The Cost of Curbing Externalities with Market Power: Alcohol Regulations and Tax Alternatives

Christopher Conlon

Department of Economics
NYU Stern School of Business and NBER
cconlon@stern.nyu.edu

Nirupama Rao

Business Economics and Public Policy
University of Michigan Ross School of Business
nirurao@umich.edu

June 2025*

Abstract
Products with negative externalities are often subject to regulations that limit competition. The single-product case may suggest that it is irrelevant for aggregate welfare whether output is restricted via corrective taxes or limiting competition. However, when products are differentiated, curbing consumption through market power can be costly. Firms with market power may not only reduce total quantity, but distort the purchase decisions of inframarginal consumers. We examine a common regulation known as post-and-hold (PH) used by a dozen states for the sale of alcoholic beverages. Theoretically, PH eliminates competitive incentives among wholesalers selling identical products. We assemble unique data on distilled spirits from Connecticut, including matched manufacturer and wholesaler prices, to evaluate the welfare consequences of PH. For similar levels of ethanol consumption, PH leads to substantially lower consumer welfare (and government revenue) compared to simple taxes, because it distorts consumption choices away from high-quality/premium brands and towards low-quality brands. Replacing PH with volumetric or ethanol-based taxes could reduce consumption by 10-11% without reducing consumer surplus, and roughly triple tax revenues.

Keywords: Excise Tax, Pigouvian Tax, Tax Efficiency, Regulation, Vertical Restraints.

JEL Classification Numbers: H21, H23, H25, L12, L13, L42, L51, L66, L81, K21.

*This replaces a previous draft circulated under the title "The Price of Liquor is Too Damn High". The authors thank seminar participants at Stanford, Columbia, Yale, Ohio State, Harvard, Kellogg, Wharton, NYU Law, NYU Stern, Drexel, Boston College, Dartmouth, INSEAD, UBC, Wisconsin, Cornell; and conference participants at Utah WBEC, FTC Microeconomics, and IIOC. We have benefited from many useful discussions from: Michael Riordan, Kate Ho, John Asker, Katja Seim, JF Houde, Alan Auerbach, Daniel Shaviro, Robin Lee, Chris Nosko, Juan Carlos Suaŕez Serrato, Rob Porter, Judy Chevalier, Jacob Burgdorf, Julie Mortimer and Adam Cole. Any remaining errors are our own. Researchers' own analyses calculated (or derived) based in part on data from Nielsen Consumer LLC and marketing databases provided through the NielsenIQ Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the NielsenIQ data are those of the researchers and do not reflect the views of NielsenIQ. NielsenIQ is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

1. Introduction

The manufacture, distribution, and selling of alcoholic beverages are big business in the United States, with sales exceeding $250 billion in 2022. Alcohol markets are also subject to an unusual degree of government intervention. Federal, state, and even local governments levy excise taxes on alcohol, raising more than $18.3 billion annually. Beyond industry-specific taxation, the sale and distribution of alcohol are also tightly regulated at the state and federal levels. A common state regulation is *post-and-hold* (PH), which governs wholesale alcohol pricing in 12 states — more than a third of states where alcohol is not sold by a state-run monopoly. These regulations discourage competition among wholesalers, leading to higher prices and lower output.

The Connecticut PH law we examine requires wholesalers to "post" a uniform price schedule to a state regulator, and then "hold" that price schedule for 30 days. All licensed retailers in the state may purchase at the posted price. Prior to sales taking place, wholesalers are offered a four day "lookback" period during which they are allowed to match but not undercut competitor prices. Theoretically, we show that PH softens competition and facilitates supra-competitive pricing in the wholesale market. Even when wholesalers offer identical products, the unique iterated weak dominant Nash equilibria of the PH pricing game leads to wholesale prices as high as a single-product monopolist would charge. Empirically, we show that PH leads to unambiguously higher prices, particularly for higher cost or inelastically demanded (higher-quality) products, and that if PH were replaced with simple tax instruments, the state could both reduce alcohol consumption and increase consumer surplus.

Understanding these policies is particularly relevant now, as Courts of Appeals are split on whether PH laws constitute a violation of the Sherman Act.[1] Proponents of PH have long argued that requiring wholesalers to commit to publicly posted prices prevents price discrimination and protects small retailers, similar to the aims of the Robinson-Patman Act. In 2022, the Federal Trade Commission (FTC), the Department of Justice Antitrust Division (DOJ), and the US Treasury Department (TTB) issued a joint report on competition in alcoholic beverage distribution that included a section largely critical of PH policies because they restrict competition and lead to higher prices (U.S. Department of Treasury, 2022). However, in 2024, the FTC reversed course and filed a Robinson-Patman suit against alcohol wholesaler Southern Glazer's, alleging harms to small retailers due to quantity discounts for large chains (the kind of behavior PH restricts). In a number of speeches, FTC Commissioner Bedoya has said "there is not one empirical analysis showing that Robinson-Patman actually raised consumer prices."[2] One interpretation of PH is that it provides a mechanism to implement the ban on wholesale price discrimination enshrined in the

[1] Courts have found that laws similar to PH violated the Sherman Act in: California (1980), Massachusetts (1998), Maryland (2004), Washington (2008); and upheld them in New York (1984) and Connecticut (2019).

[2] See `https://www.ftc.gov/system/files/ftc_gov/pdf/returning_to_fairness_prepared_remarks_commissioner_alvaro_bedoya.pdf`

Robinson-Patman Act, so that all retailers face uniform but elevated prices.[3] Our less charitable interpretation is that PH provides a mechanism for price coordination among wholesalers and is an example of regulatory capture. In Connecticut, spirits wholesalers spend nearly twice as much on state-level campaign contributions as wholesalers in California, a state with more than 10 times the population, but relatively competitive distribution. In fact, only Texas saw more political spending by spirits wholesalers than the state we study.[4]

At first glance, outsourcing price increases to private firms might seem like an attractive way to limit alcohol consumption and the associated negative externalities. Intuition from the single-product case suggests that it is irrelevant from a total welfare perspective whether we restrict supply via a Pigouvian tax or through increased market power (perhaps from lax merger approval, weaker antitrust enforcement, or market designs like PH).[5] Indeed, this argument is made by proponents of the "Green Antitrust" movement for allowing consolidation (and sometimes coordination) among fossil fuel companies, and restricting "excessive competition" has been a key feature of market design in the legalization of marijuana.[6] The interaction of market power and taxes is also a concern in attempts to address the "internalities" of sugar-sweetened beverages (Allcott et al., 2019; Dubois et al., 2020; O'Connell and Smith, 2024).

However, the intuition from the single-product case fails when products are differentiated. Put simply, we can think about alcoholic beverages as a bundle of two characteristics: ethanol and branding/quality. For example, the cheapest plastic bottle vodka and the most expensive Scotch might contain equal amounts of ethanol but differ vastly when it comes to consumer perceptions of quality or willingness to pay, which are often captured by differences in upstream prices or marginal costs. A social planner concerned only with limiting the negative externalities might levy a Pigouvian tax on ethanol alone. In a multi-product setting we show that the marginal external damage depends not just on the ethanol content of the product, but whether consumers are likely to substitute towards higher- or lower-ethanol alternatives. A firm with market power recognizes that if consumers value both characteristics, it is optimal to "tax" both characteristics proportional to the elasticity of demand, leading to higher prices on products that consumers value for non-ethanol attributes. Market power may lead not only to markups on premium products that are *too high* but markups on low-end products that are *too low*. This problem becomes particularly acute in markets like distilled spirits where costs or product quality are highly dispersed.

This means that consumers who substitute from premium products to inexpensive ones due to

[3] In Appendix E, we show that PH states have fewer retail stores and lower employment, suggesting it does little to benefit small retailers.

[4] Campaign contributions are the authors' tabulations of data from `https://www.followthemoney.org` .

[5] Levy et al. (2021) discuss public health externalities regarding the FTC investigation into the merger of cigarette maker Altria and leading e-cigarette (vape) manufacturer Juul.

[6] See Hollenbeck and Giroldo (2022); Thomas (2019) on entry restrictions in marijuana markets; Hollenbeck and Uetake (2021) on the interaction between taxes and market power in marijuana; and Hansen et al. (2020) for analysis of a (Pigouvian) "potency tax". For Green Antitrust see Kingston (2011) and Linklaters (2020) in favor and Schinkel and Treuren (2020) against.

PH prices may consume similar amounts of ethanol but be worse off. This allows combinations of simple sales and volumetric taxes to yield a triple dividend: higher consumer surplus, lower alcohol consumption, and more tax revenue. We show that even a single ethanol tax can maintain the same aggregate ethanol consumption as PH while increasing consumer surplus by more than 11%. Consumer surplus gains stem from flattening the difference between price and marginal cost across products with the same ethanol content, allowing consumers to shift away from low-priced value brands (plastic bottle Vodka) and towards premium products (Smirnoff Vodka), leaving many significantly better off, especially the higher income households who bought premium products in the first place.

To assess the welfare implications of PH and tax alternatives, we assemble new, unique data from the Connecticut Department of Consumer Protection and private data sources. These data track the monthly prices of spirits products at the manufacturer, wholesaler, and retailer level, and quarterly shipments from manufacturers to wholesalers in Connecticut from August 2007 to June 2013. Using these data, we show that retail spirits prices are higher in Connecticut than elsewhere, particularly for premium products, and that spirits consumption in Connecticut is skewed towards "lower-end" products despite it being one of the wealthiest states in the country. Following wholesale prices at the product level over time reveals that wholesalers price in parallel with little to no price dispersion, as we would expect given the incentives created by PH.

Combining the price and quantity data, we estimate a model of demand for spirits at the wholesale level that allows for correlated preferences among product categories such as gin or vodka, and heterogeneous preferences over price, package size, and overall demand that vary with income. In addition to matching aggregate purchases, we also match moments based on observed wholesaler markups and individual purchases by income. Our estimates show that the least-expensive products, which are consumed more heavily by lower-income households, feature both more elastic demands and more substitution to the outside option, making them attractive targets for reducing ethanol consumption. Unfortunately, firms with market power set the lowest markups on these products.

We assume that in the absence of the PH system, the wholesale tier would become perfectly competitive, allowing us to evaluate the welfare effects of alternative tax regimes using our demand estimates. We consider several counterfactual policies: an *ad valorem* sales tax, an ethanol tax similar to that used by the U.S. federal government, a volumetric tax (the most common state-level approach), and a minimum price per unit of ethanol.

These counterfactuals make clear that PH imposes steep welfare costs by distorting inframarginal purchase decisions. For instance, the state could reduce ethanol consumption by nearly 13% without decreasing consumer surplus by switching from PH to an ethanol tax. Meanwhile, tax revenue from alcohol would nearly triple. Scaling these gains across other PH states would imply approximately $1 billion in additional revenue nationally —accompanied by higher consumer surplus. Ethanol price floors could reduce ethanol consumption by even more (up to 25%) without sacrificing consumer surplus, though they are less effective for raising revenue.

To evaluate the trade-offs between competing policy goals, we define the frontiers of the consumer surplus vs. tax revenue and consumer surplus vs. ethanol consumption (negative externality) trade-offs. We find that conventional *ad valorem* taxes perform reasonably well in terms of revenue efficiency — they lie close to the surplus-revenue frontier — but tend to result in higher levels of ethanol consumption than Ramsey-style (product-specific) tax schedules. Conversely, a minimum price per unit of ethanol performs well in maximizing consumer surplus per unit of ethanol consumed, though it is relatively weak as a revenue-generating instrument. This helps explain the use of ethanol price floors in policy settings like Scotland (Griffith et al., 2022).

We assess several alternative modeling assumptions. One concern is that wholesale distribution is not costless. However, we find that allowing wholesalers to incur marginal costs of \$1/L or even \$2/L — relative to observed average price-cost margins of around \$3/L — has limited impact on welfare estimates. While additional distribution costs would slightly reduce the tax revenue gains from reform, they do not significantly alter the main welfare comparisons. (Note that a per-liter tax and a per-liter distribution cost are essentially isomorphic; welfare gains arise because firms do *not* price to the own-price elasticity and marginal cost.) Another concern is that profit-maximizing manufacturers might respond to a more competitive wholesale tier by raising prices, thus "undoing" some of the gains from improved market structure — a point also raised by Miravete et al. (2020). We find that while such price adjustments can raise manufacturer profits by up to 30%, they have only modest effects on tax revenue and welfare outcomes.

While it may not be surprising that replacing PH and its idiosyncratic incentives with well-designed taxes leads to efficiency gains, our analysis also offers broader insights into the use of market power as a tool for addressing externalities. Much like a sales tax, all firms with market power tend to impose higher (additive) markups on higher-quality products — those with higher marginal costs and less elastic demand. As we show, market power can serve as a "second-best" instrument for correcting externalities only if marginal external damage is positively correlated with marginal cost. When this condition fails, even simple excise taxes can outperform market power by better aligning prices with average marginal external harm, while avoiding distortions in relative prices. This issue is especially pronounced in markets such as distilled spirits, where product quality and marginal costs vary widely but are not systematically related to external damages. For example, Grey Goose costs nearly seven times as much as the cheapest vodka, yet there is little reason to believe it generates proportionally more harm. In such settings, firms with market power will tend to overprice premium products and underprice cheaper ones relative to a social planner's preferences. We expect similar distortions to arise in other markets with highly dispersed costs or product quality, such as legalized marijuana, where relying on restricted competition or firm pricing decisions may similarly fail to align private incentives with public policy goals.

2. Alcohol Regulations and Taxes in the US

2.1. State regulations regarding alcoholic beverages

While the federal government imposes substantial taxes on alcoholic beverages, the regulation of alcoholic beverage markets is almost wholly the purview of state governments.[7] Nearly all states that allow alcohol to be sold by private firms have instituted a *three-tier* system of distribution, in which the manufacture, distribution, and sale of alcoholic beverages are vertically separated by law. A common feature of nearly all systems is that retail firms (bars, restaurants, supermarkets, and liquor stores) must purchase alcoholic beverages from an in-state wholesaler.

In 18 states, known as *control states*, the state directly operates the wholesale distribution or retail tier and, in some cases, does both. In some control states, the state monopoly applies to all alcoholic beverages; in others, it applies to distilled spirits, not wine or beer.[8] Recent empirical work has focused on these control states and on understanding the behavior and welfare consequences of state-run monopolies. Seim and Waldfogel (2013) show that Pennsylvania locates more stores in rural areas and fewer stores in urban areas than a profit-maximizing firm would choose. Miravete et al. (2018) show Pennsylvania's policy of setting a uniform markup (of over 50%) on all products is set above the revenue-maximizing level, while Miravete et al. (2020) compare the uniform markup to product-specific taxes. Other studies have examined how both quantity and prices rose when Washington State privatized its state monopoly. Different authors have offered competing explanations: Illanes and Moshary (2020) explain this phenomenon with increases in product variety, while Seo (2019) focuses on increased convenience and one-stop shopping.

The majority of states are like Connecticut, where private businesses own and operate the wholesale and retail tiers. The three-tier system in *license states* prohibits manufacturers and distillers from selling directly to retailers. These license states often have regulations that restrict not only cross-tier ownership and cross-state shipping, but also a variety of other practices.[9] For example, welfare effects of both exclusive territories and exclusive dealing in the beer industry have been studied in Sass and Saurman (1993); Sass (2005); Asker (2016). The three tier system prohibits the kinds of vertical integration and arrangements found in other work on vertical integration (and anticompetitive harms) in beverage distribution (Luco and Marshall, 2020, 2021) as well as the wholesale price discrimination studied in Villas-Boas (2009).

What differentiates spirits wholesaling from beer distribution, at least in Connecticut, is that

[7] The 21st Amendment ended Prohibition by turning the power to regulate the import, distribution, and transportation of alcoholic beverages within their borders over to the states, largely exempting their regulations from scrutiny under the Commerce and the Import-Export Clauses of the U.S. Constitution. Since then, numerous Supreme Court cases have eroded state control over alcohol policy, as the Court has held that state control of alcohol is subject to federal power under the Commerce Clause, the First Amendment, and the Supremacy Clause, among others.

[8] A few control states, for example, Maine and Vermont, maintain a state monopoly on the distribution and sale of spirits but contract with private firms for retail operations (including pricing).

[9] License states may also impose other restrictions, such as which days alcoholic beverages can be sold; whether supermarkets can sell spirits, wine, or beer; and the number of retail licenses a single chain retailer can hold.

it involves a substantial amount of *common agency.* As many as four statewide wholesalers often sell the same product. Wholesalers distribute products from multiple competing distillers/manufacturers and do not divide markets geographically. Also, spirits wholesalers in Connecticut (and many other states) have a "duty to deal" and must supply all licensed retailers at posted prices. In other words, the market structure bears many of the hallmarks of competition, but the market outcomes in Connecticut under PH appear anything but competitive.

In Connecticut, under PH, manufacturers and wholesalers are prohibited from offering quantity discounts, and must charge the same prices to all purchasers. This is implemented by requiring manufacturers and wholesalers to provide the regulator with a price list for the following period (usually a month). In Connecticut, prices must be posted by the 12th day of the preceding month, and cannot be changed until the next posting period. However, some PH states, including Connecticut, also allow a *lookback* period, during which prices can be amended —but only downwards, and not below the lowest competitor price for the same item from the initial round. During this period, wholesale firms are able to observe the prices of all competitors. In Connecticut, the lookback period lasts for four business days after prices are posted. Many states, including Connecticut, also employ a formula that maps posted wholesale prices onto minimum retail prices. This limits retailers from pricing below cost (with limited exceptions to clear excess inventory).[10]

2.2. Legal Environment of Post-and-Hold

The legal status of PH laws has been challenged in several court cases, with different circuit courts drawing different conclusions as to whether §1 of the Sherman Act preempts state alcohol-pricing statutes under the 21st Amendment. In a landmark Supreme Court case, *California Retail Liquor Dealers Ass'n v. Midcal Aluminum, Inc* (1980), the court ruled that the wholesale pricing system in California was in violation of the Sherman Act. The California system at the time resembled PH, but with the additional restriction that retail prices were effectively set via a resale price maintenance agreement by wholesale distributors.[11] The court's ruling established a two-part test for determining when state actions are immune to federal preemption: 1. a law must clearly articulate a valid *state interest* (such as temperance) 2. the policy must be *actively supervised* by the state. A second California case also went to the Supreme Court and further clarified that the state action immunity did not apply if the statute led to a *per se* violation of the Sherman Act (such as collusion, market division, or refusals to deal) (See *Rice v. Norman Williams Co (1982)*).

Subsequent rulings in other courts have also struck down PH provisions as violations of the Sherman Act. In *Canterbury Liquors & Pantry v. Sullivan* (1998), the district court ruled Mas-

[10]There is a long history of policymakers being concerned about retailers using alcoholic beverages as "loss leaders." Some states allow a limited number of "post-offs," in which retailers can price below the most recent wholesale price in order to clear inventory. See `https://www.cga.ct.gov/2000/rpt/2000-R-0175.htm` for a list of various state regulations.

[11]It is worth pointing out that prior to the *Leegin* decision in 2007, minimum resale price maintenance was a *per se* violation in the United States.

sachusetts's post-and-hold scheme was a violation of §1 of the Sherman Act on summary judgment. In Maryland, the Fourth Circuit ruled in favor of a large liquor retailer in *TFWS v. Schaefer et al.* (2004, final appeal 2009), ending the state's PH system and ban on volume discounts. The Ninth Circuit's appellate decision in *Costco v. Maleng (2008)* affirmed that Washington state's "post-and-hold scheme is a hybrid restraint of trade that is not saved by the state immunity doctrine of the Twenty-first Amendment."

In contrast, the Second Circuit (which comprises Connecticut, New York, and Vermont) has twice upheld PH laws, with both decisions focusing on the *lack of coordination* required to establish a §1 collusion case (and thus a *per se* violation under *Rice v. Norman Williams*). Writing for the majority in *Battipaglia v. New York State Liquor Authority* (1984), Judge Henry Friendly found "New York wholesalers can fulfill all of their obligations under the statute without either conspiring to fix prices or engaging in 'conscious parallel' pricing. So, even more clearly, the New York law does not place 'irresistible pressure on a private party to violate the antitrust laws in order to comply' with it."

More recently, the Second Circuit's majority opinion in *Connecticut Fine Wine and Spirits, LLC v. Seagull* (2019) focused similarly on the lack of communication between wholesalers:

> Nothing about this arrangement requires, anticipates, or incents communication or collaboration among the competing wholesalers. Quite to the contrary: A post-and-hold law like Connecticut's leaves a wholesaler little reason to make contact with a competitor. The separate, unilateral acts by each wholesaler of posting and matching instead are what gives rise to any synchronicity of pricing.

The Second Circuit's dissenting opinion sharply criticized the majority's reasoning:[12]

> allow[ing] de facto state-sanctioned cartels of alcohol wholesalers to impose artificially high prices on consumers and retailers across all three states in our Circuit...The problem with Connecticut's law is not that it affirmatively compels wholesalers to collude in order to fix prices, but that it provides no incentive – or ability – for wholesalers to compete on price.

As we illustrate with our theoretical model in Section 3, both parties are partially correct. Connecticut's PH system leads to supra-competitive wholesale prices in a one-shot game via unilateral incentives, without requiring any communication or repeated cooperation among the parties.

These disparate circuit court rulings leave PH laws fully legal in some parts of the United States but prohibited elsewhere. The circuit split opens the door for the Supreme Court to resolve the issue, and highlights the importance of understanding the impact of PH laws on pricing behavior and welfare.

[12] We should disclose that we were not engaged or compensated by any parties in the Connecticut case (or any other case). However, previous versions of this paper were cited by the briefs of several parties, including the theoretical result that PH could lead to prices as high as the collusive prices in a static unilateral effects framework.

2.3. Taxes on Distilled Spirits

Federal, state, and even some municipal governments levy their own excise taxes on distilled spirits. The overwhelming majority of these taxes take the form of specific taxes, which are a fixed dollar amount per unit (either volume or alcohol content), though, in most states, the general sales tax also applies to alcohol purchases.[13]

Federal taxes are remitted by the distiller/manufacturer or upon import.[14] At the federal level, distilled spirits are generally taxed at \$13.50 per proof-gallon, where a proof-gallon is one liquid gallon that is 50 percent alcohol. Most spirits are bottled at 80-proof or 40% alcohol by volume (ABV), and incur \$2.85/L in federal taxes. Flavored spirits (generally 60-proof) incur lower taxes, and overproof spirits (often over 100-proof) pay higher taxes per liter.

Most state excise taxes, on the other hand, are volumetric, meaning they do not vary by alcohol content, and are remitted by the wholesaler. Connecticut's specific tax on spirits was raised from \$4.50 per gallon (\$1.18 per liter) to \$5.40 per gallon (\$1.42 per liter) on July 1, 2011, and again to \$5.93 per gallon (\$1.56 per liter) on October 1, 2019. We use the timing of the tax increase as an instrument in our analysis. Like most states, Connecticut includes alcohol products in its general retail sales tax base. Connecticut also increased its general sales tax rate from 6% to 6.35% when it raised its alcohol excise tax in 2011.

As a share of the overall retail price, these excise taxes can be large, particularly for the least expensive products. For example, a 1.75L bottle of 80-proof vodka in Connecticut (after 2011) includes \$7.48 in combined state and federal taxes. At the low end of the spectrum, a 1.75L plastic bottle of *Dubra Vodka* (one of the best-selling and least expensive products) typically sells for \$11.99 at retail; taxes therefore account for greater than 60% of the price. On the other end of the spectrum, a 750mL bottle of premium vodka (*Grey Goose* or *Belvedere*) or Scotch whisky (*Johnnie Walker Black*) might retail for over \$40, of which only \$3.21 (about 8%) would go to taxes.

3. Theoretical Analysis

Below we present a theoretical model that shows that the post-and-hold system functions like a "price matching game." This eliminates the incentive to cut prices to increase market share. Even when multiple firms sell identical products, the iterated weak-dominant strategy is to set the monopoly price and then match any competitor's price in the second stage. This will lead to higher prices compared to competitive wholesale markets. We consider both a simple single-product example in Section 3.1 and also a more realistic example with multi-product firms in Section 3.2. To understand the efficacy of the PH system, we compare it to the social planner's problem in

[13]This applies largely to license states. In control states, it is hard to differentiate retailer markups from *ad valorem* taxes.

[14]Imported spirits may also be subjected to additional *ad valorem* tariffs. In October 2019, President Trump imposed a 25% tariff on Scotch Whisky imports, which was later suspended for five years in June 2021 by the Biden administration.

Section 3.3 and highlight some important differences.

3.1. PH with a Single Homogenous Good

In order to understand the mechanisms of PH, we consider a two-stage game designed to resemble the PH process in Connecticut described in Section 2 and begin with a single homogeneous good.

In the first stage, each wholesaler $f \in \mathcal{F}$ simultaneously submits a uniform price p_f^0 to the regulator. In the second stage, with common knowledge of competitor prices, firms are allowed to revise their prices with two caveats: (a) prices can only be revised downward from the first-stage price $p_f \leq p_f^0$, and (b) prices cannot be revised below the lowest competitor's price for that item $p_f \geq \underline{p^0} = \min_g\{p_g^0\}$. Only after this second stage are sales realized.

$$p_f^* = \underset{p_f \in [\underline{p^0}, p_f^0]}{\arg\max}\, \pi_f = (p_f - mc_f) \cdot q_f(p_f, p_{-f})$$

We assume that the overall demand is given by $Q(P)$, where P is the "market price", and that firms charging the "market price" split the demand proportionally a la Bertrand into shares γ_f :[15]

$$q_f(p_f, p_{-f}) = \begin{cases} 0 & \text{if } p_f > \min_g p_g; \\ \gamma_f \cdot Q(P) & \text{if } p_f = \min_g p_g. \end{cases} \tag{1}$$

A dominant strategy in the second stage is to match the lowest price in the first stage $\underline{p_0}$ as long as it is above the marginal cost $p_f^* = \max\{mc_f, \underline{p^0}\}$. Given the dominant strategy in the first stage, there exists a continuum of symmetric subgame-perfect Nash equilibria (SPNE) between marginal cost and firm f's monopoly price p_f^m: $\sigma(p_f^0, p_f) = \Big([mc_f, p_f^m], \max\{mc_f, \underline{p^0}\}\Big)$.[16]

To illustrate, if all firms set $p_f^0 = mc$ in the first stage, then all firms will match this in the second stage, resulting in $p_f^* = mc$. Upward deviations in the first stage do not change $\underline{p_0}$, and thus have no effect, while downward deviations result in negative profits. Likewise, if all firms set the monopoly price $p_f^0 = p_f^m$, then upward deviations cannot result in higher profits, while downward deviations will reduce the profits of all firms but without increasing the share of firm f. All intermediate prices are also a SPNE, following the monopoly logic.

This symmetric SPNE prediction is unhelpful because it fails to rule out any price between marginal cost and monopoly, however, nearly all refinements (Pareto dominance, trembling hand, iterated weak-dominance, etc.) will select the equilibrium where each firm submits the price they would charge as monopolist p_f^m, while in the second stage competitors price match the lowest price $\underline{p_0}$. Because the monopoly price maximizes profits in a one-shot game, our analysis is not extended

[15]This generalizes the usual Bertrand assumption of splitting ties equally, but does not affect firm strategies as long as $\gamma_f \perp p_f$. For example: $\gamma = [\frac{1}{3}, \frac{1}{3}, \frac{1}{3}]$ or $\gamma = [\frac{1}{2}, \frac{1}{4}, \frac{1}{4}]$.

[16]We consider asymmetric costs and asymmetric strategies in Appendix A. These involve checking for "limit prices" that can be ruled out when costs are not "too dispersed".

to repeated games.[17] We establish the uniqueness under iterated weak dominance in Proposition 1 and provide extensions to the asymmetric case in Appendix A.

Proposition 1. *In the absence of limit pricing (or under symmetric marginal costs $mc^f = mc \; \forall f$), the unique equilibrium of the single-period game under iterated weak dominance is the monopoly price: $\sigma(p^0_f, p_f) = (p^m_f, \underline{p^0})$ where $\underline{p_0} = \min_f p^0_f$. (Proof in Appendix A.1.1).*

3.2. PH with Heterogeneous Costs and Multiproduct Firms

In our data, multiple wholesalers each sell multiple (often identical) products (e.g. three firms sell *Smirnoff Vodka 750mL*). In a typical Bertrand environment, this would lead to marginal cost pricing. However, we show that under the rules of PH the equilibrium prices are significantly higher.

Consider a multi-product wholesale firm $f \in \mathcal{F}$, which chooses prices for all products they sell $j \in \mathcal{J}_f$. Following the single-product example in Section 3.1, an iterated weak-dominant strategy is for f to set the initial price p^f_j as if it could do so unilaterally and then simply to match the lowest competitor price on that product in the second stage (assuming that it exceeds marginal cost). Our challenge is to characterize the equilibrium of *second-stage prices* when we do not necessarily observe *first-stage prices*.

As before, firm f's sales of product j are given by $q^f_j(\mathbf{P}) = \gamma^f_j \cdot Q_j(\mathbf{P})$, where $Q_j(\mathbf{P})$ represents the total demand for product j, and $\mathbf{P}$ represents the vector of the lowest (second-stage) prices for each product available in that period.[18] We write the profits of firm f (if all sellers with $\gamma^f_j > 0$ charge the "market price" P_j) as:

$$\pi_f(\mathbf{P}) = \gamma^f_j \cdot Q_j(\mathbf{P}) \cdot (P_j - mc^f_j) + \sum_{k \in \mathcal{J}_f \setminus \{j\}} \gamma^f_k \cdot Q_k(\mathbf{P}) \cdot (P_k - mc^f_k) \qquad (2)$$

If each firm f could unilaterally set the price P_j, the first order condition of (2) with respect to P_j, and divided by $\gamma^f_j > 0$ is given by (3), which we re-write in terms of marginal revenue, and

[17] A more challenging extension would be to think about a different game where prices are locked in for 30 days at a time, but firms do not have a "lookback period" because the monopoly price need not be the unique iterated weak-dominant equilibrium of the one-shot game.

[18] We continue to assume that firms which set $p^f_j > \underline{p^0_j}$ sell zero units. The substantive restriction is that γ^f_j is constant and does not depend on prices. In practice, this allows us to observe γ^f_j as the fraction of shipments of product j that go to wholesaler f each year. Allowing for differentiation among wholesalers could further soften price competition, though in practice we observe little to no price dispersion among second-stage prices, which would make estimating such differentiation difficult.

marginal costs:[19]

$$\left[Q_j + \frac{\partial Q_j}{\partial P_j}(P_j - mc_j^f)\right] + \sum_{k\in\mathcal{J}_f\setminus\{j\}} \frac{\gamma_k^f}{\gamma_j^f}\cdot\left[\frac{\partial Q_k}{\partial P_j}(P_k - mc_k^f)\right] \geq 0, \tag{3}$$

$$P_j\left(1 + 1/\epsilon_{jj}(\mathbf{P})\right) \leq mc_j^f + \sum_{k\in\mathcal{J}_f\setminus\{j\}} \frac{\gamma_k^f}{\gamma_j^f}\cdot D_{j\to k}(\mathbf{P})\cdot(P_k - mc_k^f). \tag{4}$$

All firms should choose initial prices so that marginal revenue (the left-hand side of (4)) equals marginal cost (the right-hand side of (4)). Because firms will have different marginal costs, this means that the lowest-cost firm will set the lowest initial price. This implies that in *second-stage* prices, all firms except the lowest-cost/price firm will match this price and have marginal revenue *below* marginal cost. We use the relationship in (4) to identify the firm with the lowest cost, and write (4) as a system of equations in final prices $\mathbf{p}$ rather than inequalities:

$$\kappa_{jk} \equiv \frac{\gamma_k^f}{\gamma_j^f},\ \text{such that } f = \arg\min_{f':\gamma_j^{f'}>0}\left[mc_j^{f'} + \sum_{k\in\mathcal{J}_{f'}\setminus\{j\}} \frac{\gamma_k^{f'}}{\gamma_j^{f'}}\cdot D_{j\to k}(\mathbf{p})\cdot(p_k - mc_k)\right], \tag{5}$$

$$p_j = \underbrace{\frac{1}{1 + 1/\epsilon_{jj}(\mathbf{p})}}_{=\mu_j^{PH}(\mathbf{p})}\cdot\left[mc_j + \underbrace{\sum_{k\neq j}\kappa_{jk}\cdot D_{j\to k}(\mathbf{p})\cdot(p_k - mc_k)}_{=UPP_j(\kappa)}\right]. \tag{6}$$

We express second-stage prices p_j in terms of: an inverse elasticity markup $\mu_j^{PH}(\mathbf{p}) = \frac{1}{1+1/\epsilon_{jj}(\mathbf{p})}$; a marginal cost mc_j; and the opportunity cost $UPP_j(\kappa)$ which depends on the diversion ratio from j to k and the price cost margin of substitutes $(p_k - mc_k)$.

What distinguishes (6) from the typical multi-product Bertrand oligopoly pricing equation is the presence of κ_{jk}. In the Appendix (A.1), we show how to write the markup in matrix form as $\boldsymbol{\eta} \equiv (\mathbf{p} - \mathbf{mc}) = \left(\mathcal{H}_{PH}(\kappa)\odot\Delta(\mathbf{p})\right)^{-1}\mathbf{q}(\mathbf{p})$ where $\Delta(\mathbf{p})$ are demand derivatives and $\mathcal{H}(\kappa)$ is the ownership matrix. In the usual setting, the elements of $\mathcal{H}_{jk} = 1$ if the products share an owner, and $\mathcal{H}_{jk} = 0$ otherwise. Under PH, the ownership matrix has entries $\kappa_{jk} = \gamma_k^f/\gamma_j^f$ which depend on the relative importance of products to the lowest opportunity cost firm f from (5). This means that when multiple firms sell j and only one firm sells k, we can have $\kappa_{jk} > 1$ and/or $\kappa_{kj} < 1$. When the lowest-cost seller of j does not sell k, this implies $\kappa_{jk} = 0$ (but not necessarily $\kappa_{kj} = 0$). Overall this

[19] Here $D_{j\to k} = \frac{\partial Q_k}{\partial P_j}/\left|\frac{\partial Q_j}{\partial P_j}\right|$ represents the *diversion ratio* from good j to good k, or the fraction of consumers who switch to k when they leave good j in response to a price change. Also notice that if a firm reduced its price in the first stage to $p_j^{f'}$ so that $\underline{p}^0 \leq p_j^{f'} < p_j^f$, this would have no effect on the market price in the second stage. This is the non-uniqueness of subgame perfect equilibria in Section 3.1, whereas the second-stage equilibrium is unique as long as the price-setting firm for each product j doesn't play a weakly-dominated strategy.

implies a level of competition somewhat greater than if a single firm controlled the entire wholesale market, but significantly lower than what we would expect when multiple wholesalers offer identical products.

In our empirical example, we observe (or can estimate) all of the objects in the bracketed expression from (5) and thus can determine which firm f is the "price setter" for product j. In our application, this is determined by differences in the γ_j^f terms, because $mc_j = p_j^m + \tau_j^v + \mathrm{w}_j$ does not vary by firm. Manufacturers (Bacardi, Diageo, etc.) are required to charge all wholesalers an identical price p_j^m, and excise taxes τ_j^v vary with volume but not with the identity of the wholesaler. Thus, the only substantive assumption is that w_j, the additional costs incurred by the wholesaler do not vary between firms.[20]

3.3. Comparison to the Social Planner's Problem

To understand the potential inefficiencies of the PH system, we compare the pricing equation in (6) to a social planner's problem. Under this benchmark, the planner faces the same multi-product demand system $\mathbf{q}(\mathbf{p})$ and is able to set the price of each product by choosing $\mathbf{p}$. The planner faces competing objectives: maximize social surplus, limit the external harm from ethanol consumption, and raise revenue.

The planner maximizes the difference between consumer surplus $CS(\mathbf{q})$ and total cost $C(\mathbf{q})$, and must deliver a minimum level of revenue $\mathbf{p} \cdot \mathbf{q} - C(\mathbf{q}) \geq \overline{R}$, which can be thought of as either the amount of variable profit required to sustain the industry or as tax revenue to be used for other government objectives (or any combination of the two). The planner also faces a cap on the acceptable amount of external damage from alcohol consumption $E(\mathbf{q}) \leq \overline{E}$. We make a common (though by no means necessary) assumption that the externality is *atmospheric*, or that it depends only on total ethanol consumption and not the source of the ethanol nor the identity of the consumer, such that $E(\mathbf{q}) = \mathbf{e} \cdot \mathbf{q}$, where e_j is the ethanol content of product j.[21]

A planner that can freely choose the price vector $\mathbf{p}$ that solves a "Ramsey problem" and is described by the Lagrangian with multipliers λ_e, and λ_r representing the shadow value of an additional unit of ethanol, and revenue, respectively. This gives a solution similar to the multiproduct

[20]This rules out both returns to scale, and possible geographic differences arising from transportation costs. This seems reasonable because Connecticut is a small state and most of the wholesalers are located within a very small geographic region near the center of the state. Allowing for some homogeneous (across firms and products) wholesaler cost $\mathrm{w} \in \{0, 0.5, 1, \ldots\}$ is a straightforward extension that we will consider later.

[21]This assumption would be violated if for example, if tequila generates more externalities per unit of ethanol than vodka or if 1750mL bottles generate more externalities *per liter* than 750mL bottles. This could be the case if problem drinkers preferred particular sources of ethanol. Recent work by Griffith et al. (2019) models different taxes across broad categories: beer, wine, spirits, etc. and an externality that is convex in individual consumption to capture the possibility that heavy drinkers generate more external damage.

monopoly problem:[22]

$$\max_{\mathbf{p}} CS(\mathbf{q}) - C(\mathbf{q}) + \lambda_r(\mathbf{p}\cdot\mathbf{q} - C(\mathbf{q}) - \overline{R}) - \lambda_e(\mathbf{e}\cdot\mathbf{q} - \overline{E}), \tag{7}$$

$$p_j = \underbrace{\frac{1}{1-\theta/\left|\epsilon_{jj}(\mathbf{p})\right|}}_{=\mu_j(\theta)} \left[mc_j + \underbrace{\sum_{k\neq j} D_{j\to k}(\mathbf{p})\cdot(p_k - mc_k)}_{=UPP_j} + \frac{\lambda_e}{1+\lambda_r} D^e_{j\to 0}(\mathbf{p}) \right]. \tag{8}$$

As in (6), (8) relates the optimal price p_j to: the (inverse) elasticity of demand ϵ_{jj}, the marginal cost of production mc_j, and to the "opportunity cost" that arises from multi-product pricing. Because the planner internalizes all cross-price effects, it is as if $\kappa_{jk} = 1$ for all pairs of products. Meanwhile, $\theta = \frac{\lambda_r}{1+\lambda_r}$ functions like a "conduct parameter" with $\theta = 0$ giving the perfectly competitive outcome and $\theta = 1$ corresponding to the monopoly problem (putting all the weight on revenue and ignoring both consumer surplus and the externality) for the inverse-elasticity markup $\mu_j(\theta) = [1-\theta/\left|\epsilon_{jj}\right|]^{-1}$.

The last term in brackets augments the marginal cost to capture how aggregate ethanol consumption declines as we raise the price of j:

$$D^e_{j\to 0}(\mathbf{p}) = e_j - \sum_{k\neq j} D_{j\to k}(\mathbf{p})\cdot e_k - D_{j\to 0}(\mathbf{p})\cdot e_0. \tag{9}$$

If products have similar levels of external damage $e_k \approx e$ for all k, then we can simplify (9) further as $D^e_{j\to 0} \approx (e - e_0)\cdot D_{j\to 0}$.[23] For this reason we label $D^e_{j\to 0}$ "diversion away from ethanol." This will be larger when the product contains more ethanol e_j, or when the ethanol content of the outside option e_0 is smaller (beer and wine vs water and soft drinks). This means that the planner should treat products that are more substitutable for the outside option (or lower ethanol options) as if they have higher marginal costs and set higher prices, but "taxing" products where consumers substitute to alternatives with similar (or greater) external damage will reduce welfare without reducing the externality.

There are three main differences between the planner's problem in (8) and the PH problem in (6). The first is that under PH the external damage is ignored ($\lambda_e = 0$). The second is that the planner fully internalizes all cross-product effects by setting $\kappa_{jk} = 1$ for all pairs of products. Both tend to lead to *lower prices* under PH. However, the third difference leads to *higher prices* under PH through the inverse elasticity markup $\mu_j(\theta)$. The PH markup sets $\theta = 1$, while the planner's $\theta \in [0, 1]$ is determined by the revenue target $\overline{R}$. In net, and compared to the planner's problem from (8), PH should lead to higher markups on lower effective marginal costs. This should mean

[22] See the derivation in Appendix A.2. We don't claim originality, but we could not find (8) and (9) in the literature. The most similar expressions we could find were in Sandmo (1975); Oum and Tretheway (1988).

[23] For example, most distilled spirits are 40% alcohol by volume (ABV) or 80 proof, see Table 1.

higher prices on products with higher values of mc_j, and lower prices on products with lower values of mc_j.

This creates the ambiguity that motivates our empirical exercise; PH can lead to higher prices for some products and lower prices for others compared to the planner, depending on the relevant values of $(\kappa_{jk}, \lambda_r, \lambda_e)$ and the demand elasticities. Because we do not observe the planner's weights (λ_r, λ_e), we instead focus on tax alternatives which dominate the PH outcome on all three dimensions (consumer surplus, revenue raised, and ethanol consumption). In Appendices A.2.1 and A.2.2 we provide a more detailed comparison of the PH, planner, and tax problems that we revisit in our empirical results.

4. Data and Some Descriptive Evidence

In this section, we present several stylized facts and patterns in the data consistent with the theory in Section 3. We show that: (1) Prices are higher in PH states than in other license states. When comparing Connecticut (our PH state) and Massachusetts (a nearby non-PH license state): (2) prices are higher in Connecticut; (3) relative prices are higher for "premium" products; (4) relative sales are lower for "premium" products. Finally, (5) when multiple wholesalers offer a product in Connecticut, prices move largely in lockstep.[24] *To simplify comparisons, we convert all prices and quantities to per-liter equivalents throughout this article.*

4.1. Cross State Evidence from Retail Prices

Our first set of stylized facts comes from the NielsenIQ Retail Scanner Dataset (through the Kilts Center at Chicago Booth) from 2013 (the last year in our administrative data). These data report weekly unit sales and total revenue for each product (a unique UPC) for a set of retail stores that voluntarily share their data with NielsenIQ.

To compare prices across states, we construct a price index using the 250 best-selling products weighted by q_j^{US}, the 2013 NielsenIQ national retail sales (in liters) for product j, where the price p_j^x is the 2013 revenue in state x for product j divided by its total sales (in liters). The goal is to construct a price index for a representative "liter of spirits":[25]

$$\text{Index}(x) = \frac{\sum_{j=1}^{250} p_j^r(x) \cdot q_j^{US}}{\sum_{j=1}^{250} q_j^{US}} \tag{10}$$

[24] Appendix E also extends panel data analysis by Cooper and Wright (2012) to show that aggregate sales of alcoholic beverages and employment in the retail alcoholic beverage sector are lower under PH.

[25] While coverage across states in the NielsenIQ data for supermarkets is excellent, coverage for distilled spirits varies. Some control state monopolies don't share data with NielsenIQ at all. In some license states (such as California), supermarkets are allowed to sell distilled spirits, leading to good coverage, while in others, only standalone liquor stores can sell spirits (including New York, New Jersey, and Connecticut), leading to moderate coverage. Other license states, such as Rhode Island and Delaware, where NielsenIQ records fewer than 1,000 sales, are excluded from the analysis.

Figure 2 plots the price index for control states, license states with PH, and license states without PH regulations. The dark bars on the left indicate the state excise tax burden in each license state. We do not separate out excise taxes for control states.

Figure 2 illustrates two key facts. First, PH states feature some of the highest prices. In fact, PH states outrank nearly all other license states, except Texas, which has an unusual market structure (though not PH). Second, price differences are not fully explained by differences in tax rates. Despite high prices, PH states have fairly typical tax burdens, ranking roughly in the middle of the distribution of taxes.

A simple way to think about what would happen if we eliminated PH in Connecticut would be to consider another license state as a counterfactual. For example, Illinois has prices that are approximately \$3 per liter lower, while having tax rates that are roughly double those we see in Connecticut. A more obvious comparison for Connecticut is the neighboring state of Massachusetts, which eliminated PH in 1998.[26] The two states are demographically similar[27], and are likely to have similar local wages and transportation costs. Moreover, much of Connecticut is in the shared Hartford-CT/Springfield-MA metro area, so we might expect that preferences for distilled spirits might be similar in the two states.[28] However, as Figure 2 suggests, prices are around \$1.90 per liter lower in Massachusetts, while excise taxes are only \$0.35 per liter lower.

In Figure 3, we plot the average retail price per liter in 2013 in Connecticut against the average retail price per liter in Massachusetts for each vodka brand in the NielsenIQ data. We focus on vodka because it represents around 45% of the sales volume in each state. If the prices were identical in both states, all points would lie along the 45-degree line. Instead, prices in Connecticut generally exceed prices in Massachusetts. Moreover, the price premium is larger for more expensive products. We can see that budget brand *Popov* is priced similarly in the two states. Meanwhile, *Smirnoff*, the most popular brand, is subject to a substantial Connecticut premium, and *Belvedere*, a high-end brand, is subject to an even greater premium. The best-fit line, $P_{CT} = 0.723 + 1.073 \cdot P_{MA}$, indicates

[26] There has been some confusion in the literature as to whether Massachusetts is a PH state. Cooper and Wright (2012) report that Massachusetts ended PH in 1998 while Saffer and Gehrsitz (2016) draw their data regarding PH laws from the NIAAA catalog of wholesale pricing restrictions (`https://alcoholpolicy.niaaa.nih.gov/apis-policy-topics/wholesale-pricing-practices-and-restrictions/3`) which describes Massachusetts as a PH state. To clarify the status of the PH statute in Massachusetts, we contacted the Massachusetts Alcoholic Beverage Control Commission. The General Counsel of the Massachusetts Alcoholic Beverages Control Commission explained that "The US District Court ruled the post-and-hold provision to be unconstitutional, so while it remains 'on the books,' it is not enforced so licensees do not need to post and hold (although they are still required to post prices). The case on point is Canterbury Liquors & Pantry v. Sullivan, 16 F.Supp.2d 41 (D.Mass.1998), as well as a Massachusetts Appeals Court case recognizing the District Court's ruling [in] Whitehall Company Limited v. Merrimack Valley Distributing Co., 56 Mass. App. Ct. 853 (2002)." As such, we follow Cooper and Wright (2012) and treat Massachusetts as a non-PH state after 1998.

[27] Tabulations of American Community Survey data reported in Appendix E indicate that in 2010 both states have identical shares of female respondents (52%) and very similar racial composition, with 78% identifying as white in Connecticut and 81% in Massachusetts. Average household income is slightly higher in Connecticut (\$89,500 vs. \$83,200), while mean age (measured for the 18+ population) is 47.7 in Connecticut compared to 46.8 in Massachusetts. Years of schooling are nearly identical, averaging 13.3 in Connecticut and 13.4 in Massachusetts.

[28] In addition to Hartford/Springfield, parts of Connecticut get media from New York City and/or Boston.

that on average, Connecticut consumers pay approximately \$1.45 per liter more for discount vodka, \$2.18 per liter more for mid-tier vodka, and \$3.64 per liter more for premium vodka.[29] (Recall, the tax difference is only a uniform \$0.35 per liter).

The fact that the prices are relatively higher on premium products in Connecticut can also distort *which products are purchased* in each state. In Figure 4, we categorize vodkas based on the *national average* price per liter and plot the share of sales (by volume) in each price band for each state for the two most popular sizes (750mL: upper panel, 1.75L: lower panel).[30] The idea is that the national average price captures some objective measure of "quality." The purchase patterns in Figure 4 show that relative to their Massachusetts neighbors, consumers in Connecticut are more likely to purchase products from the two lowest "quality" groups, and much less likely to purchase products from the two highest "quality" groups. Again, this is purely descriptive, and it may be that preferences for vodka in plastic bottles are higher and preferences for *Grey Goose* are lower in Connecticut for other idiosyncratic reasons.[31] We provide an alternative comparison based on CDFs in Appendix D.1.

4.2. Administrative Data from Connecticut

Our main dataset is meant to capture the universe of sales of distilled spirits at the *wholesale level* in the state of Connecticut from July 2007 to July 2013. This dataset has been collected and compiled by us (the authors), and has not been previously analyzed.

The first data source is the monthly price postings from Connecticut's Department of Consumer Protection (DCP). The PH system necessitates that all *wholesalers* submit a full price list for all products they sell.[32] A similar regulation requires that the manufacturers/distillers (firms like Bacardi, Diageo, Jim Beam, etc.) post prices each month.[33] This means that we see monthly product-level pricing for both the *manufacturer* tier and the *wholesale* tier.

There are several challenges related to data construction. The first is that the format of price filings is irregular. While some firms provide spreadsheets, others provide printed PDF reports and many provide scans of faxed-in price lists. The second challenge is that a single product such as *Johnnie Walker Red* is sold by a single manufacturer (Diageo) but by up to four wholesalers, and there is no product identifier that links the product between the manufacturer and the wholesaler or between wholesalers. This means that all product matching and assignment to a unique product identifier must be done primarily by hand. A third challenge is that reporting of product flavors can be inconsistent: we might see shipments of one flavor (Cherry), but price postings only for

[29] Here, we've defined discount, mid-tier, and premium vodkas as \$10, \$20, \$40 per liter, respectively.

[30] A similar pattern holds for 1L bottles, but we omit these from the figure since 1L bottles account for only 4% of retail liquor store sales in Massachusetts and Connecticut.

[31] Another possibility is that consumers in Connecticut drive to Massachusetts to save \$9 on Grey Goose, but not to save \$0.50 on Popov.

[32] Recall that the legislation prohibits quantity discounts, so firms are restricted to *uniform* prices.

[33] Each manufacturer/distiller is the sole seller for each of the brands they produce, unlike the wholesale tier, where multiple firms offer identical brands.

another flavor (Orange). As different flavors are often priced identically within a brand-size-proof combination, we consolidate multiple flavors so that *750mL Smirnoff Vodka (Flavored)* is a unique product, but "Orange" or "Cherry" is not.

The most serious limitation of the price-posting data is that we usually don't observe both: (a) the initial price postings; and (b) the amended or revised price postings. In some cases, we see only the initial price posting and some handwritten (or faxed) amendments. In others, we see only initial price postings and do not know whether prices were amended or not. And finally, in other cases, we observe only a list of amendments to prices and no price postings at all.[34] One limitation is that we don't have both sets of prices, which we would need to analyze the two stages of the price-posting process. We can offer anecdotal evidence that when firms amend prices, they are required to list the competitor whose price they "match," and this is verified by the DCP. However, an advantage of the model in Section 3.2 is that it requires only a single wholesale price for each product (the second stage price of the lowest opportunity-cost wholesaler).

The second data source tracks shipments of distilled spirits from manufacturers/distillers/importers to wholesalers. These data were obtained from the Distilled Spirits Council of the United States (DISCUS). The DISCUS data track shipments from member manufacturers – generally the largest distillers – to wholesalers for each product and constitute 78% of total shipments of distilled spirits (by volume) in the state of Connecticut.[35]

A key aspect of the DISCUS data is that it contains all shipments (of covered brands) to the state of Connecticut. This includes products that ultimately end up in bars and restaurants, as well as those sold in retail liquor stores. Another advantage of the DISCUS data is that we see total shipments not only by product, but also to each wholesaler. This lets us estimate the γ_j^f parameters from our theoretical model in (5) directly from the shipment data. The primary disadvantage is that for less popular products, shipments can be lumpy, with only a handful of shipments per year. For this reason, we focus our analysis primarily at the *quarterly* level of observation, and for the least popular products (one shipment per year or less, around 6% of total sales), we have to apply some further smoothing. For the 21.9% of products not included in our DISCUS sample, rather than exclude them from the analysis, we impute shipments using the NielsenIQ Retail Scanner data totals from 34 stores in Connecticut. We describe the construction of the quantity data in detail in the Data Appendix. Figure 1 plots ethanol consumption data from spirits using our constructed data and reports from the tax receipts data. The series show two important features: 1) ethanol consumption from spirits is rising over our sample period; and 2) our data both mirror

[34] We discuss the data cleaning in detail in our Data Appendix. When in doubt, we treat price postings as if they are (second stage) "as amended.' Some manufacturers tend to post only the prices of products whose prices changed from the previous month, which requires some care in constructing the full sequence of prices.

[35] DISCUS members include: Bacardi U.S.A., Inc., Beam Inc., Brown-Forman Corporation, Campari America, Constellation Brands, Inc., Diageo, Florida Caribbean Distillers, Luxco, Inc., Moet Hennessy USA, Patron Spirits Company, Pernod Ricard USA, Remy Cointreau USA, Inc., Sidney Frank Importing Co., Inc., and Suntory USA Inc. Some of the largest non-DISCUS members include: Heaven Hill Distillery and Ketel One Vodka.

this consumption pattern and generally fit consumption levels.[36]

Table 1 reports summary statistics for the 735 products we use in our analysis by category and bottle size.[37] Products are brand-flavor-proof-size combinations, such as *Smirnoff Vodka 750mL* or *Tanqueray Gin 1L*. Vodka is the largest product category, accounting for 208 products, and 44.8% of all spirits liters sold. While a plurality of products are 750mL, it is 1.75L products that account for 56.7% of sales volume. Most products are 80-proof (40% alcohol by volume), and as such proof averages near 80 for most categories and bottle sizes, with some exceptions.[38]

Table 1 also reports the average price and average price-cost margin (or additive markup) net of any taxes: $(p_j - mc_j)$ at each tier of the distribution chain (manufacturer/distiller, wholesaler, and retailer). Table 2 reports similar information except with the average Lerner markup $L = \frac{p_j - mc_j}{p_j}$ instead of the additive markup, and broken out by manufacturer/distiller instead of by size and category. To produce meaningful summary measures across differently-sized products, product prices and margins are measured in per-liter terms, and all means are weighted by liters sold. Our data are unusual because we observe prices at the manufacturer p^m, wholesaler p^w, and retailer p^r level, as well as the excise taxes τ_j paid by wholesalers. This means we directly observe input costs except at the manufacturer level.[39] The largest manufacturer, Diageo, sells 155 products and accounts for 32.7% of sales by volume, and enjoys the highest Lerner markups (around 30% on average).

The most important takeaway from Tables 1 and 2 is that the wholesale tier is significantly more profitable than other tiers. A "typical" product retails for slightly less than \$20 per liter, with a breakdown of: \$3.97/L of wholesaler margin, \$3.07/L of manufacturer margin, \$2.71/L of retailer margin, and \$1.42 in state and \$2.85 in federal taxes.[40] Moreover, prices (per liter) and markups tend to be higher (for all tiers) on 750mL products than on the less expensive (per unit) 1.75L products. Our counterfactuals will focus on the case where we remove PH, make the wholesale tier more competitive, and instead use taxes to constrain ethanol consumption and address negative externalities.

[36] Appendix C.2 provides additional comparisons over time.

[37] We restrict the sample using the following criteria: (1) we only consider the 750 best-selling products (99.9% of sales volume); (2) only products whose average wholesale price is below \$60/L (mostly excluding rare Scotch Whisky); (3) we exclude Cordials and Liqueurs (e.g. Triple Sec, Baileys, Kahlua) which are generally 20% alcohol by volume or less and possibly complements rather than substitutes for distilled spirits; (4) we exclude Cognacs (e.g. Hennesey and Courvoisier) because these products contain vintage/age statements and are nearly impossible to match across data sources.

[38] Some popular gins and imported Scotch Whiskies are over-proof. Most flavored vodkas are 60-proof, and flavored rums can be as low as 42-proof (e.g. *Malibu Coconut Rum*).

[39] Manufacturer marginal costs are backed out of the first order conditions using our demand estimates and following the procedure described in Appendix A.3. Retail prices come from the NielsenIQ Scanner Dataset for Connecticut and are available only for select retail stores, while manufacturer and wholesaler prices are statewide.

[40] The remainder being production costs.

4.3. Wholesaler Pricing Behavior

Our main focus is the pricing behavior and market power of the wholesale tier. Up to four wholesalers sell identical products at identical prices, yet each charges a substantial markup above the manufacturer's price. There are several innocuous possibilities, including the fact that wholesaling activities are costly to produce; maybe they provide valuable ancillary services; or perhaps wholesale firms are substantially differentiated in ways we cannot observe.

In the PH system, the wholesale price is given by $p_j^w = p_j^m + \tau_v + \mathrm{w}_j + \eta_j$, where p_j^m is the manufacturer's price, τ_v is the existing volumetric tax, and w_j is any additional cost incurred by the wholesaler.[41] We plot the wholesale prices and manufacturer prices in Figure 5. Rather than plot the 45-degree line, we plot the zero markup ($\eta_j = 0$) line: $p_j^w = p_j^m + \tau_v$, so that the markup η_j is the vertical distance from the line. We see that (after accounting for taxes) wholesaler markups η_j are larger on more expensive products, with products like *Grey Goose* and *Johnnie Walker Black* having very high price-cost margins. Large markups η_j are not exclusive to the most expensive products; the mid-priced product *Smirnoff Vodka* (the overall best-seller) also has a high margin, though other popular yet inexpensive products such as *Dubra Vodka* have small markups.

Figure 6 tracks the wholesale (case) prices of up to four different wholesale firms in addition to the manufacturer price for four popular spirits products: Stolichnaya Vodka (1000mL), Tullamore Dew Irish Whiskey (1750mL), Dewars White Label (750mL), and Johnnie Walker Black (1750mL). These products were selected because they are sold by different sets of wholesalers and have a lot of price variation over time. For each product, the prices set by the different wholesalers move in near lockstep with one another. This is true for a wide variety of products sold by different sets of wholesalers. While one innocuous explanation might be that this synchronous movement simply reflects changes in input prices, the manufacturer prices plotted alongside the wholesale prices do not support this reasoning. Manufacturer prices change only rarely, while wholesale prices move more frequently and together. Instead, it appears that wholesalers are pricing in parallel, which is consistent with the price-matching incentives created by PH.

Occasional price deviations are short-lived and typically involve only one of three to four wholesalers selling a product. Across the entire dataset, the average price-dispersion within a product-month across wholesalers is 2.7% and 77% of observations have no price dispersion.[42] When this happens, we interpret these deviations as cases where initial price postings rather than "amended" price postings are recorded. In the case of Johnnie Walker Black (1.75L), monthly wholesale prices oscillate between two price points, but for Eder, we observe only the higher of the two prices.[43]

[41] For now we will assume $\mathrm{w}_j = 0$, or that no additional costs are incurred by the wholesaler, we will relax this in our empirical exercise, but it ends up not affecting the economics of the problem. What really matters is that w_j isn't significantly different across products (ie: if Whiskey needs special transport relative to Vodka, etc.)

[42] We calculate this as $\frac{p_{max} - p_{min}}{\frac{1}{2}(p_{max} + p_{min})}$.

[43] For some of the months in question, we are able to confirm the dates on the submitted prices are consistent with "initial" prices. A likely explanation is that "amended" prices were submitted via fax or were not properly digitized.

When there is dispersion in our posted wholesale prices, in nearly 80% of such cases, the cause is a single wholesaler with a higher recorded price. For this reason, in our econometric model, we assume that all firms play the iterated weak dominant strategy of matching the lower price in the second stage. Moreover, because our econometric model looks at prices and quantities at the quarterly level rather than the monthly level, we end up smoothing out some of this higher frequency price variation.

5. Empirical Analysis

5.1. Econometric Model of Demand and Supply

We start with a model of simultaneous supply and demand at the wholesale level, and augment this with additional "micro moments" in spirit of Berry et al. (2004). As noted in Section 4.2, we observe monthly shipments from manufacturers to wholesalers and use these data to construct our measure of quantity sold. Because these shipments can be lumpy, we aggregate our data at the *quarter level* which eliminates some of the higher frequency price variation observed in Figure 6. This leaves us with 24 periods from 2007 Q3 - 2013 Q2. Throughout our analysis, our unit of observation is a *liter* and our prices are *per liter*. A key feature of the PH game from Section 3.2 is that multiple wholesalers offer identical products (e.g., Smirnoff Flavored Vodka 1750 mL at 60-proof) at identical prices.

In each quarter t, consumer i chooses whether to purchase a single product j (a brand-size-proof combination) or the outside option ($j = 0$) where utility is given by:

$$u_{ijt} = \beta_i^0 + \beta_{it}\,\mathrm{x}_{jt} - \alpha_{it}\,p_{jt}^w + \xi_{b(j)} + \xi_t + \Delta\xi_{jt} + \varepsilon_{ijt}(\rho). \tag{11}$$

Here, p_{jt}^w represents the minimum wholesale per-liter price, and x_{jt} represents additional product characteristics (bottle size and/or proof), and $(\xi_{b(j)}, \xi_t)$ represent brand and time fixed-effects respectively.[44] We specify the idiosyncratic shock $\varepsilon_{ijt}(\rho)$ so that demand follows the random coefficients nested logit model (Brenkers and Verboven, 2006; Grigolon and Verboven, 2014) to allow for more substitution within a product category (Gin, Rum, Tequila, North American Whiskey, Irish/Scotch Whisky, and Vodka) than across categories.[45]

We define the individual purchase probability $\sigma_{ijt} = \mathbb{P}(u_{ijt} > u_{ij't} \mid \alpha_i, \beta_i)$ for all $j \neq j'$. The share of j in market t is given by:

$$\sigma_{jt}(\boldsymbol{\xi_t}; \theta_2) = \int \sigma_{ijt}(\alpha_{it}, \beta_{it}, \boldsymbol{\xi_t}; \theta_2) f(\alpha_{it}, \beta_{it} \mid \mathrm{y}_{it}, \theta_2)\, h(\mathrm{y}_{it})\, \partial\alpha_i\, \partial\beta_i\, \partial\mathrm{y}_{it}. \tag{12}$$

We allow consumers to have heterogeneous preferences for product characteristics that are de-

[44] We let $\boldsymbol{\xi_t} = [\xi_t, \xi_{b(j)}, \Delta\xi_{jt}\, \forall j]$, the stacked vector of fixed effects and demand shocks for each market t.

[45] We've used this in earlier drafts, and it has become a popular choice for distilled spirits (Miravete et al., 2018) and beer (Miller and Weinberg, 2017).

termined by observed demographics (income y_{it} discretized into five "quintile" bins $\mathcal{I}_k$)[46] and unobserved (normally distributed) characteristics. We also require that the price coefficient α_i is log-normally distributed, so that all consumers have downward-sloping demand curves, and allow for correlation between random coefficients with variance-covariance matrix Σ:

$$\begin{pmatrix} \ln \alpha_i \\ \beta_i^0 \\ \beta_i \end{pmatrix} \sim \mathcal{N}\left(\sum_{k=1}^{5} \Pi_k \cdot \mathbb{I}\{y_i \in \mathcal{I}_k\}, \Sigma \right). \tag{13}$$

The literature establishes a unique inverse (Berry et al., 1995; Berry and Haile, 2014) for the vector $\boldsymbol{\xi_t}$, which sets predicted shares in (12) equal to the observed shares $\boldsymbol{\mathcal{S}_t} = \boldsymbol{\sigma_t}(\boldsymbol{\xi_t}, \theta_2)$. With instruments z_{jt}^D, we define conditional moment restrictions:[47]

$$\Delta\xi_{jt} = \sigma_{jt}^{-1}\left(\boldsymbol{\mathcal{S}_t}, \theta_2\right) - \xi_t - \xi_{b(j)} - \overline{\beta}\, \mathrm{x}_{jt}, \quad \mathbb{E}[\Delta\xi_{jt} \mid z_{jt}^D] = 0. \tag{14}$$

$$\omega_{jt} = p_{jt}^w - \eta_{jt}\left(\mathcal{H}_t(\kappa), \theta_2\right) - p_{jt}^m - \tau_{jt}^v - \mathrm{w}_{jt}, \quad \mathbb{E}[\omega_{jt} \mid z_{jt}^S] = 0. \tag{15}$$

The second set of conditional moment restrictions (15) matches the observed wholesaler markups $(p_{jt}^w - mc_{jt}^w)$ to those predicted under the PH model $\eta_{jt}\left(\mathcal{H}_t(\kappa), \theta_2\right)$ from (6).[48] A useful feature of our data is that we observe most components of marginal costs $mc_{jt}^w = p_{jt}^m + \tau_{jt}^v + \mathrm{w}_{jt} + \omega_{jt}$ including manufacturer prices p_{jt}^m and volumetric taxes τ_{jt}^v. Any additional marginal costs incurred by the wholesaler are captured by the unobserved cost shock ω_{jt} and its mean w_{jt}. Rather than parameterize these remaining costs w_{jt} with covariates like $\beta\, x_{jt}$, we assume $\mathrm{w}_{jt} = \mathrm{w}$ (a constant) is common across products and wholesalers, and consider robustness to several values $\{0, \$0.50, \$1.00, \$1.50, \ldots\}$ Higher wholesaling costs lead to slightly more elastic demand (and worse fit), while attempts at estimating w as wholesaler specific constants (or a single constant) result in small negative values. In our main specification, we assume $\mathrm{w} = 0$ and consider other values in robustness tests.[49]

[46] The income cutoffs for the bins are $\mathcal{I} = \{0, \$25K, \$45K, \$70K, \$100K\}$ and coincide with the NielsenIQ panelist data (with some consolidation at lower income levels).

[47] We partition parameters: $\theta = [\theta_1, \theta_2]$ where $\theta_1 = [\boldsymbol{\xi_b}, \xi_t, \overline{\beta}]$ and $\theta_2 = [\Pi, \Sigma, \rho]$ so that the average price effect is in θ_2 (and $\sigma_j^{-1}(\cdot)$).

[48] We give a non-original derivation for additive markups η_{jt} (in matrix form) from (6) in (A.1) and (A.2).

[49] It is reasonable to ask what additional marginal costs are meant to be captured in w_{jt}. Connecticut wholesalers all charge a regulated per mile, per-delivery fee on top of the wholesale price. If set appropriately, it suggests that the distance between wholesalers and retail customers is unlikely to be an additional cost component (and why setting $\mathrm{w} = 0$ instead of some other positive value might be reasonable if the fee covers the marginal cost of delivery). All but one of the largest wholesalers are located within 40 miles of New Haven, so cross-wholesaler differences in delivery fees are likely small anyway, and it seems hard to justify significant differences in storage and transportation costs across products. See `https://www.cga.ct.gov/2023/pub/chap_545.htm` (Sec 30-64a) for a description of delivery charges.

5.2. Micro Moments

We augment the moments from (14) and (15) with additional "micro moments" (Petrin, 2002; Berry et al., 2004). We use NielsenIQ panelist data to construct a set of moments year by year (2007-2013) that correspond to product characteristic-demographic interaction parameters Π from (13). We also construct another set of moments inspired by Atalay et al. (2025) meant to target the nesting parameter ρ by matching the rate at which repeated consumer purchases occur within the same category $\mathcal{J}_c$: (Vodka, Rum, Gin, North American Whiskey, UK Whiskey, Tequila).

Our notation and `PyBLP` implementation closely follow Conlon and Gortmaker (2025). We restrict the NielsenIQ panelist sample to households residing in Connecticut who purchase at least one distilled spirits product that year and use the projection weights when aggregating across households. The first two sets of moments capture the probability that a given liter of spirits is purchased by a household from income bin $\mathcal{I}_k$ conditional on: (a) making any purchase $j \neq 0$; (b) purchasing a large size bottle $\mathrm{x}_j = 1.75L$. (c) the third set of moments matches the average *wholesale price per liter* for products purchased by households in income bin $\mathcal{I}_k$. For each household, we match the product identifier of *retail purchases* to our wholesale data and use the *corresponding wholesale price* in that quarter (because our demand model is at the wholesale level). The goal of these moments is to capture *incidence* of different policies across income groups, rather than identify income effects (e.g. if a policy raises prices of cheap vodka, which groups of consumers are likely to be affected?).[50] The full set of micro-moments are given by:

$$g_M(\theta_2) = \begin{pmatrix} \mathbb{P}\left[\mathrm{y}_i \in \mathcal{I}_k \,\middle|\, j \neq 0,\, t \in \mathcal{T}_{\text{year}}\right] \\ \mathbb{P}\left[\mathrm{y}_i \in \mathcal{I}_k \,\middle|\, \mathrm{x}_j = 1750\text{mL},\, j \neq 0,\, t \in \mathcal{T}_{\text{year}}\right] \\ \mathbb{E}\left[p^w_{jt} \,\middle|\, \mathrm{y}_i \in \mathcal{I}_k,\, j \neq 0,\, t \in \mathcal{T}_{\text{year}}\right] \\ \mathbb{P}\left[k \in \mathcal{J}_{ct} \setminus \{j\} \,\middle|\, j \in \mathcal{J}_{ct},\, t \in \mathcal{T}_{2007\text{–}2013}\right] \ \forall c \end{pmatrix}_{\text{Year}=2007,\ldots,2013} \tag{16}$$

We construct a separate set of moments for each year by aggregating across markets (quarters) within that year $t \in \mathcal{T}_{year}$ in part because the set of NielsenIQ panelists (and weights) differs by year. Variation across markets in the values of micro-moments, or variation across markets in the distribution of demographics, is generally required to separate the parameters in Π from the remaining unobserved heterogeneity in Σ (see Berry and Haile (2024); Conlon and Gortmaker (2025)). Appendix B.1.1 further details these micro moments.

Our final set of moments matches the probability that a consumer whose first choice product is from a particular category $j \in \mathcal{J}_c$ (Vodka, Gin, Rum, NA Whiskey, UK Whisky, Tequila) would select a second choice product from the same category $k \in \mathcal{J}_c$ (conditional on making a purchase). This is straightforward to construct from our demand model, and highly informative about the

[50] A potential threat is that is that the NielsenIQ panelist sample is not representative in some key way. (If NielsenIQ households shop systematically at different stores, or purchase different products in retail stores than they do in bars/restaurants).

nesting parameter ρ. To calculate the target values from the NielsenIQ Panelist data, we estimate the probability that when a consumer switches from one brand to another, both brands are in the same category. We do this by resampling the order of purchases within a household and focus only on switching between *distinct* products, which allows us to separate what might be a strong brand preference for Smirnoff (Vodka) and a large idiosyncratic ε_{ijt}, from a large value of ρ. The repeat purchase rate in the panelist data is higher than the unconditional market share of each category from Table 1, suggesting a value of ρ significantly greater than zero: Vodka (0.51 vs 0.44), Gin (0.60 vs 0.07), Rum (0.20 vs 0.175), NA Whiskey (0.26 vs 0.15), UK Whiskey (0.33 vs 0.10).[51] Appendix B.1.2 provides further implementation details.

The micro moments in (16) are not necessarily the "quasi-optimal" micro moments proposed by Conlon and Gortmaker (2025), which approximate the scores of the individual likelihood. However, they are straightforward to calculate from the NielsenIQ panelist data without many additional assumptions, and are more likely to be *compatible* with the aggregate sales data.[52]

5.3. Estimation Details

Estimation takes place in `PyBLP` (Conlon and Gortmaker, 2020) and uses the micro-moment interface developed in (Conlon and Gortmaker, 2025); we follow the recommended practices described therein whenever possible. We use all four sets of moments (demand, supply, demographic micro data, repeated purchases) in (14) to (16). We estimate the parameters of the model using two-step GMM and in the second step, update the instruments with the feasible approximation to the optimal instruments: $\mathbb{E}\left[\frac{\partial \xi_{jt}}{\partial \theta} \mid z_{jt}, \widehat{\theta}\right]$ for demand, and $\mathbb{E}\left[\frac{\partial \omega_{jt}}{\partial \theta} \mid z_{jt}, \widehat{\theta}\right]$ for supply (Chamberlain, 1987; Berry et al., 1999). When constructing the weighting matrix, we allow for correlation between the blocks of supply and demand moments, but impose independence from other blocks. For the micro-moments in (16), we treat each year of the NielsenIQ panelist data as an independent sample but allow for correlation between the moments within the same year.[53]

In order to obtain a first-stage pilot estimate for $\widehat{\theta}$ so that we can construct the weighting matrix and the approximation to the optimal instruments, we need to choose initial instruments (z_{jt}^d, z_{jt}^s). For the supply instruments, we set $z_{jt}^s = [1, p_{jt}^m, \tau_{jt}]$ using only the included regressors from (15) as instruments. We experimented with using higher-order functions of p_{jt}^m to approximate the conditional moment restriction but found that it did not matter in practice.

For the initial demand moments (14), we need to choose instruments z_{jt}^d. The obvious instruments are the excluded cost variables of (15): p_{jt}^m (the manufacturer's price); and τ_{jt} (the

[51] See the discussion of the RCNL model (and the "optimal moments") in the appendix to Conlon and Gortmaker (2025). We use the fact assignments of products to categories are known (Vodka, Rum, Gin, etc.) and use repeated purchases to measure the correlation of preferences within a category ρ. Atalay et al. (2025) instead use repeated purchase data to assign products to nests (ie: do the same households that purchase Coca-Cola also purchase Pepsi?).

[52] Conlon and Gortmaker (2025) formalize this notion of "compatibility" and illustrate how incompatible micro moments can bias parameter estimates. See Appendix B.1.1 for a detailed example.

[53] See Conlon and Gortmaker (2025) for precise details about how the standard errors and weighting matrix are adjusted.

per-liter excise tax, which increased in July 2011). In addition, we follow the recipe in Gandhi and Houde (2019) and construct instruments based on quadratic interactions of differences in exogenous product characteristics (category, size, proof, flavored) $\sum_k (\mathrm{x}_{jt} - \mathrm{x}_{kt})^2$. We also include expected wholesale prices $\mathbb{E}[p^w_{jt} \mid p^m_{jt}, \tau^v_{jt}, z_{jt}]$ among the quadratic interactions. In words, these instruments convey, "How many other 750mL flavored vodkas are available?" or "How many other similarly priced whiskeys are for sale?" These are meant to capture the changes in the "crowding" of the product space over time.[54] This procedure tends to produce a large number instruments that are highly correlated with each other, we take the first 32 principal components and use them as z^D_{jt}.

In order to compute the integral in (12), we must determine a way to approximate the joint distribution of income $h(\mathrm{y}_{it})$ and unobserved heterogeneity $f(\alpha_{it}, \beta_{it} \mid \mathrm{y}_{it}, \theta_2)$. Our main specification uses 500 quasi-random draws from the discrete distribution of income and the three-dimensional (log) normal distribution of unobserved heterogeneity. When the dimension of integration is two or less, we use a product rule of the Gauss-Hermite quadrature rule and the discrete distribution on income y_{it}. The goal of estimation is to recover the Cholesky root L such that $LL' = \Sigma$.[55] For some specifications, elements of L are zero so that Σ is no longer full rank. In such cases, we reduce both the dimension of integration and restrict L so that we obtain correct inference (avoiding the problems associated with parameters on the boundary).

Prior to estimation of θ, we estimate the discrete distribution of household income $h(\mathrm{y}_i) = \mathbb{P}(\mathrm{y}_i \in \mathcal{I}_k)$ which appears both in the calculation of market shares (12) and the micro-moments (16). We estimate $h(\mathrm{y}_{it})$ separately for each year using all NielsenIQ Panelists from the state of Connecticut (not just those purchasing spirits) using the provided projection factors. Income is recorded in ranges, and we consolidate some of the lower income bins so that our distribution is $\mathcal{I} = \{< \$25K, \$25K - \$45K, \$45K - \$70K, \$70K - \$100K, \geq \$100K\}$. Because Connecticut is a high-income state, and NielsenIQ top codes income at \$100k, 29% of households are in our top "quintile" while only 8.5% are in the bottom "quintile". As one might expect, household incomes decline during the Great Financial Crisis, then rise slowly over time.

5.4. Parameter Estimates

We report our estimated parameters for our main specification in Table 3. The parameters themselves are not easily interpretable, though we can see some obvious patterns.

The demographic interactions in Π correspond directly to the micro-moments in (16). At higher income levels, consumers become less price sensitive (as one might expect), but also the taste for all spirits β^0_i declines. This implies that higher-income consumers purchase similar quantities of alcohol as lower income households, but do so at higher prices. (Because the price sensitivity parameters are log-normal, a larger negative number indicates *less price sensitivity* so that $\alpha_i = -e^{-0.736} = -0.479$ for the lowest income group and $\alpha_i = -e^{-2.291} = -0.101$ for the highest income group.)

[54] In the data, we see more U.S. whiskey products entering and fewer flavored-rum products.

[55] These are consistent with the "best practices" for `PyBLP` in Conlon and Gortmaker (2020).

We see less of a discernible pattern for the large format size, 1750mL. Our fixed effects are at the brand level (e.g., *Smirnoff Vodka 80-Proof*), so different sizes share the same $\xi_{b(j)}$ term, but may differ in the 1750mL dummy. The omitted middle-income group \$45k-\$75k exhibits a slight preference for large bottles $\beta_i^{1750} = 0.30$, while the highest and lowest income groups exhibit a slight preference ($\beta_i^{1750} < 0$) for smaller bottles (750mL and 1L). Other characteristics such as ethanol content/proof and other demographics (age, race, etc.) were either not significant or not sufficiently captured in the Connecticut NielsenIQ Panelist data to be used in our main specification.

Even after controlling for demographics, there is substantial unobserved heterogeneity in price sensitivity α_i and the overall intercept for spirits demand β_i^0 in the variance-covariance matrix Σ. However, much like in the case of the Π parameters, we estimate a strong correlation so that most households either: (a) like alcohol but dislike price; or (b) are less price sensitive but like alcohol less. This is likely driven in part by the large number of sales concentrated at relatively low price points. Together, both (Π, Σ) lead to a strong negative correlation between (α_i, β_i^0).

Higher values of the nesting parameter ρ imply a higher probability that a consumer's second choice product will be in the same category as their first choice product. At the estimated value of $\rho = 0.27$, this means that 73% of vodka buyers would switch to another vodka. For the other categories, the model predicts: Gin (45%), Rum (56%), NA Whiskey (53%), UK Whisky (46%). Overall, these are slightly higher than target moments with the exception of the Gin category (which is 15 percentage points lower).

Our demand model is more easily understood by its economic predictions. Because we target average markups in (15), the model does an excellent job here; observed and predicted Lerner markups are $\frac{P-MC}{P} = 0.23$. The observed markups are more dispersed $IQR : (0.188, 0.276)$ compared to those predicted by the model $IQR : (0.222, 0.255)$. Part of this is the variance reduction provided by the parametric model, and part is that while we can match quarterly markups on average, we don't have features in the model (or the instrument set) beyond $\Delta\xi_{jt}$ or ω_{jt}.[56]

In Table 3, we report a median own-elasticity of -4.77 with an IQR of $(-5.07, -4.48)$ and an aggregate elasticity (to a 1% tax) of -0.53. Absent the supply restriction (15), we would find significantly *less elastic* demand curves, while increasing the marginal cost of wholesaling w_{jt} would yield *more elastic* demand curves. Understanding the welfare implications of different tax policies requires measuring whether consumers respond to higher prices by switching brands or by substituting away from spirits altogether. We estimate that the diversion ratio to the outside good $D_{j\to 0}$ averages 46%.[57] When a consumer substitutes away from a liter of product j, this reduces

[56]It is worth noting that Figure 6 reports monthly prices while our demand model is estimated using (smoother) quarterly prices and quantities.

[57]Compared to some recent IO papers that estimate demand for distilled spirits in Pennsylvania (Miravete et al., 2020, 2018), our estimates suggest somewhat larger own-price elasticities (-4.77 vs. -3.75), but much less elastic aggregate elasticities (-0.53 vs. -2.48), which implies *greater substitution between brands* and much *lower diversion to the outside good*. Regression estimates of the aggregate elasticity for spirits vary considerably (both in credibility and point estimates). On the lower end Wagenaar et al. (2009) report an elasticity of -0.29 as a result of their meta-analysis, while on the higher end Leung and Phelps (1993) report an elasticity of -1.5.

ethanol from spirits $D^e_{j\to 0}$ by between $IQR : (0.17, 0.21)$ liters (from a baseline of 0.4 liters of ethanol per liter of spirits on average). Ethanol diversion is more dispersed than the diversion to the outside good $D_{j\to 0}$ and a key input into the externality correction by the planner from (8).

Perhaps the best way to validate our demand model is to examine the predicted substitution patterns. For several top products, we compute the diversion ratio from that product to its closest substitutes and report the name, average wholesale prices, and diversion ratios in Table 4. For the most part, products appear to compete with similarly priced products within the same category (largely due to the nesting parameter ρ). For example, Dubra Vodka (1.75L), the least expensive product in our sample, appears to compete most closely with the other discount vodka brands (Popov, Sobieski, Gray's Peak, and Bellows) as well as Smirnoff vodka (a mid-range vodka and the best selling product overall). Belvedere (a super premium vodka) appears to compete with Grey Goose, Absolut, and Ketel One (all upscale vodkas). Woodford Reserve, a premium bourbon, competes primarily with Maker's Mark and Jack Daniels, the two best-selling American whiskeys. Because of the nesting structure, we see that Captain Morgan's competes primarily with Bacardi Rum, and Beefeater Gin competes largely with other gins (as well as best-selling Smirnoff vodka). We also see that consumers largely substitute from 1.75L bottles to 1.75L bottles; or 750mL and 1L bottles to 750mL and 1L bottles.

In many of our counterfactual experiments, what matters for welfare is which products have less elastic demand (and thus higher markups under PH), and which products are most substitutable away from ethanol (often to the outside option) as opposed to towards other distilled spirits. If these products coincide, then the PH system is likely to target the "correct" products with higher markups. Unfortunately, this is not the case.

Products exhibit significant heterogeneity in these key predictions depending on their wholesale prices, and we illustrate this heterogeneity in Figure 7 (around 95% of sales are for products priced under $\$33/L$ — indicating a long-tail of seldom-purchased yet expensive products). As an example, we see that for most products, lower prices are associated with more elastic demand (lower markups) and greater diversion towards the outside option. Our measure of marginal ethanol reduction is highest among the least expensive products, suggesting that attempts to reduce ethanol consumption would be most effective by targeting these products. Products around $\$20/L$ are associated with the lowest values of $D^e_{j\to 0}$, suggesting the greatest degree of substitution to other spirits (or higher alcohol content products). As we showed in Figure 5

Figure 7 also highlights how our demand model provides sufficient heterogeneity to produce substantially different predictions from the plain (IIA) logit. In that model, the own elasticities are an increasing function of prices $\epsilon_{jj} = -\alpha p_j \cdot (1 - s_j)$. Instead, we see an inverted U-shape where the lowest and highest priced products have the most elastic demands, and products priced around \$32/L have the least elastic demands (95% of all sales occur below this price).

5.5. Alternative Specifications and Discussion

While our model of consumer demand in Section 5.1 is similar to the prior literature on alcoholic beverages (Miller and Weinberg, 2017; Miravete et al., 2018, 2020), our identification strategy is somewhat different than most BLP applications. We rely more heavily on the supply restrictions from (15) to capture the price sensitivity and the micro moments from (16) to capture the heterogeneity, and less on cross-market variation in prices and product assortment.[58] By imposing moments from the supply side, we fully leverage the fact that we observe (upstream) manufacturer prices p_{jt}^m and variation in the excise tax τ_{jt}^v, but we require that firms set prices according to the iterated weak dominant strategy of the PH game in (6). This is different (but not more restrictive) from the typical assumption that firms play a static Bertrand-Nash game (such as in Berry et al. (1995, 1999)) or other alternatives such as double marginalization (Villas-Boas, 2007).[59]

In a sense, we are asking the demand system to do less than the usual BLP application, because we observe both manufacturer and wholesaler prices (and don't need to recover markups). We still use the demand estimates to understand how consumers will adjust purchase patterns under counterfactual pricing with competition among wholesalers and where different tax instruments change the relative prices.

We highlight some key features of our demand model that are important in our counterfactual welfare calculations. First, we estimate a high degree of correlation between the price sensitivity α_{it} and random coefficient on the constant β_{it}^0 that arises from both the interactions with income Π and the unobserved heterogeneity Σ. This means that some people really like spirits but really dislike prices, while others are price-insensitive but less interested in spirits. It avoids the scenario where the least price sensitive (highest income) individuals purchase all (or most) of the spirits.

The second feature is that rather than impose a parametric (monotone) relationship between price sensitivity and income, we flexibly estimate Π using a series of income bins.[60] Allowing for non-monotonicity is important because in the micro data the highest- and lowest-income households tend to purchase slightly more spirits (per capita) than middle-income households (leading to a slight U-shape).[61] One disadvantage of this sort of flexible "preference shifter" formulation is that we lose the ability to estimate interpretable income elasticities or construct Engel Curves.[62] However, our goal is not to understand how the market for distilled spirits would look under a different income distribution, but rather how different tax policies might impact the relative prices

[58]See Berry and Haile (2024) for the non-parametric treatment of identification with micro data.

[59]Conlon and Gortmaker (2020) show that strong instruments (such as p_{jt}^m) seem to perform well even when the supply model is mis-specified.

[60]Common monotone relationships include Berry et al. (1995): $\alpha_i = \alpha \cdot \log(y_i - p_j)$, Berry et al. (1999): $\alpha_i = \frac{\alpha}{y_i}$ and Nevo (2001): $\alpha_i = \alpha + \pi_0 y_i + \pi_1 y_i^2 + \sigma \nu_i$ or the Box-Cox transform proposed by Miravete et al. (2023): $\alpha_i = \alpha \cdot \frac{(y_i - p_j)^\lambda - 1}{\lambda}$.

[61]We explore this phenomenon in more detail in (Conlon et al., 2024) and show that using national data middle income households tend to consume more beer and less spirits than the highest and lowest income groups.

[62]An ongoing literature including Griffith et al. (2018); Miravete et al. (2023); Birchall et al. (2024) considers different ways to model the relationship between income and price sensitivity.

and consumption choices of different types of consumers.

The third feature is that we assume the price sensitivity α_{it} follows a log-normal distribution. This is important because it rules out upward-sloping demand curves, even for extreme draws from the distribution. In addition to the correlated random coefficients and flexible bins in Π, it also allows for a more flexible pass-through. In Table 3, we estimate the average pass-through rate to be $\rho = 1.3$ which is consistent with regression estimates from Conlon and Rao (2020), but would be ruled out under a simple logit and possibly some simpler mixed logits with quasi-linear utility (Miravete et al., 2023).[63]

In Appendix D.3, we consider the robustness of our results to two key assumptions. In Table D.2 we explore how our results vary by profiling the nesting parameter ρ and re-estimating the remaining parameters. In Table D.2, we find that larger values of ρ imply larger own-elasticities, less diversion to the outside option, and a smaller overall elasticity to a 1% tax on spirits. We also illustrate how our pseudo second-choice moments "select" a value of ρ, by matching the probability of "repeat purchases" within the same category (such as Vodka $\rightarrow$ Vodka). To test the robustness of our estimates to the specification of marginal cost, we re-estimate the model (and all counterfactuals) assuming that wholesalers incur an additional per-liter cost $\mathrm{w}_{jt} = 1$ instead of 0 in Appendix D.3.2, which leads to smaller markups and more elastic demand (but worse fit), but otherwise similar welfare implications.

We have also estimated a variety of restricted and/or expanded versions of the specification in Table 3. Excluding the Π or Σ parameters significantly worsens the overall fit of the estimates. Eliminating the nesting parameter ρ leads to predictably non-sensical (logit-like) substitution patterns. Additional parameters in Σ, such as the variance-covariance terms for the 1750mL dummy, are difficult to separate from the variance-covariance term on the constant and do not significantly improve the overall fit.

6. Welfare Under Counterfactual Policies

Proponents of PH often make one of two arguments. The first is that PH protects the profits of small retailers from competition with larger chains by requiring public posting of wholesale prices (and holding for 30 days to discourage discounts). This is the typical Robinson-Patman argument against wholesale price discrimination. In Appendix Table E.3, we show that while PH may redirect surplus from large to small retailers, states which terminate PH see growth in both the number of retail (liquor) stores and retail (liquor) employment.

The other argument in favor of PH and similar policies is that raising prices is a feature and not a bug, and by reducing competitive incentives among wholesalers, one can reduce consumption of sin goods like distilled spirits and provide a "second best" correction to a negative externality. As such, we focus our welfare analysis on how effective PH is in reducing alcohol consumption,

[63] Others have found even larger estimates for distilled spirits using different data ≈ 1.6 (Kenkel, 2005).

measured by forgone tax revenues and lost consumer surplus.

The starting point for our welfare analysis is that in the absence of PH, the wholesale tier would become perfectly competitive. One reason to use perfect competition as a benchmark is that we frequently observe multiple wholesalers distributing identical products (e.g. see *Johnnie Walker Black, 1.75L* in Figure 6). Unlike beer distribution, the market for distilled spirits in Connecticut does not have *franchise laws* which restrict wholesalers to *exclusive territories*, and all wholesalers service the entire state.[64] The second reason is that in order to quantify the welfare consequences of using market power (rather than taxation) to limit ethanol consumption, zero market power provides an obvious comparison.

We describe several simple tax instruments that we consider as alternatives to the PH system in Table 5: (a) a volumetric tax (similar to the one that Connecticut and most license states use currently); (b) an ethanol specific tax (similar to the one used by the federal government); (c) an *ad-valorem* tax (similar to the general sales tax or the fixed-markup rule used in control states like Pennsylvania (Miravete et al., 2018)); (d) a price floor per unit of ethanol (similar to that enacted in Scotland and examined by Griffith et al. (2022)). We also provide some benchmarks to illustrate the full range of potential policies: (e) the perfectly competitive price absent any taxes; (f) a profit-maximizing multi-product monopolist (similar to the privatized monopoly of Maine); and (g) a product-specific (Ramsey) tax that maximizes consumer surplus subject to either a revenue or aggregate ethanol constraint. Under each of our policy alternatives, we do not change the baseline federal excise tax (paid by manufacturers), which we include in the manufacturer price p^m_{jt}, but we do replace the existing state volumetric tax $\tau_{jt} = \$1.42/L$ with the policy alternative. In our baseline scenario, we assume that wholesalers incur no additional marginal costs ($\mathrm{w}_{jt} = 0$). Later, we allow for a \$1 per liter wholesaling cost ($\mathrm{w}_{jt} = 1$), and consider a variety of alternative costs in Appendix D.3.

$$mc_{jt} = p^m_{jt} + \underbrace{\tau_{jt}}_{=0} + \underbrace{\mathrm{w}}_{\in\{0,\frac{1}{2},1,\ldots\}} \tag{17}$$

Our baseline scenario also holds the upstream price p^m_{jt} fixed, though later, we provide additional results that allow manufacturers to re-optimize prices after the wholesaler markup is eliminated. We view these as upper and lower bounds on how manufacturers might respond. The main constraint on manufacturer price adjustment is likely that they sell the same products in neighboring states (New York, New Jersey, Connecticut, and Rhode Island) and may be limited in their ability to price discriminate.

[64] See Asker (2016) on exclusivity in beer distribution. In Connecticut, all of the major spirits wholesalers are located near one another in the center of the state. While Bertrand competition among two firms might result in marginal cost pricing, some products are sold by a single wholesaler. In a world without PH, the manufacturer could eliminate double marginalization by selling through a second wholesaler. Alternatively, wholesale markups might be eliminated if manufacturers directly supplied retailers rather than using wholesalers as intermediaries.

Because our demand estimates reflect derived demand at the *wholesale level* and abstract away from retail pricing, the area under the demand curve corresponds to the joint welfare of both retailers (bars, restaurants, and liquor stores) as well as households. The advantage is that we consider the entire market for spirits (both on- and off-premise), but the disadvantage is that we cannot separate the surplus of final consumers from small and large retailers. We compute both the aggregate impact and the impact on each of the five income "quintiles".

The primary motivation for limiting the consumption of distilled spirits is the associated negative externalities. There is little agreement on the magnitude of the externality.[65] Because our analysis focuses largely on the wholesale tier, we are limited in our ability to model *who* does the drinking. Instead, we treat the externality as if it were *atmospheric* (i.e., it depends on only the aggregate level of ethanol consumption). This would be problematic if we were concerned that there were larger negative externalities associated with drinking tequila rather than vodka, or that lost productivity was greater for households earning over \$100K.[66] Rather than take a stand on the externality, we consider three policy targets: (a) keeping ethanol consumption fixed at the existing level under PH; (b) increasing ethanol consumption by 10%; (c) reducing ethanol consumption by 10%. In our final exercise, we ask: How much can we reduce ethanol consumption without reducing consumer surplus?

6.1. Comparing Tax Instruments

Our first goal is to understand how the different tax instruments affect the relative prices of products. To do this, we eliminate the wholesaler markup and the existing volumetric tax and then find the level of each tax instrument that holds overall ethanol consumption *fixed at the PH level*. We then examine how the counterfactual prices compare to those observed under the PH system in Figure 8. Product prices that lie below the black 45-degree line become less expensive under the alternative policy than under PH, while prices above the line become more expensive under the alternative policy. We pair tax instruments with similar economic implications together in Figure 8 in order to highlight the connections between them.

The first panel of Figure 8 compares prices under volumetric and ethanol taxes with PH prices. There is little difference between taxing volume and taxing ethanol content as the bulk of products are around 80 proof (40% alcohol by volume).[67] To hold ethanol consumption fixed, these taxes effectively add a fixed $\tau^v = \$5.52$ (per liter) or $\tau^e \cdot e_j = \$13.70 \cdot e_j$ (per liter of ethanol) to each product. This leads to mostly higher prices for products less than $\$20/L$ (under PH) and lower

[65] See `https://www.ias.org.uk/wp-content/uploads/2020/12/The-costs-of-alcohol-to-society.pdf` and `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8200347`

[66] One serious concern is that the heaviest drinkers account for the bulk of the external damage (Griffith et al., 2019). Conlon et al. (2024) find that in the U.S. the heaviest drinkers are concentrated among the highest and lowest income groups.

[67] Remember we focus on base spirits (Gin, Rum, Whiskey, Vodka, Tequila) and ignore lower proof cordials and liqueurs, as well as non-spirits (fortified wines, beer, wine).

prices for products above $\$20/L$, which is "flatter" than the 45-degree line.[68] As we document in Appendix C.1, this shifts purchase away from the least expensive products (often 1.75L bottles of vodka) and towards premium products (Smirnoff Vodka, Jack Daniels, etc.).

The second panel of Figure 8 illustrates that the minimum unit price amounts to setting a price floor while otherwise pricing at marginal cost: $p^w_{jt} = \min\left\{\$18.90/L \cdot \frac{Proof_j}{80}, p^m_{jt}\right\}$. We also see that product-specific Ramsey prices designed to maximize consumer surplus for the current level of ethanol consumption (and ignoring revenue) would follow a similar strategy to the minimum unit price, effectively selling more expensive products at marginal cost $p^w_{jt} = p^m_{jt}$. At lower prices the two prices diverge with Ramsey prices incorporating marginal cost and elasticity information for some additional "slope" below the price floor. Both alternatives both lead to much lower prices on the most expensive products and significantly higher prices on products currently priced below $\$18/L$. This similarity provides some insight as to why Scotland might have enacted a minimum unit price in lieu of a tax: it is effective at curbing consumption while providing as much consumer surplus as possible (though it raises very little revenue).[69]

The final panel of Figure 8 examines a uniform sales tax. This leads to an even "steeper" relationship than PH. In part this stems from the fact that the marginal costs (manufacturer prices p^m_{jt}) at the low end of the distribution are quite low and the (eliminated) volumetric taxes $\tau^v \approx \$1.42$ are a significant component of marginal cost. Taken together, this means that raising prices to a sufficient level to curb ethanol consumption requires very high sales tax rates (72% without existing excise taxes, and 41% if we don't eliminate existing excise taxes), which exceed the typical markups under PH (around 23%), especially among high-end products.[70] Product-specific (Ramsey) taxes, which maximize consumer surplus subject to a revenue constraint (ignoring ethanol consumption), look similar to the uniform sales taxes in the middle of the price distribution, but have higher prices than PH for the least expensive products, and lower prices than PH for the most expensive products (producing a slight "S-shape").

6.2. Welfare Results

We compare PH and alternative tax policies across our three welfare measures: (a) consumer surplus; (b) revenue raised; and (c) total ethanol consumption (external damage). In Figure 9, we map out the welfare tradeoffs for all possible levels of each tax instrument and report the percentage changes in welfare measures relative to the status quo policy (PH), which we locate at the origin of the graphs. For each tax instrument, we denote the point on the curve that leaves ethanol

[68] A combined tax rate of \$8.37/L may seem high compared to the existing state tax of \$1.56/L (and federal tax of \$2.85/L). To put things in perspective, taxes on spirits in the UK are roughly twice as large at £12.65/L (or \$16.35/L) at 40% ABV.

[69] For comparison, the minimum unit price in Scotland is £26.00/L at 80 proof, which works out to over \$35/L, and much higher than the \$18.90/L floor needed to match current consumption in Connecticut.

[70] This in line with (slightly smaller) than the Pennsylvania state-run monopoly studied in Miravete et al. (2018, 2020) which charges a \$2 per bottle fee and 30% markup with an 18% sales tax for a combined markup of $p_{jt} = 1.53 \cdot (p^m_{jt} + 2.00)$

consumption unchanged ($*$), increases ethanol consumption by 10% ($+$), and decreases ethanol consumption by 10% ($\times$).

Although it is unlikely to be the preferred policy of lawmakers, if we eliminated PH and existing volumetric taxes, ethanol consumption would increase by 126%, and consumer surplus would rise by 110%. We denote this point by ($P = MC$). If we maintained the existing volumetric taxes, ethanol consumption would increase by 84% and consumer surplus by 76%, and tax revenue would rise by 83% (from additional sales). Although this would represent a dramatic increase, the predicted per capita sales of spirits would be similar to Delaware or Washington, DC (and lower than New Hampshire).

The left panel of Figure 9 considers the trade-off between overall ethanol consumption (the source of the negative externality) and consumer surplus. Here, the frontier is defined by Ramsey (ethanol), which maximizes the consumer surplus at each level of ethanol consumption by setting product-specific taxes (and ignoring tax revenue). As was the case in Figure 8, the minimum ethanol unit price is remarkably close to the frontier. The existing PH system is dominated by simple taxes on volume or ethanol content, which allow higher levels of consumer surplus at each level of ethanol consumption. However, in this sense, the PH system performs better than a uniform sales tax rate. Under a uniform sales tax, raising prices at the low end of the distribution enough to discourage consumption requires extremely high sales tax rates, leading to even higher prices at the high end of the distribution (a "steeper" curve in Figure 8). One advantage of the PH system is that profit-maximizing wholesalers can choose different markups for different products depending on their elasticities, rather than being constrained to a single tax rate.

The frontier in the right panel of Figure 9 is defined by the Ramsey (Revenue) scenario, which uses product-specific taxes to maximize consumer surplus at each level of tax revenue (while ignoring ethanol consumption). This traces out a curve from the perfectly competitive price (with no additional taxes) to the monopoly price, which achieves the highest possible revenue increase of 345% (but reduces consumer surplus by 16.9% and ethanol consumption by 10.2%). The uniform sales tax gets surprisingly close to this frontier, though it requires significantly higher levels of ethanol consumption (around 10%) to achieve similar levels of revenue and consumer surplus as the Ramsey frontier (because it sets prices that are "too steep"). As in Figure 8, taxing volume or taxing ethanol content yields nearly identical results, but would be less effective at raising revenue than sales taxes, and are thus inside the frontier. An obvious limitation of PH is that differences between wholesale prices and marginal costs are captured as wholesaler profits rather than tax revenue. Indeed, under PH, the only source of tax revenue is the \$1.42/L volumetric tax. However, even if we could extract wholesaler profits via a lump-sum tax, not only is this point (denoted by ($PH + PS$)) dominated by volumetric/ethanol taxes, but the corresponding taxes would actually reduce ethanol consumption by more than 10%. The highly similar product-specific Ramsey (Ethanol) taxes and the minimum-unit price are both less effective at raising tax revenue than $(PH + PS)$, but still raise more revenue than the existing PH system.[71]

From a policy perspective the most important case may be the set of policies that dominate PH on all three fronts: (a) lower ethanol consumption; (b) greater consumer surplus; (c) greater tax revenue. In order to focus on cases that increase social surplus, we use $(PH + PS)$ as our revenue benchmark. Most states rely on some combination of sales taxes and volumetric/excise taxes on spirits, and we calculate all combinations of tax rates that dominate PH (northwest of $PH+PS$ in both panels of Figure 9). We plot the set of tax rates that dominate PH for each year in Figure 10.[72] In 2011-2013, it is possible to dominate PH using just a volumetric tax (or a volumetric tax and the existing 6.35% sales tax rate). In earlier years, wholesale profits are higher and it is hard to generate the same amount of revenue (per capita) as could be raised by lump-sum taxation of wholesale profits without increasing the sales tax rate to around 15%. In Appendix Figure C.5, if we relax the revenue constraint to match only current tax collections (and not wholesale profits), we no longer need sales taxes in order to dominate PH. In Appendix C.2, we address the issues of "nominal taxation" raised in Seim and Thurk (2023) and illustrate in more detail how the tax rates needed to dominate PH change over time, particularly during a period where demand for spirits is growing as demonstrated in Figure 1.

6.3. Distributional Analysis and Endogenous Responses

The main takeaway from Figures 8 and 9 is that PH is dominated by simple tax instruments such as volumetric or ethanol taxes, which produce greater consumer surplus and tax revenue at lower levels of ethanol consumption. However, those alternative policies tend to produce "flatter" relationships which increase prices at the low end and reduce prices at the high end of the market. This raises potential distributional concerns, particularly since the micro-moments suggest that the least expensive products are purchased disproportionately by the lowest-income households.[73]

In Table 6, we decompose the percentage change in consumer surplus in Figure 9 for each of our five income bins. We report this for our three scenarios: (a) holding ethanol fixed at the PH level; (b) increasing ethanol consumption by 10%; (c) reducing ethanol consumption by 10%. Though we report the effects for all of the alternative tax policies, we focus our attention primarily on the volumetric and ethanol taxes. Even under a 10% reduction in overall ethanol consumption, the volumetric and ethanol taxes increase overall consumer surplus relative to PH. However, all of

[71] We (generously) assume that under the minimum-unit price, the state collects the difference between the marginal cost and the minimum unit price as revenue. In practice, collecting revenue from a minimum-unit price or a lump-sum tax on wholesalers could prove challenging, and these should be thought of as a theoretical benchmark. One approach might be to set a sufficiently large license fee for wholesalers or to auction off wholesale licenses. Assessing a tax based on the difference between the manufacturer price and some "minimum unit price" might be possible (but could likely be undone if manufacturers raised prices).

[72] Because our data start in Q3 of 2007 and end in Q2 of 2013 we define a "year" to coincide with a typical academic year instead of a calendar year.

[73] The distributional analysis is complicated by the fact that our measures of consumer welfare implicitly include retailer surplus because we model demand at the wholesale level. As such, any surplus losses and gains likely partly accrue to retail establishment owners.

the consumer surplus gains accrue to the highest income (> \$100k) group (+10.2% under ethanol taxes and +9.7% under volumetric taxes), which accounts for roughly 30% of the population of Connecticut. Meanwhile, other income groups are actually worse off.[74] We see a similar pattern if we hold ethanol fixed at existing levels, where the majority of the gains accrue to the highest income group. Under an ethanol tax, all groups are slightly better off, while under a volumetric tax, households earning below \$70,000 in income are slightly worse off.

Our welfare analysis has considered a wide range of tax rates and instruments, but thus far, we have assumed that in the absence of PH, the wholesale tier would be perfectly competitive. We focus on an ethanol tax for the remainder because it performs the best in Table 6, is already implemented by the federal government, and most directly addresses the externality associated with alcohol consumption. In Table 7, we relax perfect competition in two ways: (a) we allow for uniform $\mathrm{w}_{jt} = \$1$ per-liter marginal cost incurred by wholesalers; (b) we allow for manufacturers to adjust prices after the wholesale markups have been eliminated.[75] The first panel describes an ethanol tax that leaves total ethanol consumption unchanged from PH (as in the top panel of Table 6). In the second panel, we see how much we can reduce ethanol consumption (and associated negative externalities) without reducing consumer surplus in aggregate.

The main finding from Table 7 is that under an ethanol tax, it is possible to reduce ethanol consumption from spirits by 12.87% while increasing tax revenue by 293% and maintaining the PH level of overall consumer surplus. Allowing for an additional $\mathrm{w}_{jt} = \$1/\mathrm{L}$ marginal cost for wholesalers does not change the economics of the problem and simply reduces the tax rate from \$6.50/L to \$5.47/L (and the corresponding revenue by almost exactly \$1/L). Allowing upstream manufacturers to raise their prices significantly increases their estimated profits (an increase of 29% compared to an increase of 9%), but still allows for a nearly 12% reduction in ethanol consumption without reducing (aggregate) consumer surplus — again, this functions largely as a transfer from tax revenue to manufacturers.[76]

Middle-income households (between \$45,000−\$75,000) are nearly indifferent between an ethanol tax that holds aggregate ethanol consumption fixed at the PH level and the existing PH system.

[74] This is partly explained by the kinds of products that experience price increases and decreases when we replace PH with an ethanol tax. Appendix Table C.1 lists the 15 products that experience the largest sales gains (biggest winners) and sales losses (biggest losers) when PH is replaced with an ethanol tax. The biggest winners are a combination of popular products like Smirnoff (the best-selling product overall) and Jack Daniel Black Label, and high-end products like Grey Goose and Tullamore Dew 12-Year; these products have low own-price elasticities and were marked-up accordingly under PH. These are products, particularly the high-end products, disproportionately purchased by high-income households. These products sales rise because their prices fall substantially under the tax. The losers comprise the lowest end products, largely low-cost vodkas exemplified by Dubra Vodka, which sees a more than 60% price increase. In a sense, replacing PH with an ethanol tax effectively raises the prices of all products inferior to Smirnoff enough to effectively eliminate them as options (at similar prices nearly all consumers prefer Smirnoff). In general, large 1.75L containers see large price increases and sales declines because under PH they sold at disproportionate discounts relative to 750mL bottles.

[75] We provide a full welfare analysis with a wholesaling cost of \$1/L in Appendix D.3.2.

[76] Allowing manufacturers to re-optimize prices (against elasticities) also leads to larger increases on high-end/premium products so that the welfare gains are less concentrated among the wealthiest households.

These households tend to prefer beer and have the lowest per capita consumption of spirits (see Conlon et al. (2024)). They also serve as a constraint on policymakers, as any policy that reduces ethanol consumption relative to PH is likely to reduce the consumer surplus of this group (or the households earning less than $25,000).

One approach to addressing the distributional effects of tax alternatives to PH might be to transfer some of the additional tax revenue in order to hold harmless the lowest-income groups (such as by reducing taxes on wage income or expanding the EITC) while still reducing the overall level of ethanol consumption in Connecticut. However, such transfers may be complicated by political considerations. An additional complication is that demand for spirits is not spread uniformly across households within an income group (not all households purchase spirits), so that true "Pareto Improvements" may not be feasible (it certainly will not be feasible for households with a high idiosyncratic preference ε_{ijt} for *Dubra Vodka* which sells for less than $8/L).

A deeper question is whether lower levels of consumer surplus which arise from reduced consumption of spirits (rather than consumption of less preferred products) should be treated as a welfare loss. In the literature on sugar-sweetened beverage (SSB) taxes and "internalities," the main motivation for SSB taxes is to reduce consumption by low-income households (see Allcott et al. (2019)). Similarly, if we think that the heaviest drinkers or those who generate the most internal/external damage are more likely to seek out the least expensive forms of ethanol, then our alternative policies *understate* the reduction in external damage for a given level of aggregate ethanol consumption. A key limitation is that our top-line number represents a 12.87% reduction in *ethanol from spirits*, but some of these spirits buyers may switch to beer or wine instead of away from alcoholic beverages entirely.

6.4. Discussion: Why does PH perform so poorly?

One might have expected the post-and-hold system to perform better. Firms with market power have the ability to choose prices more flexibly than a single tax rate would allow. Moreover, they can (and do) choose prices with knowledge of own- and cross-price elasticities. Indeed, it has been known since Ramsey (1927) that there is a duality between the optimal tax problem and the multi-product monopoly problem, and that the monopolist minimizes deadweight loss for a particular level of revenue. This raises the question, how does the PH problem solved by wholesale firms differ from the social planner's problem?

Broadly speaking, under PH, wholesaler market power leads firms to set proportional markups that are *too high*, on effective marginal costs that are *too low* because they ignore the externality and do not fully incorporate all cross-price effects.[77] This ends up being particularly acute for distilled spirits because of the extreme dispersion in wholesaler marginal costs (manufacturer prices); the

[77]The cross-price effects are particular to the PH game in (6) and arise because the marginal wholesaler may not sell all competing brands, this wedge would disappear with a single monopolist. Ignoring the externality is not unique to the PH game.

manufacturer price of Dubra Vodka is around $\$4/L$, while for Grey Goose Vodka it is $\$30/L$. Both products have similar proportional wholesale markups in the data ($p_j^w/p_j^m \approx 1.4$), but in dollar terms the Grey Goose markup is \$10 larger. From the perspective of a social planner, this would only make sense if the external damage associated with Grey Goose were much larger than for Dubra. If we calculate our measure for the marginal external damage $D^e_{j\to 0}$, which captures how much a consumer substituting away from a product reduces their ethanol consumption it is actually 33% larger for Dubra (0.24 liters of ethanol versus 0.18).[78] This is a single example of a broader phenomenon: the least expensive products tend to have the largest marginal ethanol reduction $D^e_{j\to 0}$ (see Figure 7), and the lowest (dollar-value) markups in Figure 5. The typical "tagging" argument in public finance is that we would like to indirectly raise the relative prices for products with larger negative externalities (Allcott et al., 2015). However, PH effectively "tags" the most expensive products rather than the least expensive. This ends up reducing consumer surplus by too much for each unit of ethanol consumption eliminated. Simple taxes on volume or ethanol can match the average external damage without also distorting the relative prices in the wrong direction.

We can formalize intuition this by considering the difference between prices set under the PH system in (6) and the planner's problem from (8). We work out this difference in Appendix A.2.1 in (A.7):

$$\frac{p_j^{PH} - p_j(\lambda_e,\lambda_r)}{\mu_j(\theta)} = \underbrace{\left(\frac{\mu_j^{PH}}{\mu_j(\theta)} - 1\right)(mc_j + UPP_j)}_{>0} + \underbrace{\frac{\mu_j^{PH}}{\mu_j(\theta)} \cdot \Delta UPP_j^{PH}}_{<0} - \underbrace{\lambda_e^* \cdot D^e_{j\to 0}}_{<0} \tag{18}$$

The first term is unambiguously positive (and leads to higher prices under PH) because the PH markup is greater than the planner's markup when applied to the same mc_j and UPP_j terms $\mu_j^{PH} \geq \mu_j(\theta)$. As the planner places more weight on revenue $\frac{\lambda_r}{1+\lambda_r} = \theta \to 1$, this first term shrinks.

The second term is mostly negative (reducing prices under PH) because under PH wholesalers fail to incorporate all of the cross-product effects that a multi-product monopolist would. This is because $UPP_j(\kappa) = \sum_k \kappa_{jk} \cdot D_{j\to k} \cdot (p_k - mc_k)$ and the planner effectively sets $\kappa = 1$ for all pairs of products so that $\Delta UPP_j^{PH} = UPP_j^{PH}(\kappa) - UPP_j(1) \leq 0$.[79] This term is specific to the PH game in (5) and might differ or disappear in a different setting (such as Maine contracting with a private firm to be the monopoly seller of all spirits).

The third term is negative so long as the planner cares about reducing ethanol consumption $\lambda_e^* > 0$ and if raising the price of j reduces ethanol consumption $D^e_{j\to 0} > 0$. There are a few

[78] We perform similar calculations for a wider set of products and report those most impacted in Table C.1. Also note that $D^e_{j\to 0}$ is convenient because it only depends on the demand estimates and not the level of external damage/planner's weight λ_e^* which we avoid taking a stand on.

[79] In theory, it is possible that $\kappa_{jk} > 1$, and for a large close substitute $D_{j\to k}$ is large enough, then $\Delta UPP_j(\kappa)$ could be negative. In practice $\kappa_{jk} = 0$ is a significant fraction of the time, and only 2/636 products have $\Delta UPP_j < 0$.

products in our data for which $D^e_{j\to 0} < 0$ and consumers increase their ethanol consumption as they leave product j (e.g. the lowest ethanol products such as *Malibu Rum 21% ABV*).

In general, PH sets markups that are too large on effective marginal costs that are too small. For some products, the PH prices may be higher than the planner would prefer, while for others they may be lower. The first two terms in (18) push in opposite directions and can be calculated with a guess of the "conduct parameter"/revenue constraint θ; we label this sum $\text{Distortion}_j(\theta)$ because it represents the difference in pricing incentives between PH and a planner with revenue parameter θ (and $\lambda_e^* = 0$). For small values of θ, prices under PH will be higher than the planner would set, while for larger values of θ, the planner's prices approach the multi-product monopolist, and the distortion becomes negative.

The best-case scenario for PH occurs when the planner's weights (θ, λ_e^*) happen to make the expression in (18) equal to zero *on average*—that is, when the first two terms offset the third $\sum_j \text{Distortion}_j(\theta) \approx \sum_j \lambda_e^* \cdot D^e_{j\to 0}$.[80] However, there is no reason to expect that profit-maximizing firms will *accidentally* choose markups that align with the planner's weights. Yet, as we show in (A.12), it is relatively easy to match the external damage *on average* using a simple volumetric tax. To improve upon a volumetric tax, PH would need to set *relative prices* that better reflect differences in external harm—specifically, by charging higher prices on products with greater $D^e_{j\to 0}$. This approach is known in the public finance literature as "tagging." We evaluate PH's implicit tagging ability by comparing $D^e_{j\to 0}$ with $\text{Distortion}_j(\theta)$ across different values of θ. For all $\theta \in [0, 0.5]$, the correlation between these two measures is never more than 0.05, and the highest adjusted R^2 is below 0.01. This suggests PH does a poor job targeting products with higher marginal externalities and thus is unlikely to outperform a flat volumetric tax. Indeed, without the ability to "tag" the right products, PH may perform *worse* than a volumetric tax. While the volumetric tax applies uniformly across products, PH's distortions vary widely (with a standard deviation of roughly 2.15 when $\theta = 0$), increasing the likelihood that prices will deviate significantly from the planner's preferred ones, even if the average aligns.[81]

The approach above in (18) and the related challenges extend beyond the PH game. In many settings, firms with market power set prices based on the inverse elasticity of demand, marginal cost, and opportunity cost from multiproduct pricing:

$$p_j = \underbrace{\frac{1}{1 + \theta^*/\epsilon_{jj}}}_{\mu_j(\theta^*)} [mc_j + UPP_j(\kappa)].$$

The framework described above captures the incentives of a single multi-product monopolist (such as the Maine alcohol retailer), but it also applies to many multi-product oligopoly settings. For instance, in the upstream market for distilled spirits, firms like Diageo and Bacardi are sole sellers

[80] This also assumes that the planner extracts wholesaler profits through lump-sum transfers.

[81] See further discussion of the tax alternative in Appendix A.2, and comparisons using our data in Appendix C.1.

of their respective product portfolios. The framework even encompasses more complex settings such as common ownership (Backus et al., 2021a), and it can be extended to vertical structures like double marginalization (Villas-Boas, 2007) or vertical integration (Luco and Marshall, 2021).

Each of these environments differs in terms of ownership structure (κ) and may feature planner preferences $\theta^* \neq 1$. What they share is a common pricing logic: firms apply a proportional markup $\mu_j(\theta^*)$ to an augmented marginal cost, given by $(mc_j + UPP_j(\kappa))$. This implies that price-cost margins $p_j - mc_j$ will generally be larger for products with higher marginal costs. In this sense, market power acts like a tax on "product quality" or "branding" when those manifest as higher marginal costs.

Market power will more effectively target externalities when the external harm $D^e_{j\to 0}$ is positively correlated with $(mc_j + UPP_j(\kappa))$.[82] However, if $D^e_{j\to 0}$ is uncorrelated or negatively correlated with marginal costs then this will lead to high-end products being priced *too high*, while low-end products being priced *too low*, echoing the inefficiencies found in the PH game. This misalignment becomes more severe when marginal costs vary widely, as they do in our data. For example, the manufacturer price for the cheapest product (Dubra Vodka) is around \$4/L, while prices above \$30/L are not uncommon (see Figure 5).

The primary advantage of market power, relative to simple tax instruments, is the ability to set product-specific markups $\mu_j(\theta^*)$ based on inverse demand elasticities, and is the focus of previous studies (Miravete et al., 2020; O'Connell and Smith, 2024). This flexibility may help raise revenue more efficiently than a uniform sales tax, especially if elasticities are heterogeneous. In (A.11), we show that the an approximation of the optimal sales tax rate depends on a cost-weighted harmonic mean of the planner's markups or elasticities: $\mathbb{E}_{mc}\left[\frac{1}{\mu_j(\theta)}\right] \quad \text{or} \quad 1 + \theta \cdot \mathbb{E}_{mc}\left[\frac{1}{\epsilon_{jj}}\right]$. In our data (see Figure 7), own-price elasticities range from $\epsilon_{jj} \in (-5.6, -4.2)$, but the implied range of relative μ_j is small:$\frac{1-\frac{1}{4.2}}{1-\frac{1}{5.6}} < 1.08$. Even under a much wider elasticity range, such as $(-9, -2.5)$, this ratio remains below 1.5, suggesting the value of flexible markups may be dominated by dispersion in marginal costs.

Overall, when marginal costs vary significantly, we expect market power to be a poor substitute for externality-correcting taxes unless the marginal external harm $D^e_{j\to 0}$ is positively correlated with marginal cost. Conversely, if elasticities are highly heterogeneous and the planner is focused primarily on raising revenue, outsourcing pricing to firms with market power may offer some modest advantage.

7. Conclusion

We show that the post-and-hold system employed by Connecticut is not effective at discouraging the consumption of ethanol or raising tax revenues when compared to simple, commonly used tax instruments. There exist many combinations of commonly used sales taxes and excise taxes that

[82] In our data, the correlation between $UPP_j = \sum_{k\neq j} D_{j\to k}(p_k - mc_k)$ and mc_j is high (around 0.85), because expensive, high-margin products tend to have high-margin substitutes.

increase consumer surplus, reduce alcohol consumption, and raise more revenue (even compared to a hypothetical lump-sum tax on wholesale profits) relative to the post and hold system. Indeed, it is possible to reduce overall ethanol consumption (and associated externalities) by more than 12.87% without reducing consumer surplus, and while increasing tax revenues by nearly 300% (or around \$180 million per year).

Our results shed additional light on previous studies of alcoholic beverages because we are able to trace out a wide range of policy instruments over a variety of different values. As an example, we show that the minimum ethanol unit price adopted by Scotland (and analyzed by Griffith et al. (2022)) is very similar to the solution of a social planner who wishes to maximize consumer surplus subject to an upper bound on aggregate ethanol consumption. While this policy is effective at limiting consumption, it is ineffective at raising tax revenues, which perhaps explains why it has not been more widely adopted. Likewise, we show that a uniform sales tax rate does a relatively good job approximating the problem of a social planner who maximizes consumer surplus subject to a revenue constraint. However, while the uniform sales tax is able to generate similar levels of consumer surplus and tax revenue as the "Ramsey" planner, it does so at significantly higher levels of ethanol consumption (and hence negative externalities). This helps to reconcile our results with prior studies of uniform markup rules (which operate like sales or *ad valorem* taxes) set by the state-run monopolist in Pennsylvania in Miravete et al. (2018, 2020).

Our findings are enabled by unusually comprehensive data. By combining wholesaler prices with upstream (manufacturer/distiller) input prices, we are able to measure wholesale markups set by profit-maximizing firms. We find that these additive markups generally increase with input costs (see Figure 5). Matching these observed markups — along with micro-moments that indicate that lower-income consumers typically pay lower prices and consume slightly *less*, not more, alcohol than higher-income consumers[83] — allows us to estimate demand in a way that directly informs our counterfactual policy analysis.

We construct a diversion-based measure that captures the marginal reduction in ethanol consumption for each product $D^e_{j\to 0}$ and illustrate that the least expensive products tend to have the largest values (see Figure 7). Thus, by raising the prices of these low-end products and reducing the prices of premium ones, we can undo the distortion in relative prices introduced by the PH system which sets higher markups for products with higher marginal costs (and less elastic demand). This reallocation improves consumer surplus while reducing ethanol consumption.

The seemingly "free lunch" arises because firms with market power face very different incentives than a social planner. When products are differentiated, especially when marginal costs vary significantly, relying on firms to regulate externalities through market power may be far from optimal. In our setting, consumers value product quality, but market power leads firms to impose an implicit "tax" that is too high on marginal costs (product quality) and too low on externalities.

[83]We confirm this pattern in related work (Conlon et al., 2024)

This distorts the choices of infra-marginal consumers. The idea that *any* policy reducing sin-good consumption is equally effective is misleading: optimal market design must account for both the planner's multiple objectives and consumer preferences.

These findings offer a cautionary message to policymakers who hope to outsource the mitigation of negative externalities to private firms. They also apply more broadly to other sin-good markets with significant marginal cost or quality variation. For example, in legalized marijuana markets, many states restrict competition by limiting entry (Thomas, 2019; Hollenbeck et al., 2024), or impose *ad valorem* taxes at different stages of the supply chain (Hansen et al., 2022). These policies may underperform relative to simple taxes on volume or THC content (Hansen et al., 2020) in both externality correction and revenue generation. Particularly in differentiated product markets when costs or product quality are dispersed, restricting competition may not deliver the benefits policymakers expect.

References

Allcott, H., C. Knittel, and D. Taubinsky (2015): "Tagging and Targeting of Energy Efficiency Subsidies," *American Economic Review*, 105, 187–191.

Allcott, H., B. B. Lockwood, and D. Taubinsky (2019): "Regressive Sin Taxes, with an Application to the Optimal Soda Tax*," *The Quarterly Journal of Economics*, 134, 1557–1626.

Asker, J. (2016): "Diagnosing Foreclosure due to Exclusive Dealing," *The Journal of Industrial Economics*, 64, 375–410.

Atalay, E., E. Frost, A. T. Sorensen, C. J. Sullivan, and W. Zhu (2025): "Scalable Demand and Markups," *Journal of Political Economy*, Forthcoming.

Backus, M., C. Conlon, and M. Sinkinson (2021a): "Common Ownership and Competition in the Ready-to-Eat Cereal Industry," Working Paper 28350, National Bureau of Economic Research.

——— (2021b): "Common Ownership in America: 1980–2017," *American Economic Journal: Microeconomics*, 13, 273–308.

Berry, S., J. Levinsohn, and A. Pakes (1995): "Automobile Prices in Market Equilibrium," *Econometrica*, 63, 841–890.

——— (1999): "Voluntary Export Restraints on Automobiles: Evaluating a Trade Policy," *American Economic Review*, 89, 400–430.

——— (2004): "Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market," *Journal of Political Economy*, 112, 68–105.

Berry, S. T. and P. A. Haile (2014): "Identification in Differentiated Products Markets Using Market Level Data," *Econometrica*, 82, 1749–1797.

——— (2024): "Nonparametric Identification of Differentiated Products Demand Using Micro Data," *Econometrica*.

Birchall, C., D. Mohapatra, and F. Verboven (2024): "Estimating Substitution Patterns and Demand Curvature in Discrete-Choice Models of Product Differentiation," *The Review of Economics and Statistics*, 1–40.

Blanchette, J. G., C. S. Ross, and T. S. Naimi (2020): "The Rise and Fall of Alcohol Excise Taxes in U.S. States, 1933–2018," *Journal of Studies on Alcohol and Drugs*, 81, 331–338.

Brenkers, R. and F. Verboven (2006): "Liberalizing A Distribution System: The European Car Market," *Journal of the European Economic Association, MIT Press*, 4, 216–251.

Bresnahan, T. F. (1987): "Competition and Collusion in the American Automobile Industry: The 1955 Price War," *The Journal of Industrial Economics*, 35, 457–482.

Chamberlain, G. (1987): "Asymptotic efficiency in estimation with conditional moment restrictions," *Journal of Econometrics*, 34, 305–334.

Conlon, C. and J. Gortmaker (2020): "Best practices for differentiated products demand estimation with PyBLP," *The RAND Journal of Economics*, 51, 1108–1161.

——— (2025): "Incorporating Micro Data into Differentiated Products Demand Estimation with PyBLP," *Journal of Econometrics*, 105926.

Conlon, C. and J. H. Mortimer (2021): "Empirical properties of diversion ratios," *The RAND Journal of Economics*, 52, 693–726.

Conlon, C., N. Rao, and Y. Wang (2024): "Who Pays Sin Taxes? Understanding the Overlapping Burdens of Corrective Taxes," *The Review of Economics and Statistics*, 106, 1719–1729.

Conlon, C. T. and N. L. Rao (2020): "Discrete Prices and the Incidence and Efficiency of Excise Taxes," *American Economic Journal: Economic Policy*, 12, 111–143.

Cooper, J. C. and J. D. Wright (2012): "Alcohol, Antitrust, and the 21st Amendment: An Empirical Examination of Post and Hold Laws," *International Review of Law and Economics*, 32, 379–392.

Dubois, P., R. Griffith, and M. O'Connell (2020): "How Well Targeted Are Soda Taxes?" *American Economic Review*, 110, 3661–3704.

Gandhi, A. and J.-F. Houde (2019): "Measuring Substitution Patterns in Differentiated-Products Industries," Working Paper 26375, National Bureau of Economic Research.

Griffith, R., L. Nesheim, and M. O'Connell (2018): "Income effects and the welfare consequences of tax in differentiated product oligopoly," *Quantitative Economics*, 9, 305–341.

Griffith, R., M. O'Connell, and K. Smith (2019): "Tax design in the alcohol market," *Journal of Public Economics*, 172, 20 – 35.

Griffith, R., M. O'Connell, and K. Smith (2022): "Price Floors and Externality Correction," *The Economic Journal*, 132, 2273–2289.

Grigolon, L. and F. Verboven (2014): "Nested Logit or Random Coefficients Logit? A Comparison of Alternative Discrete Choice Models of Product Differentiation," *The Review of Economics and Statistics*, 96, 916–935.

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*, 73, 511–544.

Hansen, B., K. Miller, and C. Weber (2022): "Vertical integration and production inefficiency in the presence of a gross receipts tax," *Journal of Public Economics*, 212, 104693.

Hollenbeck, B. and R. Z. Giroldo (2022): "Winning Big: Scale and Success in Retail Entrepreneurship," *Marketing Science*, 41, 271–293.

Hollenbeck, B., S. Hristakeva, and K. Uetake (2024): "Retailer Price Competition and Assortment Differentiation: Evidence from Entry Lotteries," SSRN Scholarly Paper 4857371, Social Science Research Network, Rochester, NY.

Hollenbeck, B. and K. Uetake (2021): "Taxation and market power in the legal marijuana industry," *The RAND Journal of Economics*, 52, 559–595.

Illanes, G. and S. Moshary (2020): "Market Structure and Product Assortment: Evidence from a Natural Experiment in Liquor Licensure," Working Paper 27016, National Bureau of Economic Research.

Jaffe, S. and E. G. Weyl (2013): "The First-Order Approach to Merger Analysis," *American Economic Journal: Microeconomics*, 5, 188–218.

Kenkel, D. S. (2005): "Are Alcohol Tax Hikes Fully Passed Through to Prices? Evidence from Alaska," *American Economic Review*, 95, 273–277.

Kingston, S. (2011): *Greening EU Competition Law and Policy*, Antitrust and Competition Law, Cambridge: Cambridge University Press.

Leung, S. F. and C. E. Phelps (1993): "My kingdom for a drink . . .? A review of estimates of the price sensitivity of demand for alcoholic beverages," in *Economics and the Prevention of Alcohol-Related Problems: Proceedings of a Workshop on Economic and Socioeconomic Issues in the Prevention of Alcohol- Related Problems*, ed. by M. E. Hilton and G. Bloss, Bethesda, MD: National Institute on Alcohol Abuse and Alcoholism, no. NIH Pub. No. 93-3513 in NIAAA Research Monograph No. 25., 1–32.

Levy, D. T., L. M. Sánchez-Romero, C. E. Douglas, and D. T. Sweanor (2021): "An Analysis of the Altria-Juul Labs Deal: Antitrust and Population Health Implications," *Journal of Competition Law & Economics*, 17, 458–492.

Linklaters (2020): "92 percent of businesses call for changes to competition rules to boost climate change collaboration," .

Luco, F. and G. Marshall (2020): "The Competitive Impact of Vertical Integration by Multiproduct Firms," *American Economic Review*, 110, 2041–2064.

——— (2021): "Diagnosing Anticompetitive Effects of Vertical Integration by Multiproduct Firms," *Review of Industrial Organization*, 59, 381–392.

Miller, N. H., G. Sheu, and M. C. Weinberg (2021): "Oligopolistic Price Leadership and Mergers: The United States Beer Industry," *American Economic Review*, 111, 3123–3159.

Miller, N. H. and M. C. Weinberg (2017): "Understanding the Price Effects of the MillerCoors Joint Venture," *Econometrica*, 85, 1763–1791.

Miravete, E. J., K. Seim, and J. Thurk (2018): "Market Power and the Laffer Curve," *Econometrica*, 86, 1651–1687.

——— (2020): "One Markup to Rule Them All: Taxation by Liquor Pricing Regulation," *American Economic Journal: Microeconomics*, 12, 1–41.

——— (2023): "Elasticity and Curvature of Discrete Choice Demand Models," .

Nevo, A. (2000): "Mergers with Differentiated Products: The Case of the Ready-to-Eat Cereal Industry," *The RAND Journal of Economics*, 31, 395–421.

——— (2001): "Measuring Market Power in the Ready-to-Eat Cereal Industry," *Econometrica*, 69, 307–342.

O'Connell, M. and K. Smith (2024): "Optimal Sin Taxation and Market Power," *American Economic Journal: Applied Economics*, 16, 34–70.

Oum, T. H. and M. W. Tretheway (1988): "Ramsey Pricing in the Presence of Externality Costs," *Journal of Transport Economics and Policy*, 22, 307–317.

Petrin, A. (2002): "Quantifying the Benefits of New Products: The Case of the Minivan," *Journal of Political Economy*, 110, 705–729, publisher: The University of Chicago Press.

Ramsey, F. P. (1927): "A Contribution to the Theory of Taxation," *The Economic Journal*, 37, 47–61.

Saffer, H. and M. Gehrsitz (2016): "The effect of post-and-hold laws on alcohol consumption," *International Review of Law and Economics*, 45, 63–71.

Sandmo, A. (1975): "Optimal Taxation in the Presence of Externalities," *The Swedish Journal of Economics*, 77, 86–98.

Sass, T. (2005): "The competitive effects of exclusive dealing: Evidence from the U.S. beer industry," *International Journal of Industrial Organization*, 23, 203–225.

Sass, T. and D. S. Saurman (1993): "Mandated Exclusive Territories and Economic Efficiency: An Empirical Analysis of the Malt-Beverage Industry," *Journal of Law and Economics*, 36, 153–77.

Schinkel, M. P. and L. Treuren (2020): "Green Antitrust: (More) Friendly Fire in the Fight against Climate Change," SSRN Scholarly Paper ID 3749147, Social Science Research Network, Rochester, NY.

Seim, K. and J. Thurk (2023): "Inflation, Indexation, and Market Power," Working Paper, University of Georgia.

Seim, K. and J. Waldfogel (2013): "Public Monopoly and Economic Efficiency: Evidence from the Pennsylvania Liquor Control Board's Entry Decisions," *American Economic Review*, 103, 831–862.

Seo, B. (2019): "Firm Scope and the Value of One-Stop Shopping in Washington State's Deregulated Liquor Market," SSRN Scholarly Paper ID 2863094, Social Science Research Network, Rochester, NY.

Thomas, D. (2019): "License Quotas and the Inefficient Regulation of Sin Goods: Evidence From the Washington Recreational Marijuana Market," SSRN Scholarly Paper ID 3312960, Social Science Research Network, Rochester, NY.

U.S. Department of Treasury (2022): "Competition in the Markets for Beer, Wine, and Spirits," Tech. rep., Department of the Treasury.

Villas-Boas, S. B. (2007): "Vertical Relationships between Manufacturers and Retailers: Inference with Limited Data," *The Review of Economic Studies*, 74, 625–652.

——— (2009): "An empirical investigation of the welfare effects of banning wholesale price discrimination," *The RAND Journal of Economics*, 40, 20–46.

Wagenaar, A. C., M. J. Salois, and K. A. Komro (2009): "Effects of beverage alcohol price and tax levels on drinking: a meta-analysis of 1003 estimates from 112 studies," *Addiction (Abingdon, England)*, 104, 179–190.

Table 1: Summary Statistics: Wholesale and Manufacturer Price Connecticut Q3 2007 - Q2 2013

	# Obs	Share	Proof	% Flavored	Manufacturer Price	Manufacturer Margin*	Wholesaler Price	Wholesaler Margin	Retailer Price	Retailer Margin
Gin	59	7.40	87.07	0.02	11.15	3.01	16.21	3.79	18.72	2.34
Rum	147	17.50	73.63	0.21	10.17	2.60	15.08	3.65	17.60	2.52
Tequila	92	4.90	80.04	0.00	15.17	4.07	22.05	5.60	28.51	4.70
Vodka	208	44.80	79.19	0.15	10.73	2.79	15.42	3.42	18.05	2.54
NA Whiskey	127	15.20	81.80	0.00	11.59	3.18	17.41	4.54	20.08	2.76
UK Whiskey	102	10.20	80.79	0.00	18.36	4.51	25.04	5.41	28.15	3.12
750mL	310	20.10	79.05	0.18	16.44	4.32	23.57	5.85	28.32	4.74
1L	174	23.20	79.32	0.12	13.80	3.73	19.92	4.85	24.85	4.35
1.75L	251	56.70	79.55	0.08	9.32	2.36	13.53	2.94	14.91	1.36
All	735	100.00	79.40	0.11	11.79	3.07	17.03	3.97	19.82	2.71

Note: The table above describes manufacturer, wholesale, and retail prices and margins for 735 of 1,502 products (used in our estimation procedure) by category and size. The number of products corresponds to brand-size combinations, such as Smirnoff Vodka-750mL or Tanqueray Gin-1L. All averages are weighted by total liters sold. *Share* describes the share of total liters sold. The average *Proof* and percentage *Flavored* is reported. The average prices and margins are reported on a *per liter* basis.

The *Manufacturer Margin* is the difference between the manufacturer price and the estimated manufacturer marginal cost from the demand and supply model (net of federal excise taxes). All other columns in this table are observed rather than estimated.

Retailer Margin is the difference between the retail price and the wholesale price.

Wholesaler Margin is the difference between the wholesale price and manufacturer price plus state excise tax.

Federal alcohol excise taxes of $2.85 per liter of 80-proof spirits are levied on manufacturers. Connecticut state alcohol taxes, which are remitted by wholesalers, were raised from $1.18 to $1.42 per liter regardless of proof in July 2011.

Source: Harmonized Price and Quantity Data (top 750 products, average price under $60 per liter).

Table 2: Manufacturer Summary

	# Obs	Share	750mL	1L	1.75L	Manufacturer Price	Manufacturer Lerner	Wholesaler Price	Wholesaler Lerner	Retailer Price	Retailer Lerner
Diageo	155	32.70	0.16	0.21	0.63	11.75	0.30	17.00	0.23	19.26	0.11
Bacardi	48	14.20	0.21	0.34	0.45	14.30	0.24	20.03	0.23	22.63	0.11
Pernod	68	14.20	0.20	0.33	0.47	15.03	0.25	20.74	0.21	23.96	0.13
Jim Beam	102	8.30	0.18	0.23	0.59	9.59	0.27	14.55	0.24	17.56	0.14
Brown Forman	32	5.20	0.23	0.30	0.47	14.83	0.28	22.49	0.28	26.01	0.13
Skyy	26	2.90	0.27	0.06	0.67	11.18	0.22	16.00	0.21	18.58	0.13
Constellation Brands	6	2.80	0.19	0.11	0.71	7.43	0.28	12.45	0.29	14.37	0.13
Constellation	24	2.10	0.05	0.12	0.83	4.91	0.28	8.09	0.22	9.72	0.14
Star Industries	16	2.10	0.13	0.29	0.58	4.67	0.28	7.88	0.24	9.53	0.17
Imperial	6	2.10	0.19	0.10	0.71	5.72	0.27	9.16	0.23	12.16	0.24
MHW	44	2.00	0.41	0.16	0.43	11.68	0.24	16.97	0.23	20.77	0.17
Black Prince	7	2.00	0.10	0.29	0.62	3.97	0.28	5.93	0.11	7.15	0.17
Heaven Hill	21	1.50	0.18	0.05	0.77	6.65	0.24	10.12	0.20	12.05	0.15
White Rock	8	1.30	0.24	0.00	0.76	7.04	0.24	10.53	0.21	13.48	0.21
William Grant	17	1.30	0.22	0.12	0.65	10.40	0.26	16.02	0.25	18.62	0.11
Other	36	1.00	0.42	0.16	0.42	9.57	0.24	13.86	0.21	18.34	0.23
Remy-Cointreau	16	1.00	0.35	0.13	0.52	18.09	0.22	24.74	0.20	28.94	0.15
US Distributors	6	0.80	0.23	0.00	0.77	7.02	0.22	9.47	0.11	14.52	0.33
Sazerac	20	0.70	0.34	0.24	0.42	9.92	0.25	14.52	0.20	18.69	0.22
Moet Hennessy	10	0.60	0.41	0.37	0.22	24.35	0.23	31.02	0.17	37.43	0.17
LuxCo	19	0.60	0.17	0.38	0.45	7.13	0.25	10.97	0.23	14.60	0.23
MS Walker	10	0.20	0.08	0.48	0.45	5.33	0.22	7.32	0.09	10.99	0.25
McCormick	7	0.20	0.07	0.56	0.37	5.09	0.27	7.59	0.17	11.96	0.23
Proximo	3	0.10	1.00	0.00	0.00	20.02	0.26	29.27	0.26	37.64	0.21
Duggans	2	0.10	0.00	0.26	0.74	8.00	0.24	12.93	0.28	15.30	0.15
Infinium	1	0.00	0.00	0.00	1.00	5.54	0.26	8.84	0.24	10.78	0.17
Castle Brands	1	0.00	1.00	0.00	0.00	11.86	0.23	16.79	0.21	22.03	0.23

Note: The table above reports product shares, average prices, and Lerner markups by manufacturer for 735 of 1,502 products (used in our estimation procedure). The number of products corresponds to brand-size combinations, such as Smirnoff Vodka-750mL or Tanqueray Gin-1L. Average prices and Lerner markups are reported on a *per liter* basis. All averages are weighted by total liters sold.

Share describes the share of total liters sold by each manufacturer.

Manufacturer Lerner is the difference between the manufacturer price and the estimated manufacturer marginal cost from the demand and supply model (net of federal excise taxes) scaled by the estimated manufacturer marginal cost. All other columns in this table are observed rather than estimated.

Retail Lerner is the difference between the retail price and the wholesale price scaled by the retail price.

Wholesale Lerner is the difference between the wholesale price and manufacturer price plus state excise tax scaled by the wholesale price.

Federal alcohol excise taxes of $2.85 per liter of 80-proof spirits are levied on manufacturers. Connecticut state alcohol taxes, which are remitted by wholesalers, were raised from $1.18 to $1.42 per liter regardless of proof in July 2011.

Source: Harmonized Price and Quantity Data (top 750 products, average under $60 per liter).

Table 3: Parameter Estimates: Full Model

Π	Const	Price	1750mL
Below \$25k	2.433	-0.736	-0.442
	(0.287)	(0.056)	(0.083)
\$25k-\$45k	0.243	-0.720	-0.258
	(0.328)	(0.095)	(0.097)
\$45k-\$70k	0.000	-0.768	0.000
	(0.000)	(0.094)	(0.000)
\$70k-\$100k	-0.960	-1.032	-0.275
	(0.324)	(0.094)	(0.096)
Above \$100k	-3.762	-2.291	-0.794
	(0.262)	(0.074)	(0.077)
Σ^2			
Const	3.868	1.271	0.000
	(0.740)	(0.150)	(0.000)
Price	1.271	0.418	0.000
	(0.150)	(0.031)	(0.000)
1750mL	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)
Nesting Parameter ρ		0.27	
		(0.021)	
Fixed Effects		Brand+Quarter	
Model Predictions	25%	50%	75%
Own Elasticity	-5.072	-4.772	-4.484
Aggregate Elasticity	-0.545	-0.530	-0.506
Observed Wholesale Markup (PH)	0.188	0.233	0.276
Predicted Wholesale Markup (PH)	0.222	0.238	0.255
Outside Good Diversion D_{j0}	0.428	0.458	0.496
Ethanol Diversion D^e_{j0}	0.171	0.188	0.211

Note: The table above reports parameter estimates from our RCNL model. The price coefficient is log-normally distributed so that $\alpha_i = -e^{\pi_k^p + \Sigma \cdot \nu_i}$ is always negative and more negative for values of π_k^p closer to zero. High-income consumers $\pi^p = -2.291$ have smaller coefficients than low-income consumers -0.736 and are thus *less* price sensitive.
Own pass-through is the change in equilibrium prices for product j (under PH) in response to a \$1.00 increase in the price of good j.
Aggregate elasticity is the change in total spirits volume in response to a 1% price increase for all products.
Source: Harmonized Price and Quantity Data (top 750 products, average wholesale price below \$60 per liter), 24 quarterly periods. Authors' calculations.

Table 4: Best Substitutes: Diversion Ratios 2013 Q2

	Median Price	% Substitution		Median Price	% Substitution
Capt Morgan Spiced 1.75 L ($15.85)			Cuervo Gold 1.75 L ($18.33)		
Bacardi Superior Lt Dry Rum 1.75 L	12.52	7.59	Cuervo Gold 1.0 L	21.32	3.26
Bacardi Superior Lt Dry Rum 1.0 L	15.03	2.06	Sauza Giro Tequila Gold 1.0 L	8.83	2.15
Smirnoff 1.75 L	11.85	1.87	Don Julio Silver 1.75 L	22.81	2.12
Bacardi Dark Rum 1.75 L	12.52	1.57	Smirnoff 1.75 L	11.85	1.80
Lady Bligh Spiced V Island Rum 1.75 L	9.43	1.46	Cuervo Gold 0.75 L	23.44	1.44
Woodford 0.75 L ($34.55)			Beefeater Gin 1.75 L ($17.09)		
Jack Daniel Black Label 1.0 L	27.08	4.25	Tanqueray 1.75 L	17.09	7.11
Jack Daniel Black Label 1.75 L	21.85	4.19	Gordons 1.75 L	11.19	2.55
Jack Daniel Black Label 0.75 L	29.21	2.66	Seagrams Gin 1.75 L	10.23	1.84
Makers Mark 1.0 L	32.79	2.46	Smirnoff 1.75 L	11.85	1.82
Makers Mark 0.75 L	31.88	1.53	Gilbey Gin 1.75 L	9.30	1.56
Dubra Vdk Dom 80P 1.75 L ($5.88)			Belvedere Vodka 0.75 L ($30.55)		
Popov Vodka 1.75 L	7.66	3.88	Absolut Vodka 1.75 L	15.94	3.34
Smirnoff 1.75 L	11.85	2.79	Grey Goose 1.0 L	32.08	2.71
Sobieski Poland 1.75 L	9.09	1.93	Smirnoff 1.75 L	11.85	2.36
Grays Peak Vdk Dom 1.75 L	9.16	1.78	Ktl1 Vdk Im 1.75 L	20.71	1.49
Bellows Vodka 1.0 L	6.21	1.49	Absolut Vodka 1.0 L	24.91	1.47

Note: The table above reports diversion rates for five popular products. Per liter wholesale prices are reported for 2013Q2. We compute the diversion ratio for a small price change $D_{j\to k} = \frac{\partial q_k}{\partial q_j} / \left| \frac{\partial q_j}{\partial q_j} \right|$.

A plain logit would predict the best substitute as the product with the largest overall share: Smirnoff Vodka (80-Proof, 1.75L) **with** $s_{jt} = 1.2\%$ **or** 4.37% **of "inside" sales.**

Source: Authors' calculations

Table 5: Counterfactual Policies to Limit Ethanol Consumption

Policy	Product Prices
Volumetric Tax	$p_{jt} = mc_{jt} + \tau_v$
Ethanol Tax	$p_{jt} = mc_{jt} + \tau_e \cdot e_j$
Sales Tax	$p_{jt} = mc_{jt} \cdot (1 + \tau_r)$
Minimum Unit Price	$p_{jt} = \max\{mc_{jt}, \tau_u \cdot e_j\}$
Monopoly	$\mathbf{p} = \arg\max_{\mathbf{p}} (\mathbf{p} - \mathbf{mc}) \cdot \mathbf{q}(\mathbf{p})$
Ramsey (Revenue)	$\mathbf{p}(\overline{R}, 0) = \arg\max_{\mathbf{p} \geq \mathbf{mc}} CS(\mathbf{p})$ s.t. $(\mathbf{p} - \mathbf{mc}) \cdot \mathbf{q}(\mathbf{p}) > \overline{R}$
Ramsey (Ethanol)	$\mathbf{p}(0, \overline{E}) = \arg\max_{\mathbf{p} \geq \mathbf{mc}} CS(\mathbf{p})$ s.t. $\mathbf{e} \cdot \mathbf{q} \leq \overline{E}$

Note: We examine seven policy alternatives to PH. In all counterfactuals PH pricing is replaced with taxes levied on a competitive wholesale market. *Sales* levies a single-rate sales tax (τ_r) on all spirits products to achieve the desired aggregate ethanol consumption level. Similarly, *Volume* and *Ethanol* model the impact of volumetric (τ_v) and ethanol-based (τ_e) taxes set to limit ethanol consumption. A *Minimum Price* enforces a floor based on ethanol content ($\tau_u \cdot e_j\}$) but otherwise prices products competitively.
Finally, we examine the impacts of Ramsey prices where individual product prices are set to maximize consumer surplus while meeting different constraints. The first set of Ramsey prices are set to generate a required revenue (regardless of ethanol consumption). The second set of Ramsey prices is set to cap aggregate ethanol consumption (regardless of revenue generated).

Table 6: Distributional Impacts of Counterfactual Policies

		% Change in CS					
No Change in Ethanol	% Total Revenue	% Overall	Below $25k	$25k-$45k	$45k-$70k	$70k-$100k	Above $100k
Ramsey (Ethanol)	41.5	29.9	6.2	5.6	0.8	17.1	43.4
Minimum Price	52.9	29.8	5.9	6.4	3.1	17.9	42.9
Ethanol	280.4	11.2	1.2	0.9	0.4	5.4	16.7
Volume	283.8	10.1	-0.7	-2.0	-2.0	2.9	16.3
Sales	336.2	-16.1	-2.6	-0.9	-3.6	-9.9	-23.4
Ramsey (Revenue)	340.7	-6.3	-1.1	0.1	-1.3	-4.7	-9.1
-10% Ethanol							
Ramsey (Ethanol)	66.1	19.4	-5.2	-11.5	-14.9	-0.1	35.0
Minimum Price	74.2	19.4	-5.3	-10.7	-12.9	0.7	34.6
Ethanol	290.7	2.5	-8.9	-14.4	-13.9	-8.6	10.2
Volume	293.9	1.4	-11.0	-17.2	-16.3	-11.1	9.7
Sales	333.5	-24.4	-11.6	-14.4	-16.1	-21.5	-30.3
Ramsey (Revenue)	345.0	-16.9	-12.3	-16.6	-16.8	-19.5	-18.0
+10% Ethanol							
Ramsey (Ethanol)	22.0	39.7	17.4	24.4	18.2	35.5	50.4
Minimum Price	27.2	39.7	17.1	25.2	20.7	36.4	50.0
Ethanol	266.9	19.5	11.0	16.9	15.1	19.7	22.7
Volume	270.5	18.5	9.1	14.0	12.8	17.3	22.3
Ramsey (Revenue)	332.0	1.9	7.6	14.2	11.6	7.6	-2.5
Sales	333.6	-7.7	6.2	13.5	9.5	2.4	-16.5

Note: The table above reports estimates of the impacts of the counterfactual policy alternatives described in Table 5 on tax revenue collected, overall consumer surplus and the distribution of consumer surplus across the five income bins. All effects are reported as percentage changes relative to the PH baseline. The top panel describes the impact of alternative policies that limit ethanol consumption to the same aggregate level as under PH while panels B and C report the effects of alternative policies that reduce and increase ethanol consumption by 10%, respectively. Revenue is calculated as the additional tax revenue raised by the state compared to the existing excise tax collections.
Source: Authors' calculations

Table 7: Reducing Overall Ethanol Consumption (Ethanol Taxes)

	No Change to Ethanol			No Change to Overall CS		
	Base	$wc = 1$	p^m	Base	$wc = 1$	p^m
% Δ Ethanol	0.00	-0.00	-0.00	-12.87	-12.62	-11.97
% Δ Tax Revenue	280.41	211.16	248.82	292.99	232.17	256.15
% Δ Manufacturer Profit	21.47	21.24	39.57	8.94	8.97	29.34
% Δ Total CS	11.18	10.94	10.09	-0.00	0.00	0.00
% Δ CS by Income						
Below \$25k	1.23	0.79	1.31	-11.82	-12.00	-10.73
\$25k-\$45k	0.90	0.25	0.56	-18.57	-18.76	-17.44
\$45k-\$70k	0.36	-0.16	-0.64	-17.91	-18.02	-17.41
\$70k-\$100k	5.37	4.83	4.45	-12.59	-12.71	-11.97
Above \$100k	16.73	16.64	15.21	8.25	8.34	7.72
Tax per Liter	5.48	4.48	5.02	6.50	5.47	5.83

Note: The table above reports welfare estimates for the impacts of a counterfactual ethanol tax under two scenarios: (a) no change in overall ethanol consumption (b) minimizing ethanol consumption without reducing aggregate consumer surplus.

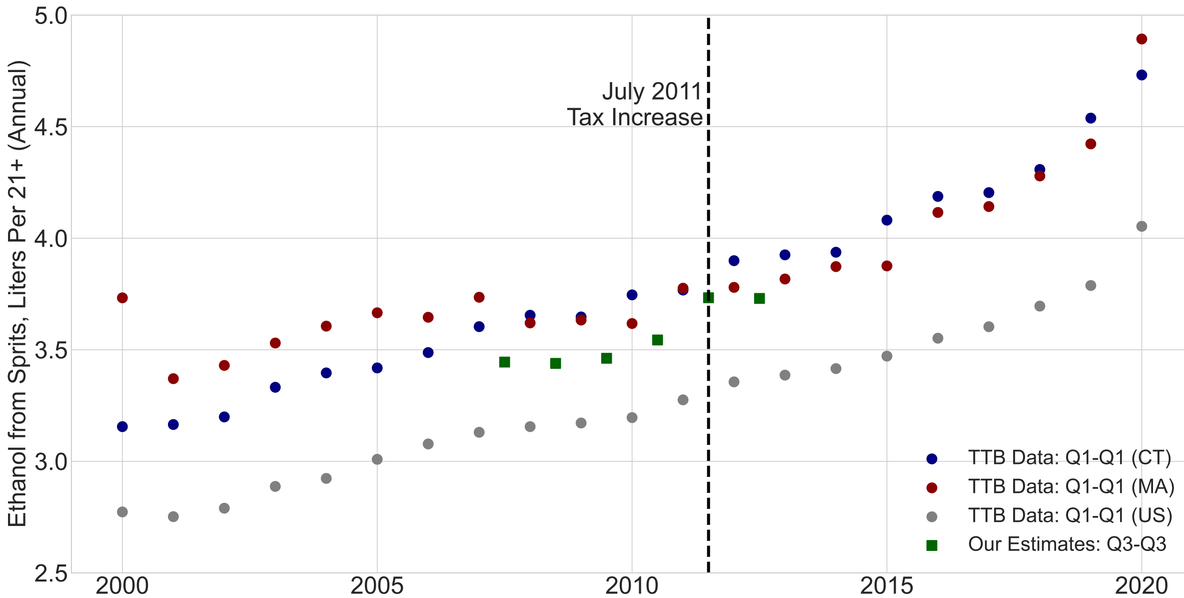
Under the Base scenario we set the wholesale price equal to the manufacturer price plus the taxes from Table 5. In the next columns, we allow for an additional \$1 per liter wholesaling cost ($wc = 1$), or we allow manufacturers to endogenously set prices (p^m) in response to counterfactual taxes but with perfectly competitive wholesaling.

Manufacturer profits increase even when prices are held fixed because absent PH, consumers substitute to higher margin/quality products.

Tax Per Liter is reported as the tax on 1L of spirits at 80-Proof (40% Alcohol by Volume)

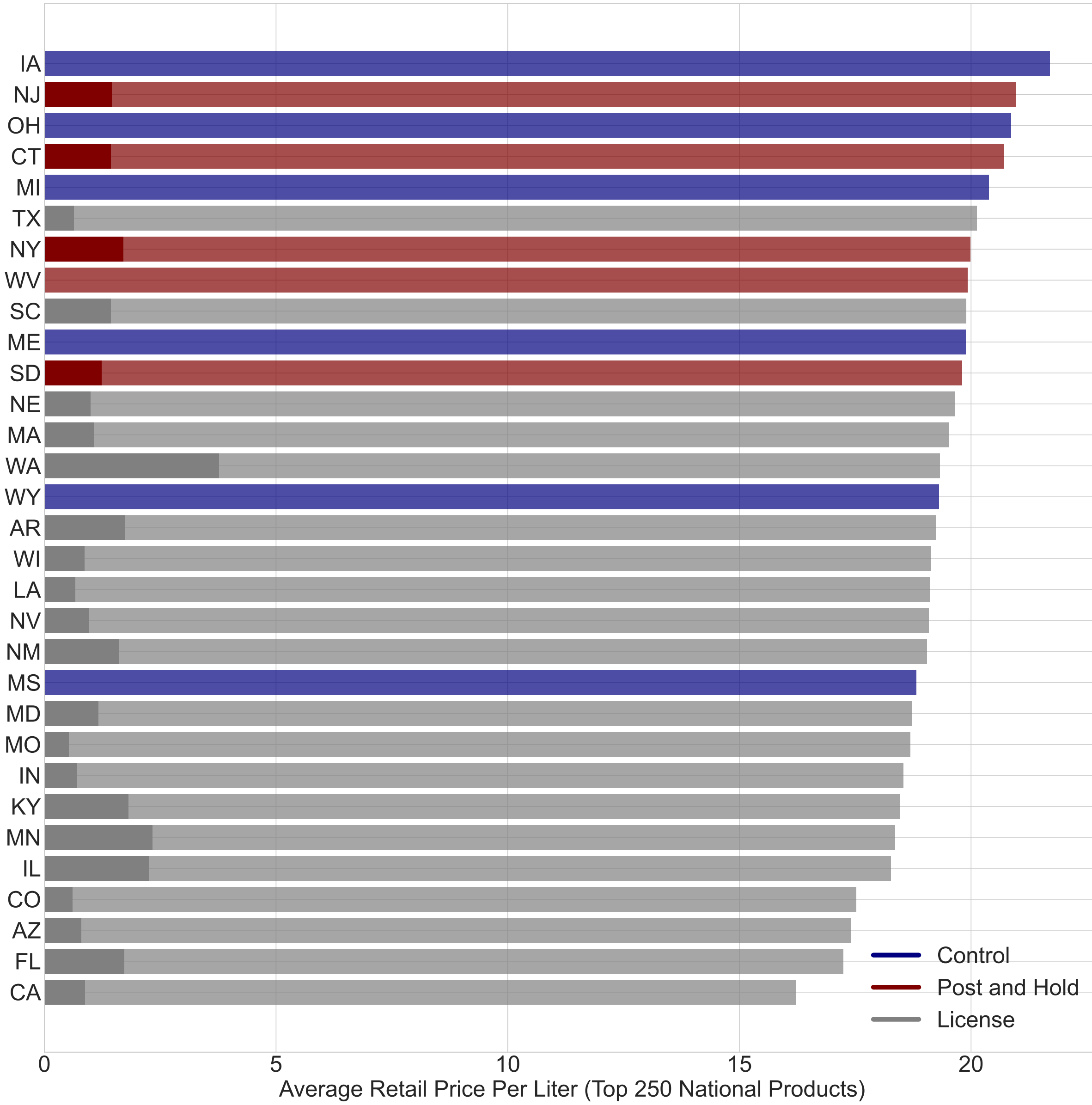
Source: Authors' calculations

Figure 1: Aggregate Trends in Spirits Consumption



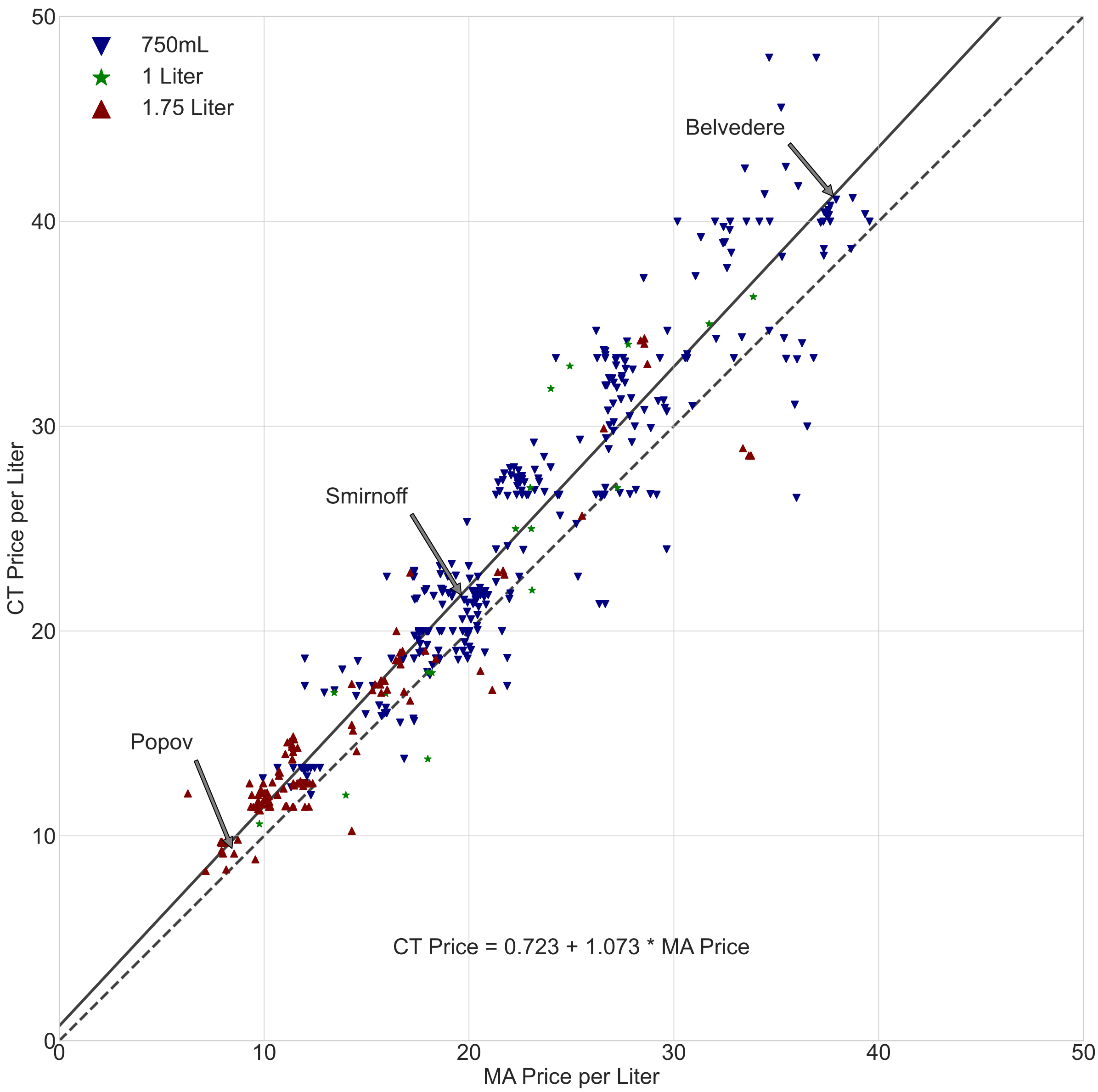
Note: Discrepancies between our data and TTB data arise from: (a) Our data measures ethanol product by product, TTB assumes a fixed proportion of ethanol by volume; (b) Our data exclude small bottles (375mL or less), cognac and liqueurs; (around 11% of sales volume) (c) because years do not align, seasonal differences in shipments and consumption.
TTB Data: Liters of Ethanol Per 21+ population for MA, CT, and US overall.
TTB Data: Reported annually in in calendar years (January-December)
Our Data: Calculated quarterly and aggregated to "Fiscal Year" of July-June (Q3-Q2).
Source: NIAAA/TTB Data; Author's calculations and Shipment data.

Figure 2: Price Indices by State, National Consumption Bundle (2013)



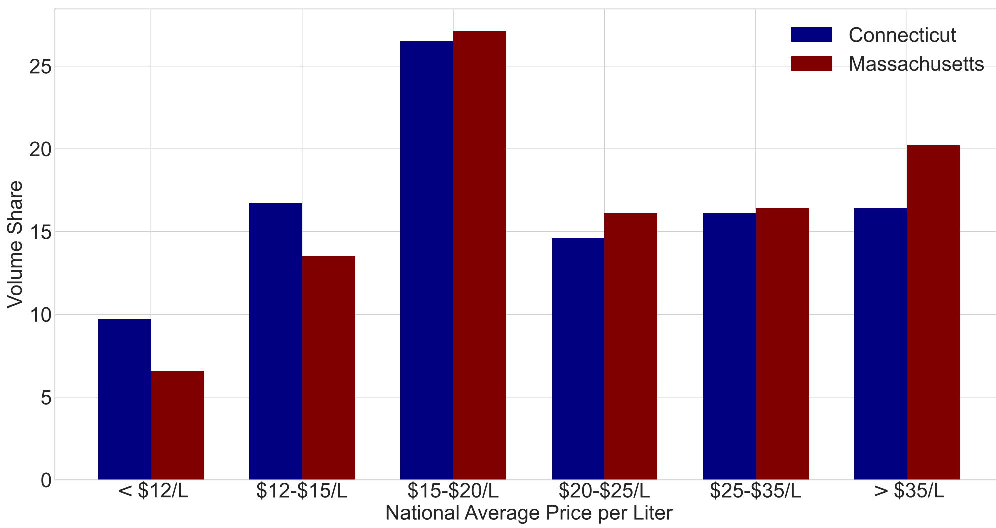
Note: The figure above plots the average retail price by state of the 250 best-selling products nation-wide. Retail prices in each state are weighted by the product's share within the top 250 national bundle by volume. As such, sales weights are constant across states so that the indices reflect only the differences in prices for the national bundle. License states such as Rhode Island and Delaware where we lack data describing sales of at least 1,000 products are excluded. Control states are shaded in blue, post-and-hold states in red and license states without post-and-hold regulations in grey. Darkly shaded bars on the left indicate state excise tax levied on the national bundle in license states (control states generally do not levy taxes on top of state markups).
Source: NielsenIQ Scanner Dataset.

Figure 3: Retail Prices for Vodka Products in Connecticut vs. Massachusetts (2013)

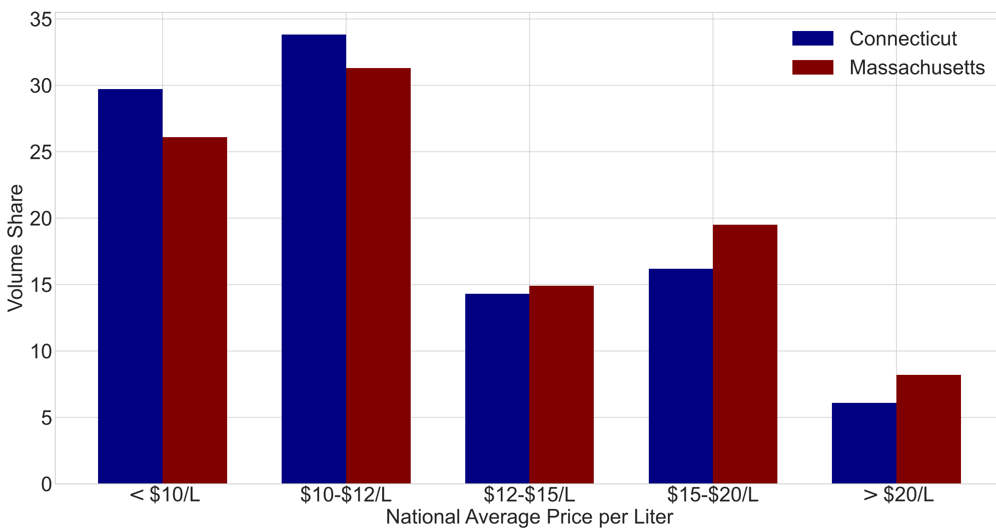


Note: The figure above compares the retail prices of individual products in Connecticut and the neighboring state of Massachusetts. Massachusetts prices are plotted on the x-axis and Connecticut prices are plotted on the y-axis with each dot representing brand-size combination, such as Smirnoff Vodka-750mL or Tanqueray Gin-1L. Prices are converted into dollars per liter and different colored markers denote 750mL (blue), 1000mL (green) and 1750mL (red) products. The dashed line plots the linear best fit and its coefficients are reported. The 45-degree line, corresponding to equal prices in Connecticut and Massachusetts, is shown as well.
Source: NielsenIQ Scanner Dataset (2013).

Figure 4: Vodka Consumption in Connecticut and Massachusetts by National Price Per Liter (2013)



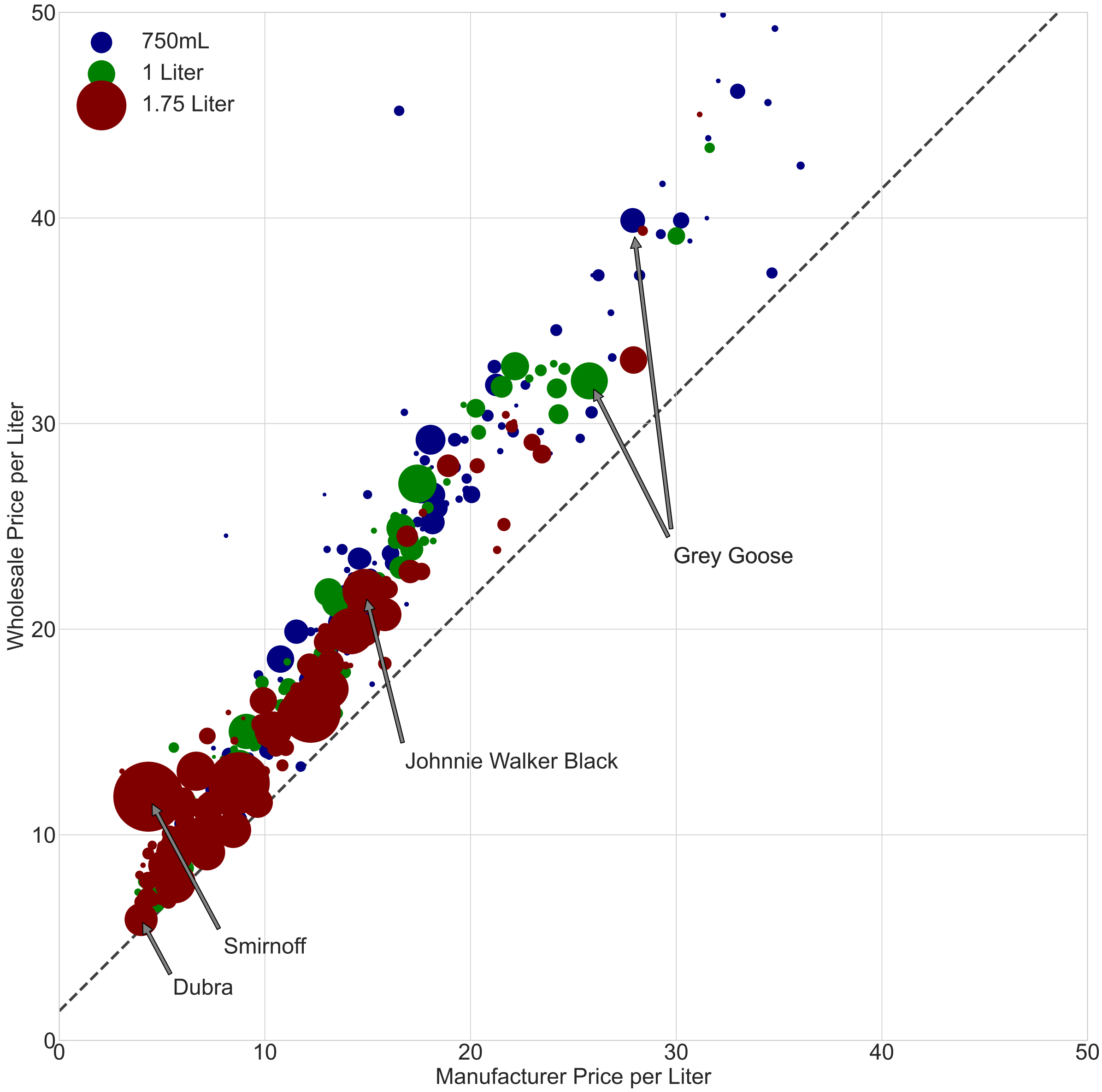
(a) 750mL Products



(b) 1.75L Products

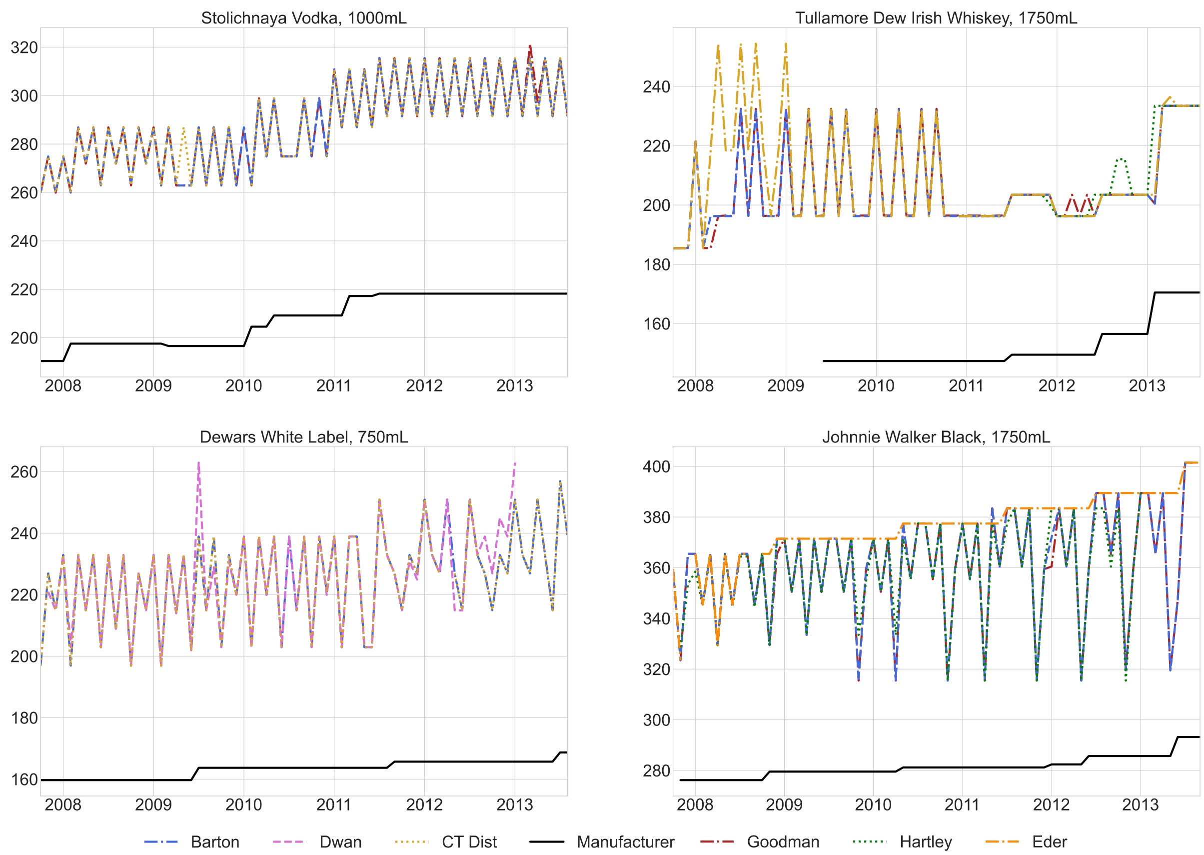
Note: The charts above show the share of vodka consumption by volume in Connecticut (blue) and Massachusetts (red) for 750mL and 1.75L products by national price per liter category. A product's national price category is determined using the average price per liter across all NielsenIQ markets outside of Connecticut designated market areas. For products only sold in Connecticut or Massachusetts the state price is used in place of the national price to calculate price per liter.
Source: NielsenIQ Scanner Data (2013).

Figure 5: Manufacturer and Wholesale Prices Q2 2013



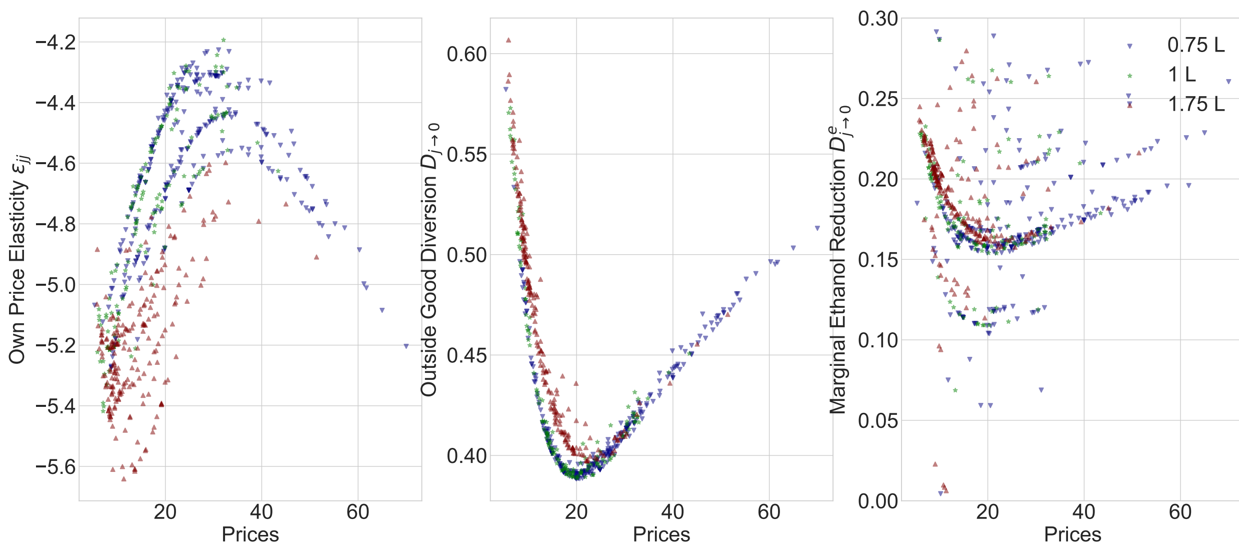
Note: The figure above plots the wholesale price against the manufacturer price, capturing how the ratio of wholesale to manufacturer price rises with manufacturer price. Prices are dollars per liter and different colored markers denote 750mL (blue), 1000mL (green) and 1750mL (red) products. Marker sizes are proportional to quarterly sales totals. The 45-degree line, corresponding to zero wholesale markup, is shown as well.
Source: Harmonized Price and Quantity Data. Period from 2013-04-01 to 2013-06-30.

Figure 6: Case Price by Wholesaler and Manufacturer Price, Four Top Selling Products



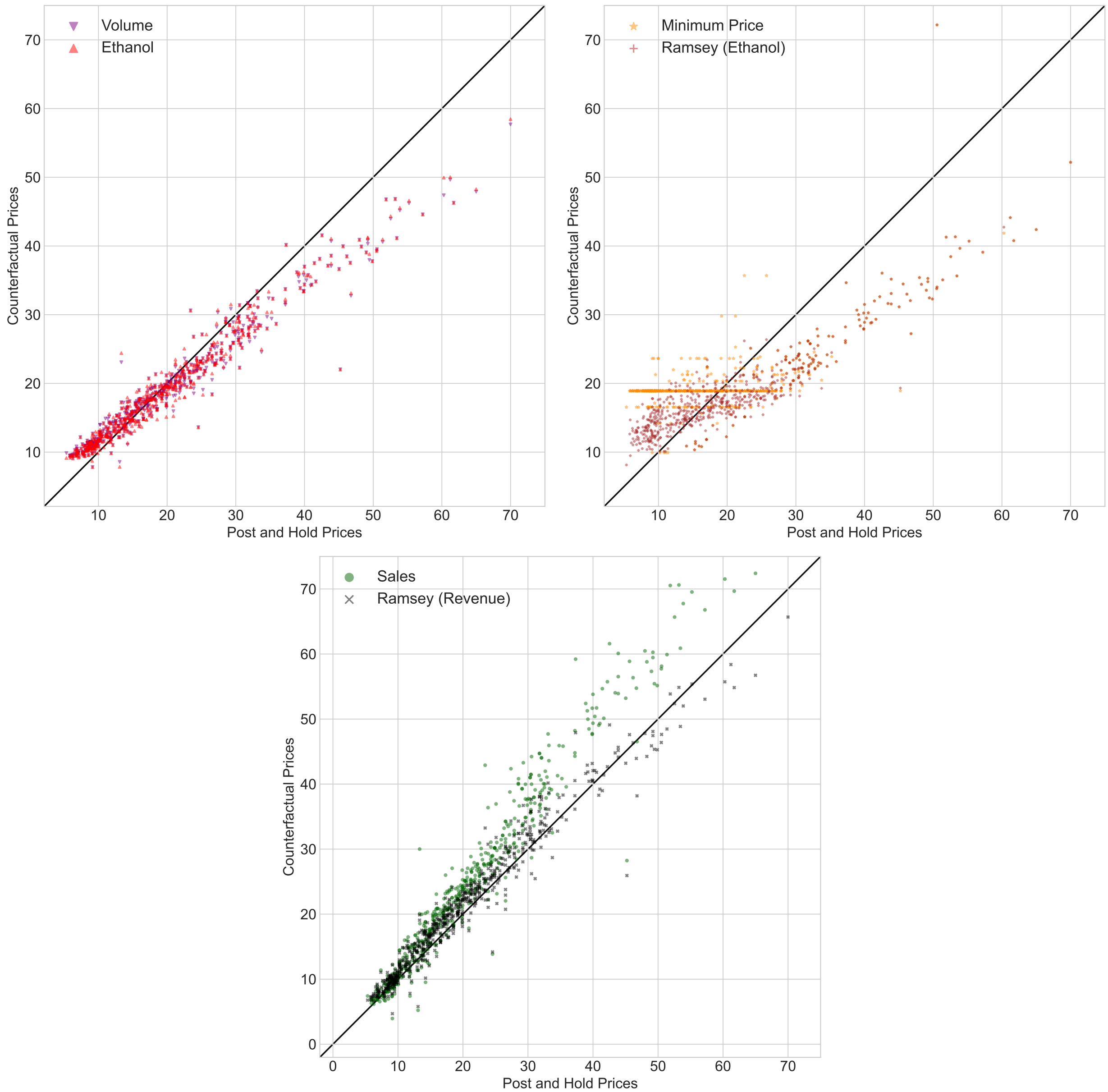
Note: The figure above plots monthly wholesale prices as well as the manufacturer price for four popular products between October 2007 and August 2013. Three wholesalers offer Stolichnaya Vodka, 1000mL (Goodman, Barton and CT Dist) and Dewars White Label, 750mL (Barton, CT Dist and Dwan), while four wholesalers sell Tullamore Dew, 1750mL (Barton, CT Dist, Goodman and Hartley) and Johnnie Walker Black, 1750mL (Barton, Eder, Goodman and Dwan) over the period. Prices offered by these distinct wholesalers overlap in the vast majority of months. While we might expect correlated wholesale price increases when manufacturer prices rise, which we observe, prices also exhibit considerable month-to-month changes between manufacturer price adjustments that happen in near lockstep across wholesalers.
Source: Harmonized Price and Quantity Data. Period from 2013-04-01 to 2013-06-30.

Figure 7: Estimated Own Elasticities and Diversion to the Outside Good



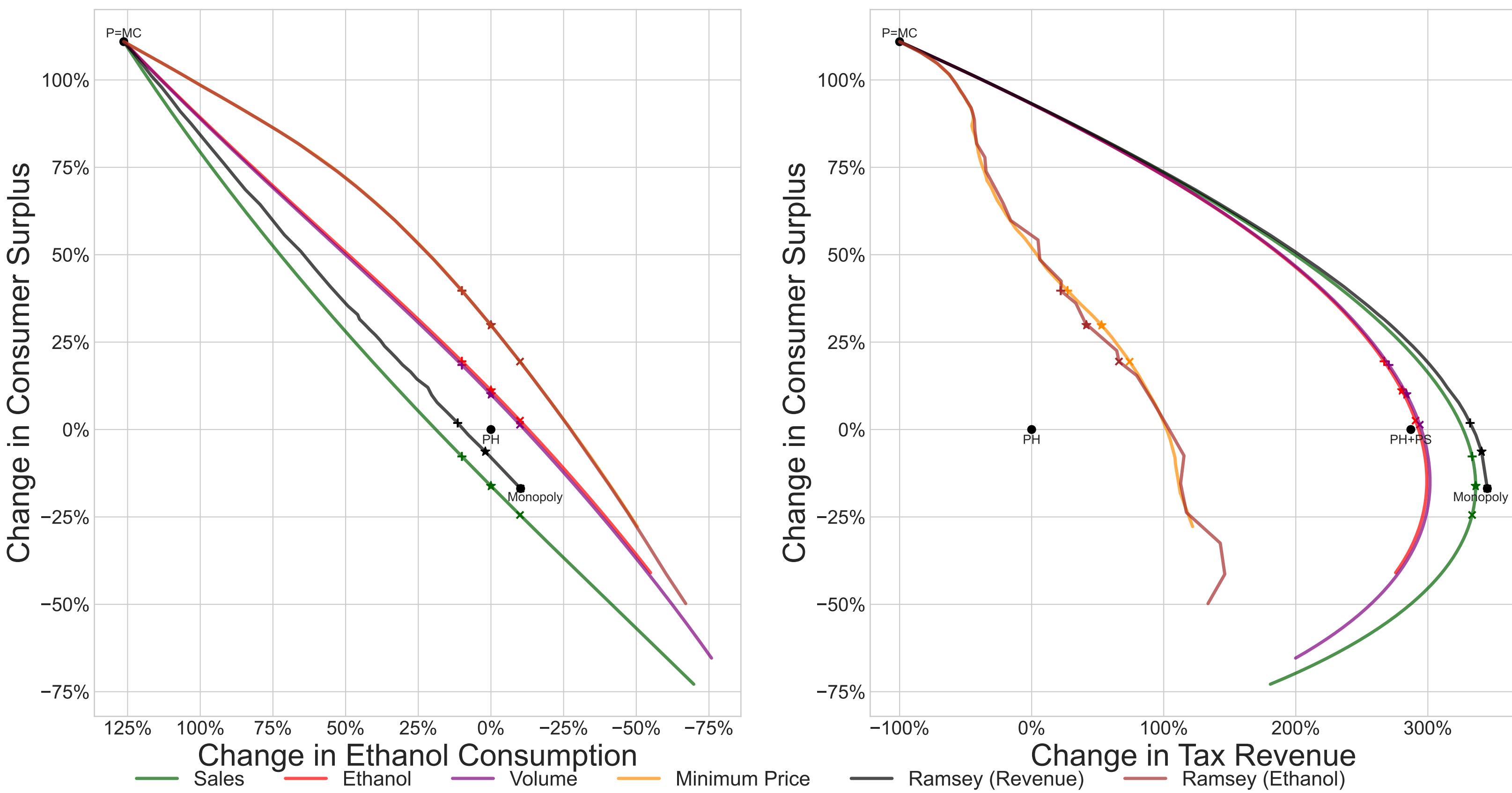
Note: The figure above plots own-price elasticities (left panel) against price where each observation is a product in 2013Q2. The center panel plots Diversion to the outside good, while the third panel plots Diversion away from ethanol. The products with the lowest prices have both the most elastic demands, and also the highest diversion away from ethanol indicating that taxing these products are the most effective at reducing ethanol consumption. Definitions: Elasticities: $e_{jj} = \frac{\partial s_j}{\partial p_j} \cdot \frac{p_j}{s_j}$; $D_{j\to 0} = \frac{\partial s_0}{\partial p_j} / \left|\frac{\partial s_j}{\partial p_j}\right|$; $D^e_{j\to 0} = e_j - \sum_{j\neq k} D_{j\to k} \cdot e_k$.
95% of sales volume are for products priced under $\$33/L$.
Source: Authors Calculations from demand estimates.

Figure 8: Prices Under PH vs. Other Policy Alternatives



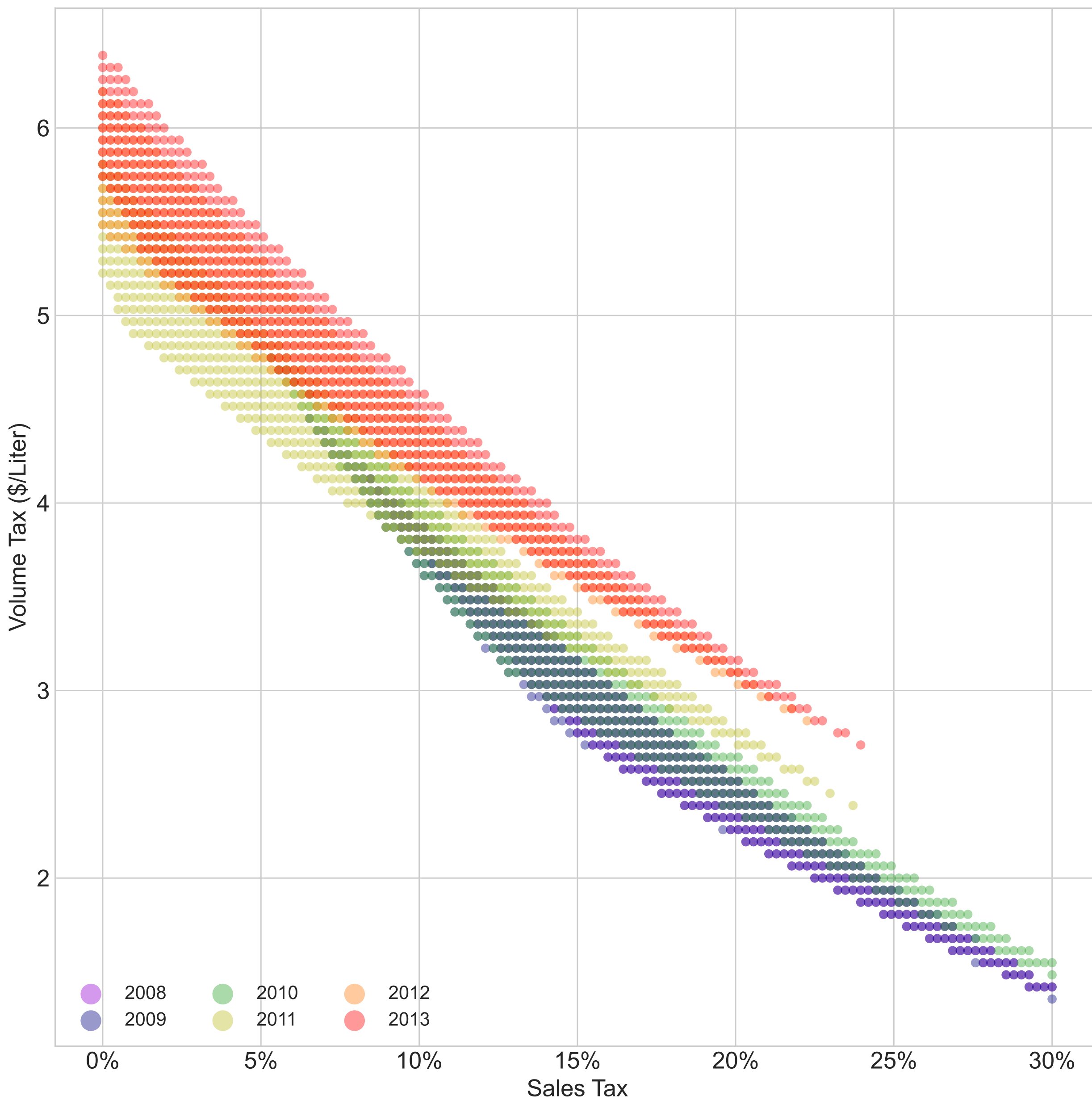
Note: The figure above plots product prices under PH against prices under our counterfactual policy alternatives. In each of our counterfactual scenarios we consider a tax rate that would keep the overall level of ethanol fixed at the status quo. Our taxes follow the definitions in Table 5, and are levied on a competitive market where wholesale price equals manufacturer price. The solid black 45-degree line illustrates prices unchanged from PH.

Figure 9: Consumer Surplus vs. Tax Revenue and Ethanol Consumption Under Alternative Policies



Note: The figure above plots the change tax revenue (left panel) and ethanol consumption (right panel) against the change in consumer surplus for each of the policy alternatives to PH detailed in Table 5. The frontiers trace the trade-off between consumer surplus and tax revenue or ethanol consumption for each policy instrument. The stars indicate an aggregate level of ethanol consumption equal to PH, while ($\times$) denotes 10% less and (+) denotes 10% more ethanol consumption (in the left panel, higher ethanol consumption corresponds to less tax revenue). We also mark competitive prices without taxes (denoted by $P = MC$) and PH pricing. In the left panel we indicate the revenue generate by existing excise taxes under PH pricing as well as the sum of tax revenue and wholesale profits generated by PH (denoted by PH+PS).

Figure 10: Combinations of Simple Taxes that Dominate PH



Note: Each dot signifies a point in the space of sales and volumetric tax rates (τ_s, τ_v) that dominates the PH outcome in all three categories: greater consumer surplus, lower ethanol consumption, greater tax revenue.
Our criteria for tax revenue is that it exceeds the amount that could be **extracted under lump sum taxation of wholesaler profits** under PH: $(PH + PS$ from Figure 9).
We calculate each point using the Q2 (April, May, June) estimates for that calendar year.
Source: Authors Calcuations.

Appendices

A. Additional Theoretical Results [For Publication]

A.1. Additional Theoretical PH Results

A.1.1. PH with a Single Product and Homogeneous Costs

We address this case in the main text and show that the first stage admits a dominant strategy of matching the lowest priced competitor so long as it is above your marginal cost.

Proof for Proposition 1

Consider a two-stage strategy of the form $\sigma_i(p_i^0, p_i^1)$. The second stage admits the unique dominant strategy where all players set $p_i^{1*} = \max\{c_i, \underline{p_i^0}\}$ where $\underline{p_i^0} = \min_i p_i^0$. For strategies of the form: $\sigma_i(p_i^0, \underline{p_i^0})$: $\sigma_i(p_i + \epsilon, \underline{p_i^0}) \geq \sigma_i(p_i, \underline{p_i^0})$ for $p_i \in [c_i, p_i^m)$ where $\geq$ denotes weakly greater profits. By induction the unique Nash Equilibrium to survive iterated weak dominance is $\sigma_i(p_i^m, \underline{p_i^0})$.

A.1.2. PH with a Single Product and Heterogeneous Costs

In the case of heterogeneous costs, the first stage becomes a bit more complicated. Begin by ordering the firms by marginal costs $c_1 \leq c_2 \cdots \leq c_N$. The market price $\hat{p}$ will be set by the lowest-cost firm (player 1). Other players play the iterated-weak-dominant-strategy $\sigma(p_i^0, p_i) = (p_i^m, \max\{\underline{p^0}, c_i\})$. Player 1 chooses p_i^0 to maximize the residual profit function:

$$\hat{p} = \underset{p_1^0 \in \{p_1^m, c_2, \ldots, c_n\}}{\arg\max} \pi_i(p_1^0) = \frac{(p_1^0 - c_1) \cdot Q(p_1^0)}{\sum_k \mathbb{I}[c_k \leq p_1^0]}$$

Player 1 can choose either to play its monopoly price and split the market evenly with the number of firms for which $c_i \leq p_1^m$, or it can set a lower price to reduce the number of firms who split the market. When the cost advantage of player 1 is small, we expect to see outcomes similar to the monopoly price. As the cost advantage increases, it becomes more attractive for player 1 to engage in limit-pricing behavior. Because our wholesalers buy the same products from the upstream manufacturer/distillers in roughly similar quantities, we ignore the possibility of heterogeneous marginal costs in our empirical example. In practice, as long as the dispersion between heterogeneous costs is not too large, firms will not have an incentive to engage in limit-pricing. Furthermore, adding firms will not necessarily lead to lower prices unless $c_{new} < c_1$.

A.1.3. Multiproduct Firms in Matrix Form

If we combine (3) and (5) and include the per-liter volumetric tax τ_v:

$$Q_j = -\frac{\partial Q_j}{\partial p_j^w}(p_j^w - mc_j - \tau_v) - \sum_{k \neq j} \kappa_{jk} \cdot \left[\frac{\partial Q_k}{\partial p_j^w}(p_k^w - mc_k - \tau_v)\right] = 0$$

We first re-write the wholesaler first-order conditions from (6) in matrix form:

$$\mathbf{q}(\mathbf{p^w}) = (\mathcal{H}_{PH}(\kappa) \odot \Delta(\mathbf{p^w}, \theta_2)) \cdot (\mathbf{p^w} - \mathbf{mc} - \tau_v).$$

The elements of the demand derivative matrix are given by $\Delta_{(j,k)} = -\frac{\partial Q_j}{\partial P_k}$ and $\odot$ denotes the Hadamard product. The entries of the ownership matrix are defined in (5), so that $\mathcal{H}_{(j,k)} = \kappa_{jk} = \gamma_k^f / \gamma_j^f$, which can be interpreted as profit weights or how the firm setting the price of j treats \$1 of (market-level) profit from k relative to \$1 of (market-level) profit from j. The profit weights depend on the relative share of the market controlled by f for products j and k. The profit weight for your own product is $\kappa_{jj} = 1$. Following a long literature in industrial organization, we can solve the linear system in (8) for the (additive) markups:[84]

$$\boldsymbol{\eta}(\theta_2) \equiv (\mathbf{p^w} - \mathbf{mc} - \tau_v) = (\mathcal{H}_{PH}(\kappa) \odot \Delta(\mathbf{p^w}, \theta_2))^{-1} \mathbf{q}(\mathbf{p^w}). \tag{A.1}$$

This gives us a definition for $\eta_{jt}(\theta_2, \mathcal{H}_{PH}(\kappa)$ to use in our supply moments:

$$p_{jt}^w - \eta_{jt}(\theta_2, \mathcal{H}_{PH}(\kappa)) = \underbrace{p_{jt}^m + \mathrm{w}_{jt} + \tau_v}_{mc_{jt}^w} + \omega_{jt}. \tag{A.2}$$

A.2. Solving the Planner's Problem

We could not find this derivation elsewhere in the literature, though we make no claims of originality here. We illustrate how one starts with the Lagrangian (7) and obtains (8).

A social planner solves a constrained optimization problem defined by demand $\mathbf{q}(\mathbf{p})$, and sets the prices $p_j \in \mathbf{p}$ of all products to maximize total surplus subject to two additional constraints: a minimum level of revenue $\overline{R}$, and a maximum level of externalities arising from ethanol consumption $\overline{E}$.

$$\begin{aligned} \max_{\mathbf{p}} \quad & CS(\mathbf{q}(\mathbf{p})) - C(\mathbf{q}(\mathbf{p})) \\ \text{subject to} \quad & \mathbf{p} \cdot \mathbf{q}(\mathbf{p}) - C(\mathbf{q}(\mathbf{p})) \geq \overline{R} \\ \text{and} \quad & E(\mathbf{q}(\mathbf{p})) \leq \overline{E}. \end{aligned} \tag{A.3}$$

where the social benefit of consumption is the same as the private benefit defined as the sum of the areas under the demand curves: $CS(\mathbf{q}(\mathbf{p})) = \sum_{k \in \mathcal{J}} \int_0^{q_k} p_k(q_1, q_2, ..., q_{k-1}, Z_k, q_{k+1}, ..., q_n)\, d\, Z_k$. The cost of producing alcoholic beverages is captured by $C(\mathbf{q}(\mathbf{p}))$. We can write the social planner's Lagrangian from (7):

$$\mathcal{L}(\mathbf{p}) = CS(\mathbf{q}(\mathbf{p})) - C(\mathbf{q}(\mathbf{p})) + \lambda_r(\mathbf{p} \cdot \mathbf{q}(\mathbf{p}) - C(\mathbf{q}(\mathbf{p})) - \overline{R}) - \lambda_e(E(\mathbf{q}(\mathbf{p})) - \overline{E}). \tag{A.4}$$

The Lagrange multiplier λ_r measures the social value of an additional dollar of revenue, while λ_e measures the shadow cost of an extra unit of external damage caused by alcohol consumption. A common assumption (though by no means necessary) is that the externality is *atmospheric*, or that it depends only on the total consumption of ethanol and not the source of the ethanol or the identity of the consumer, such that $E(\mathbf{q}(\mathbf{p})) = \sum_j (e_j - e_0) \cdot q_j(\mathbf{p})$, where e_0 is the external damage associated with the outside good, which could be zero or positive if the outside good were beer or wine. (Ideally $e_j > e_0$).

[84] See other examples from the IO literature going back to Bresnahan (1987) and Nevo (2001, 2000) for mergers, Villas-Boas (2007) for double marginalization, Miller and Weinberg (2017); Miller et al. (2021) for coordinated effects, and Backus et al. (2021a,b) for partial (common) ownership.

The first-order conditions return the two constraints and the derivative of the Lagrangian:

$$\frac{\partial \mathcal{L}}{\partial \lambda_r} : \quad \mathbf{p} \cdot \mathbf{q}(\mathbf{p}) - C(\mathbf{q}(\mathbf{p})) = \overline{R}$$

$$\frac{\partial \mathcal{L}}{\partial \lambda_e} : \quad E(\mathbf{q}(\mathbf{p})) = \sum_{k \in \mathcal{J}} e_k \cdot q_k(\mathbf{p}) + e_0 \cdot q_0(\mathbf{p}) = \overline{E}$$

$$\frac{\partial \mathcal{L}}{\partial p_j} : \quad \sum_{k \in \mathcal{J}} (p_k - mc_k) \frac{\partial q_k}{\partial p_j} + \lambda_r \left(q_j + \sum_{k \in \mathcal{J}} (p_k - mc_k) \frac{\partial q_k}{\partial p_j} \right) - \lambda_e \sum_{k \in \mathcal{J}} e_k \frac{\partial q_k}{\partial p_j} - \lambda_e e_0 \frac{\partial q_0}{\partial p_j} = 0.$$

Isolating product j, dividing through by $\frac{\partial q_j}{\partial p_j}$ and re-writing the expression in terms of the diversion ratio, $D_{j \to k} = -\frac{\partial q_k}{\partial p_j} / \frac{\partial q_j}{\partial p_j}$ and own price elasticity $\epsilon_{jj} = \frac{\partial q_j}{\partial p_j} \cdot \frac{p_j}{q_j}$ gives:

$$(1 + \lambda_r)(p_j - mc_j) - \lambda_r p_j \frac{1}{|\epsilon_{jj}|} - \lambda_e e_j - (1 + \lambda_r) \sum_{k \neq j} D_{j \to k} \cdot (p_k - mc_k) + \lambda_e \left[\sum_{k \neq j} D_{j \to k} \cdot e_k + D_{j \to 0} \cdot e_0 \right] = 0.$$

which can be solved for p_j as the social planner's pricing rule:

$$p_j = \frac{|\epsilon_{jj}|}{|\epsilon_{jj}| - \frac{\lambda_r}{1+\lambda_r}} \left(mc_j + \frac{\lambda_e}{1 + \lambda_r} \left[e_j - D_{j \to 0} \cdot e_0 \right] + \sum_{k \neq j} D_{j \to k} \left[p_k - mc_k - \frac{\lambda_e}{1 + \lambda_r} e_k \right] \right)$$

This illustrates that for each product, the marginal cost is effectively increased by the external damage term $\frac{\lambda_e}{1+\lambda_r} e_k$. It is helpful to re-arrange these terms to express diversion away from ethanol for product j as a single term:

$$D^e_{j \to 0} = e_j - D_{j \to 0} \cdot e_0 - \sum_{k \neq j} D_{j \to k} \cdot e_k \tag{A.5}$$

In the special case where external damage is similar across products (such as all products containing 40% alcohol by volume) $e_k \approx e$ for all k, we can approximate (A.5) as $D^e_{j \to 0} \approx D_{j \to 0} \cdot (e - e_0)$. For this reason we denote this term $D^e_{j \to 0}$ and label it "diversion away from ethanol".

We can simplify further with the following definitions by defining the "conduct parameter" $\theta = \frac{\lambda_r}{1+\lambda_r}$. This is a "conduct parameter" in the sense that $\theta = 0$ corresponds to perfect competition and $\theta = 1$ corresponds to monopoly. This gives us (8) in the main text:

$$p_j = \frac{1}{1 - \theta / |\epsilon_{jj}|} \left[mc_j + \sum_{k \neq j} D_{j \to k} \cdot (p_k - mc_k) + \frac{\lambda_e}{1 + \lambda_r} D^e_{j \to 0} \right]. \tag{8}$$

Absent any revenue constraint ($\lambda_r = 0$), the solution to the planner's problem is to set prices at their Pigouvian rates $p_k = mc_k + \lambda_e \cdot D^e_{j \to 0} + \sum_{k \neq j} D_{j \to k} \cdot (p_k - mc_k)$. Notice that the multi-product version depends not only on the ethanol content: e_j, but also ethanol content of diverted sales (including the outside option): $\sum_{\{k \neq j\} \cup \{0\}} D_{j \to k} \cdot e_k$. Also observe the planner's solution in (8) depends on the opportunity cost that arises from diversion to competing products (because we're maximizing *social surplus*). In the case where ($\lambda_r = \lambda_e = 0$) this still reduces to $p_j = mc_j$

because the opportunity cost term also vanishes.

A.2.1. Comparison Between PH and Planner

We can consider the difference between the prices under PH given in (6) and the prices under the planner's problem from (8). To simplify things we begin with the following definitions:

$$
\begin{aligned}
\mu_j(\theta) &= \frac{1}{1+\theta/\left|\epsilon_{jj}\right|} \\
\lambda_e^* &= \frac{\lambda_e}{1+\lambda_r} \\
UPP_j &= \sum_{k\neq j} D_{j\to k}\cdot(p_k - mc_k) \\
UPP_j^{PH}(\kappa) &= \sum_{k\neq j} \kappa_{jk}\cdot D_{j\to k}\cdot(p_k - mc_k) \\
\Delta UPP_j^{PH} &= UPP_j^{PH}(\kappa) - UPP_j \leq 0
\end{aligned}
$$

This allows us to write prices in terms of the markup and the effective marginal cost:

$$
\begin{aligned}
p_j(\lambda_r,\lambda_e) &= \mu_j(\theta)\cdot\left[mc_j + UPP_j + \frac{\lambda_e}{1+\lambda_r}\cdot D_{j\to 0}^e\right] \\
p_j^{PH} &= \mu_j^{PH}\cdot\left[mc_j + UPP_j^{PH}\right]
\end{aligned}
$$

So, the difference in prices can be expressed as:

$$
\begin{aligned}
p_j^{PH} - p_j(\lambda_e,\lambda_r) &= (\mu_j^{PH} - \mu_j(\theta))\cdot mc_j + \mu_j^{PH}\cdot UPP_j^{PH} - \mu_j(\theta)\cdot UPP_j - \mu_j(\theta)\cdot\lambda_e^*\cdot D_{j\to 0}^e \\
&= (\mu_j^{PH} - \mu_j(\theta))\cdot mc_j + \mu_j^{PH}\cdot(UPP_j + \Delta UPP_j^{PH}) - \mu_j(\theta)\cdot UPP_j - \mu_j(\theta)\cdot\lambda_e^*\cdot D_{j\to 0}^e \\
&= (\mu_j^{PH} - \mu_j(\theta))\cdot mc_j + (\mu_j^{PH} - \mu_j(\theta))\cdot UPP_j + \mu_j^{PH}\cdot\Delta UPP_j^{PH} - \mu_j(\theta)\cdot\lambda_e^*\cdot D_{j\to 0}^e \\
&= \underbrace{(\mu_j^{PH} - \mu_j(\theta))(mc_j + UPP_j)}_{>0} + \underbrace{\mu_j^{PH}\cdot\Delta UPP_j^{PH}}_{<0} - \underbrace{\mu_j(\theta)\cdot\lambda_e^*\cdot D_{j\to 0}^e}_{<0} \qquad \text{(A.6)}
\end{aligned}
$$

$$
\frac{p_j^{PH} - p_j(\lambda_e,\lambda_r)}{\mu_j(\theta)} = \underbrace{\left(\mu_j^{PH}/\mu_j(\theta) - 1\right)(mc_j + UPP_j)}_{>0} + \underbrace{\frac{\mu_j^{PH}}{\mu_j(\theta)}\cdot\Delta UPP_j^{PH}}_{<0} - \underbrace{\lambda_e^*\cdot D_{j\to 0}^e}_{<0} \qquad \text{(A.7)}
$$

The final line is divided by the preferred markup of the planner $\mu_j(\theta)$ which implies that the units in (A.7) are in terms of "marginal costs" instead of prices. This also lets us express things in terms of the relative markups $\mu_j^{PH}/\mu_j(\theta) = \frac{\epsilon_{jj}+\theta}{\epsilon_{jj}+1} \geq 1$. The relative markup is greater under PH for any $\theta < 1$, since under PH $\theta = 1$.

A.2.2. Comparison Between Tax Instruments and Planner

We perform a similar comparison for the simple tax instruments under perfect comeptition:

$$p_j(\tau) = (1+\tau_r)[mc_j + \tau_v + \tau_e\cdot e_j].$$

Consider how well simple taxes can approximate the planner's solution:

$$p_j(\theta, \lambda_e^*) = \mu_j(\theta) \cdot \Big[mc_j + UPP_j(\theta, \lambda_e^*) + \lambda_e^* \cdot D_{j\to 0}^e \Big].$$

One way to approach this is to ask, which tax rates $\tau = [\tau_r, \tau_v, \tau_e]$ minimize the difference between the planner's prices $p_j(\theta, \lambda_e^*)$ and those we can achieve with the simple tax measures $p_j(\tau)$?

$$\min_{\tau} \sum_j \Big[p_j(\theta, \lambda_e^*) - (1+\tau_r)\big(mc_j + \tau_v + \tau_e \cdot e_j\big)\Big]^2 \tag{A.8}$$

This suggests a simple (weighted) least-squares regression approach to recover τ where $\tau_r = \frac{1}{1+\alpha}, \tau_v = \frac{\beta_0}{\alpha}, \tau_e = \frac{\beta_1}{\alpha}$:[85]

$$p_j(\theta, \lambda_e^*) \sim \beta_0 + \beta_1 \cdot e_j + \alpha \cdot mc_j + \varepsilon_j \tag{A.9}$$

Thus, the simple tax measures will well approximate the planner's problem if the planner's prices can be explained with a linear regression onto the variables targeted by the tax, such as the ethanol content (because prices are per liter, volume corresponds to the constant). This gives us an easy way to see how well "tagging" products based on observable characteristics will approximate the planner's solution. We can also estimate coefficients of the form $\beta_1(e_j)$ as piecewise-linear splines if, for example, we want to allow for nonlinear taxation of ethanol content. Doing so would provide a sense of how much nonlinear taxation would improve upon a single tax rate. The drawback of this approach is that we need to know the planner's preferences (θ, λ_e^e) and calculate the planner's prices $p_j(\theta, \lambda_e^*)$ in order to calculate the dependent variable in the regression above.

Alternatively, we could try to gain some intuition from a heuristic solution to (A.8) after rescaling by the planner's inverse elasticity markup $\mu_j(\theta)$ like in (A.7):

$$\frac{p_j(\tau) - p_j(\lambda_e^*, \theta)}{\mu_j(\theta)} = \underbrace{\left(\frac{(1+\tau_r)}{\mu_j(\theta)} - 1\right)}_{?} \cdot mc_j - \underbrace{UPP_j}_{<0} - \underbrace{\lambda_e^* \cdot D_{j\to 0}^e}_{<0} + \underbrace{(1+\tau_r)(\tau_v + \tau_e \cdot e_j)}_{>0} \tag{A.10}$$

In the heuristic solution we could apply the "principle of targeting" and try to use one tax instrument — an excise tax on volume τ_v or ethanol τ_e — to address the external damage on average and another — the sales tax τ_r — to address the revenue constraint on average. This amounts to breaking up (A.10) into two parts and setting both parts to zero *on average*. This is at best an approximation to (A.8) because across multiple products j, the terms in (A.10) are likely to be correlated with one another and a better solution would (like the regression approach in (A.9)) account for this covariance.

There are two obvious possibilities. In the first, we find τ_r to set the first term to zero on average and the excise tax τ_v to set the last three terms to zero on average.[86]

$$(1+\tau_r) \approx \frac{1}{\mathbb{E}_{mc}\left[\frac{1}{\mu_j(\theta)}\right]} = \frac{1}{1 + \theta \cdot \mathbb{E}_{mc}\left[\frac{1}{\epsilon_{jj}}\right]}, \quad \tau_v \cdot (1+\widehat{\tau_r}) \approx \cdot \frac{1}{J} \sum_{j \in \mathcal{J}} \Big(\lambda_e^* \cdot D_{j\to 0}^e + UPP_j \Big). \tag{A.11}$$

[85] We may want to allow for weights in our least squares problem, an obvious choice might be the quantity sold.

[86] We could use $\tau_e \cdot e_j$ in lieu of τ_v but we know from Figure 8 they work almost identically so we illustrate with τ_v.

In this case, the sales tax is set to match the (marginal cost weighted) harmonic average of the markups (or the inverse elasticities), and the volumetric tax corrects for the average external damage term and opportunity cost from multi-product pricing UPP_j.

However, $UPP_j = \sum_{k \neq j} D_{j \to k}(p_k - mc_k)$ tends to be higher for products with higher prices/marginal costs (because of the substitution patterns in Table 4, and the correlation coefficient is 0.87). So instead we may want to group the first two terms and the last two terms together and find the heuristic solution:

$$(1+\tau_r) \approx \frac{\sum_j (mc_j + UPP_j)}{\sum_j mc_j/\mu_j(\theta)} = \frac{1 + \frac{\overline{UPP}}{\overline{mc}}}{\mathbb{E}_{mc}\left[\frac{1}{\mu_j(\theta)}\right]}, \quad \tau_v \cdot (1 + \widehat{\tau}_r(\theta)) \approx \lambda_e^* \cdot \frac{1}{J} \sum_{j \in \mathcal{J}} D^e_{j \to 0}. \tag{A.12}$$

This heuristic solution sets a higher sales tax rate than (A.11) because it uses the sales tax (instead of the excise tax) to correct for the opportunity cost by augmenting the previous solution with the average UPP_j over the average mc_j term, but a lower volumetric tax rate for the same reason. One familiar feature is that we get the expected "Pigouvian" correction where we set the volumetric tax equal to the average of the marginal external damage terms (scaled by the marginal damage per unit of ethanol λ_e^*). As always $D^e_{j \to 0}$ depends not only on the ethanol content of product j but on how much ethanol consumption is reduced when consumers substitute away. Also note that in (A.12) even at $\theta = 0$ (no revenue constraint) the planner still sets $\tau_r \approx \frac{\overline{UPP}}{\overline{mc}}$ (which works out to about 14% in our data) to correct for the opportunity cost of selling other products.

The heuristic solutions in (A.11) and (A.12) and even the regression solution in (A.9) can only match the planner's problem on average, and are likely to approximate the planner's solution best when there is not too much variation across products. This will be the case if there is not too much dispersion in the elasticities ϵ_{jj}, the external damage terms $D^e_{j \to 0}$ (see Figure 7), and the opportunity cost of selling other products UPP_j. We know that there is a large amount of dispersion in the marginal costs of wholesalers $mc_j = p_j^m + \tau_j^v$ in the data, which can be partially addressed with sales taxes in (A.9).

A.3. Recovering Manufacturer Marginal Costs

In Table 7, we allow multi-product distillers/manufacturers (e.g. Bacardi, Diageo) to adjust their prices. We also report estimated manufacturer costs in Table 1. These require estimates not only of *manufacturer prices* and the manufacturer ownership matrix $\mathcal{H}_M$, which we observe, but also of *manufacturer marginal costs* which we do not.

This next part builds on Jaffe and Weyl (2013) and Appendix E from Miller and Weinberg (2017) and almost exactly follows the implementation in Backus et al. (2021a); Conlon and Gortmaker (2020). The wrinkle here is that we observe the manufacturer prices $\mathbf{p^m}$ which simplify matters considerably, and we have the addition of the existing excise tax $\tau_{v,0}$, which we show does not create any new issues.

We write the manufacturer's first order conditions as:[87]

$$\mathbf{p^m} = \mathbf{mc^m} + \left(\mathcal{H}_M \odot \left(\frac{\partial \mathbf{p^w}}{\partial \mathbf{p^m}} \cdot \Delta(\mathbf{p^w})\right)\right)^{-1} \mathbf{q}(\mathbf{p^w}) \tag{A.13}$$

This requires that we estimate the pass-through matrix $\frac{\partial \mathbf{p^w}}{\partial \mathbf{p^m}}$.

In order to do so, we re-examine the wholealers' problem: a system of J first order conditions and J prices $\mathbf{p^w}$, with manufacturer prices $\mathbf{p^m}$ and wholesaling costs (including taxes) τ_0 serving as parameters:[88]

$$f(\mathbf{p^w}, \mathbf{p^m}, \tau_{v,0}) \equiv \mathbf{p^w} - \underbrace{(\mathbf{p^m} + \tau_{v,0})}_{=\mathbf{mc^w}} - (\underbrace{\mathcal{H}_{PH}(\kappa) \odot \Omega(\mathbf{p^w})}_{\equiv \Omega(\mathbf{p^w})})^{-1} \mathbf{q}(\mathbf{p^w}) = 0 \tag{A.14}$$

Where $\Omega(\mathbf{p^w}) \equiv \mathcal{H}_{PH} \odot \Delta(\mathbf{p^w})$ is the PH augmented matrix of demand derivatives.

We differentiate the wholesalers' system of FOC's with respect to p_l, to get the $J \times J$ matrix with columns l given by:

$$\frac{\partial f(\mathbf{p^w}, \mathbf{p^m}, \tau_0)}{\partial p^w_\ell} \equiv e_\ell - \Omega^{-1}(\mathbf{p^w}) \left[\mathcal{H}_{PH} \odot \frac{\partial\, \Omega(\mathbf{p^w})}{\partial\, p^w_\ell}\right] \Omega^{-1}(\mathbf{p^w})\, \mathbf{s}(\mathbf{p^w}) - \Omega^{-1}(\mathbf{p^w})\, \frac{\partial \mathbf{s}(\mathbf{p^w})}{\partial p^w_\ell}. \tag{A.15}$$

The complicated piece is the demand Hessian: a $J \times J \times J$ tensor with elements (j, k, ℓ), $\frac{\partial^2 s_j}{\partial p^w_k \partial p^w_\ell} = \frac{\partial^2 \mathbf{s}}{\partial \mathbf{p^w} \partial p^w_\ell} = \frac{\partial\, \Delta(\mathbf{p^w})}{\partial\, p^w_\ell}$.

We can follow Jaffe and Weyl (2013) and apply the multivariate IFT. The multivariate IFT says that for some system of J nonlinear equations $f(\mathbf{p^w}, \mathbf{p^m}, \tau_0) = [F_1(\mathbf{p^w}, \mathbf{p^m}, \tau_0), \ldots, F_J(\mathbf{p^w}, \mathbf{p^m}, \tau_0)] = [0, \ldots, 0]$ with J endogenous variables $\mathbf{p^w}$ and J exogenous parameters $\mathbf{p^m}$.

$$\frac{\partial \mathbf{p^w}}{\partial \mathbf{p^m}} = -\begin{pmatrix} \frac{\partial F_1}{\partial p^w_1} & \cdots & \frac{\partial F_1}{\partial p^w_J} \\ \cdots & \cdots & \cdots \\ \frac{\partial F_J}{\partial p^w_1} & \cdots & \frac{\partial F_J}{\partial p^w_J} \end{pmatrix}^{-1} \cdot \underbrace{\begin{pmatrix} \frac{\partial F_1}{\partial p^m_k} \\ \cdots \\ \frac{\partial F_J}{\partial p^m_k} \end{pmatrix}}_{=-\mathbb{I}_J} \tag{PTR}$$

Because the system of equations is additive in $\mathbf{p^m}$ and τ_0 this simplifies dramatically $\frac{\partial f(\mathbf{p^w}, \mathbf{p^m}, \tau_0)}{\partial \mathbf{p^m}} = -\mathbb{I}_J$. The pass-through matrix (PTR) is merely the inverse of the matrix whose columns are defined in (A.15).[89]

In the counterfactual world, with competitive wholesaling, the pass-through matrix reduces to

[87] With some additional modifications, we could follow Miller and Weinberg (2017) and interpolate between no manufacturer response and the fully flexible manufacturer response. We find that the two outcomes are not far enough apart for this to matter.

[88] Because the marginal costs are additively separable we can also define the system as $f(\mathbf{p}, 0, 0) + \mathbf{c} + \tau_{v,0} = 0$.

[89] Our average product-level own pass-through rate is 1.3 which is *overshifted*, but consistent with regression estimates in our prior work (Conlon and Rao, 2020).

the identity matrix plus any ad valorem taxes $\frac{\partial \mathbf{p^w}}{\partial \mathbf{p^m}} = I_J \cdot (1 + \tau_r)$, while the effective marginal cost becomes: $\mathbf{c_m} + \tau_v + \tau_e \cdot \mathbf{e}$, where τ_v are any per-unit taxes, and τ_e are any ethanol taxes.

Implmentation Notes:

1. PyBLP method `compute_passthrough()` will deliver (PTR) (this is very time consuming).
2. PyBLP method `compute_demand_jacobians()` will deliver $\Delta(\mathbf{p^w})$.
3. $\mathcal{H}_m$ is the ownership matrix at the manufacturer level (ie: 1's if both products are owned by Diageo, Bacardi, etc.).
4. $\mathbf{s_t}$ are observed shares and we can plug into (A.13) to get $\mathbf{mc^m}$.
5. Because $\mathbf{mc^m}$ is backed out of (A.13) it is the combination of production costs and federal excise taxes. We never need to separate the two for any counterfactuals.
6. Once we recover $\mathbf{c_m}$, we can re-solve (A.13) for the optimal manufacturer prices $\mathbf{p^m}(\mathbf{mc^m} + \tau)$ at each proposed level of taxes. PyBLP method `compute_prices()` will work fine using $\mathcal{H}_m$ and the tax-augmented marginal cost.[90]

[90] This works because excise and volumetric taxes are independent of prices and the fact that statutory incidence is downstream of manufacturers is irrelevant.

B. Empirical Implementation Details [Online Only]

B.1. Micro Moments

B.1.1. Demographic Interactions

We borrow notation from Conlon and Gortmaker (2025) to match the micro moment implementation in `PyBLP`. All micro moments take the following form, where we match $\overline{v}_{m,year}(\theta_2) = \frac{1}{4}\sum_{t\in\mathcal{T}_{year}} v_{mt}(\theta_2)$ with the model simulated analogue by averaging over the quarters in each calendar year $\mathcal{T}_{year}$. We use the same number of Monte Carlo draws in each market t so that $w_{it} = \frac{1}{I}$ and the general formula simplifies:

$$v_{mt}(\theta_2) = \frac{\sum_{i\in\mathcal{I}_t}\sum_{j\in J_t\cup\{0\}} \sigma_{ijt}(\theta_2)\cdot w_{d_m ijt}\cdot v_{mijt}}{\sum_{i\in\mathcal{I}_t}\sum_{j\in J_t\cup\{0\}} \sigma_{ijt}(\theta_2)\cdot w_{d_m ijt}} \tag{B.1}$$

Where $w_{d_m ijt}$ are the survey weights and v_{mijt} is the value. Notice that only the individual choice probabilities $\sigma_{ijt}(\theta_2)$ vary with the parameters θ_2. We match the following moments, where y_i denotes the individual's income, and $\mathcal{I}_k$ denotes each of our five income bins, and the event where a purchase is made is denoted by not selecting the outside option ($j \neq 0$):

1. $w_{dijt} = 1\{j \neq 0\}$ and $v_{mijt} = 1\{\mathrm{y}_i \in \mathcal{I}_k\}$ for each market $t \in T$ and "inside" goods only. This allows us to match:
$$\mathbb{P}\left[\mathrm{y}_i \in \mathcal{I}_k \mid j \neq 0\right]$$
2. $w_{dijt} = 1\left\{j \neq 0, \mathrm{x}_j = 750mL\right\}$ and $v_{mijt} = 1\{\mathrm{y}_i \in \mathcal{I}_k\}$ for each market $t \in T$ and "inside" goods only. This allows us to match:
$$\mathbb{P}\left[\mathrm{y}_i \in \mathcal{I}_k \mid \mathrm{x}_j = 750mL\right]$$
3. $w_{dijt} = 1\left\{j \neq 0, \mathrm{x}_j = 1750mL\right\}$ and $v_{mijt} = 1\{\mathrm{y}_i \in \mathcal{I}_k\}$ for each market $t \in T$ and "inside" goods only. This allows us to match:
$$\mathbb{P}\left[\mathrm{y}_i \in \mathcal{I}_k \mid \mathrm{x}_j = 1750mL\right]$$
4. $w_{dijt} = 1\{j \neq 0, \mathrm{y}_i \in \mathcal{I}_k\}$ and $v_{mijt} = p^w_{jt}$ for each market $t \in T$ and "inside" goods only. This allows us to match:
$$\mathbb{E}\left[p^w_{jt} \mid \mathrm{y}_i \in \mathcal{I}_k \text{ and } j \neq 0\right]$$

We match a different set of values for each income bin. To avoid colinearity (probabilities sum to one) we exclude the middle income bin for the first three sets of moments. We match a different set of moments for each year $\mathcal{T}_{year}$ from 2007-2013, rather than each *market* (a quarter). This is because the NielsenIQ Household Panelist data samples different households each year (and uses different projection weights).

These moments are straightforward to calculate from the NielsenIQ Household Panelist data, and don't require any other data sources beyond the NielsenIQ data. The exception is that for each product, NielsenIQ reports the *retail price* and we must find the corresponding *wholesale* *price* because the model is defined in terms of *Wholesale Demand*. This requires looking up the corresponding wholesale price (per liter) for each purchase we observed purchased at retail in the NielseIQ Panelist data.

We report an aggregated version of our demographic moments in Table B.1 (averaged across years). This is meant to highlight the patterns in the data that discipline the Π parameters, and approximate the goodness of fit. As an example, we do a good job matching the distribution of income conditional on purchase, and conditional on purchasing a larger product, though we struggle a bit to capture the demand from the lowest income group. Because this group is so small, the GMM weighting matrix ends up placing a very small weight on matching the behavior of the lowest income group.

We tend to consistently over-estimate the average price paid by each income group because the distribution of prices (even conditional on income) of purchases by NielsenIQ panelists is significantly lower than the overall distribution of prices in the shipment data. This is the *compatability* issue raised in Conlon and Gortmaker (2025). In general, we get the correlation between income and price paid correct, although the levels in the NielsenIQ data would be impossible to match given the overall market shares observed in our data. The GMM estimator tries to miss each moment by a similar amount (weighted by the variance of the moment). In our case the problem is less acute because we rely on matching the average markup from the supply moments to get the price sensitivity correct.

Another example of *compatibility* is that we worry that the fraction of 1.75L bottles purchased by households in the NielsenIQ Panelist data is significantly higher than the fraction of 1.75L bottles (by volume) in the shipment data. (Bars and restaurants tend not to use 1.75L bottles and prefer 750mL or 1L bottles). Thus the marginal distribution of $\mathbb{P}(\mathrm{x}_{jt} = 1.75L)$ is not the same across the two datasets, and we instead use a moment that conditions on the purchase of a bottle size, rather than the expectation $\mathbb{E}[\mathrm{x}_{jt} \cdot \mathrm{y}_i \mid \text{ purchase }]$. Calculating moments like these (or moments not conditional on purchase) also require additional assumptions on trip frequency and potential purchase opportunities. This leads us to prefer the more robust but less efficient expectation of *income* conditional on purchasing a large bottle $\mathrm{x}_j = 1.75L$.

Income	$\mathbb{P}$(Income|Purchase)	Estimated	$\mathbb{P}$(Income|1750)	Estimated	$\mathbb{E}[p^w_{jt}\|\text{Income}]$	Estimated
Below $25k	0.22	0.17	0.20	0.21	9.97	12.56
$25k-$45k	0.14	0.12	0.13	0.16	11.91	13.42
$45k-$70k	0.17	0.17	0.18	0.23	12.39	13.35
$70k-$100k	0.12	0.13	0.12	0.15	13.66	14.70
Above $100k	0.35	0.41	0.36	0.25	17.98	22.17

Table B.1: Micro Moment Fit

We examine the possibility of including other consumer demographics in y_i. We don't see enough Black or Hispanic households purchasing spirits in Connecticut to accurately estimate micro-moments in these sub-populations. The age of the head of household doesn't seem to vary in a meaningful way with any of the product characteristics in our data, and education is highly correlated with income.[91]

[91] See Conlon et al. (2024) for an in-depth examination of the interaction between household demographics and purchases of sin goods. In the national sample, we find households over 55 are more likely to be heavy consumers of distilled spirits, though that is less evident in the Connecticut data.

B.1.2. Second-Choice Moments

These moments are relatively straightforward to define in `PyBLP`, and the construction of the moments from the NielsenIQ dataset is described in detail in the text of the paper. We provide the implementation details below which closely follow (Conlon and Gortmaker, 2025).

We define $w_{dijkt} \propto \mathcal{M}_t \cdot 1\{j, k \neq 0\}$ which corresponds to a random sample of consumers whose first and second choices were both inside alternatives. We then define two parts a $v_{ijkt}^{top}(\theta) = 1\{j \in \mathcal{J}_g \text{ and } k \in \mathcal{J}_c\}$ and $v_{ijkt}^{bottom}(\theta) = 1\{k \in \mathcal{J}_c\}$ where $\mathcal{J}_c$ are the set of products in the category (such as "Vodka") and define the moment as the ratio of the two micro-moment parts: $f(\theta_2) = v^{top}(\theta_2)/v^{bottom}(\theta_2)$.

We use (Eq 21) from Conlon and Gortmaker (2025) in place of (B.1) to define the "parts":

$$v^p(\theta_2) = \frac{\sum_{t \in \mathcal{T}} \sum_{i \in \mathcal{I}_t} \sum_{j \in \mathcal{J}_t \cup \{0\}} \sum_{k \in \mathcal{J}_t \cup \{0\} \setminus \{j\}} w_{it} \cdot \sigma_{ijkt}(\theta_2) \cdot w_{d_p ijkt} \cdot v_{ijkt}^p}{\sum_{t \in \mathcal{T}} \sum_{i \in \mathcal{I}_t} \sum_{j \in \mathcal{J}_t \cup \{0\}} \sum_{k \in \mathcal{J}_t \cup \{0\} \setminus \{j\}} w_{it} \cdot \sigma_{ijkt}(\theta_2) \cdot w_{d_p ijkt}}. \tag{B.2}$$

All that remains is to define $\sigma_{ijkt}(\theta_2)$ (the probability that an individual will have first-choice product and j will have second choice product k). This is easy to compute within the model.

The idea is that for every pair of products (j, k) we compute the joint probability that j is first-choice and k is the second $\sigma_{ijkt}(\theta_2)$ given the parameters and then the v_{ijkt}^p are simply indicator functions for whether both products are Vodkas (top) or the first-choice product is Vodka (bottom). We can repeat this for each of the product categories. We illustrate how these moments are used to estimate ρ in Appendix D.3.

Constructing the target values from the data is actually more involved than calculating the predicted second choice values under the model. We use the NielsenIQ panelist data and look at households in Connecticut that purchase multiple spirits products and (unlike the demographic moments) pool across all years (2007-2013). We consider an experiment where we reshuffle the purchases made by an individual household and then ask if product j and product k are distinct products, does k come from the same category as j? or a different category?

As an example, suppose we see a household make five purchases: Smirnoff (Vodka), Smirnoff (Vodka), Smirnoff (Vodka), Tanqueray (Gin), Absolut (Vodka). We could conclude that the same category repurchase rate is 0.5 for Smirnoff (Vodka) and 0.75 for Absolut (Vodka), so that when we weight by initial purchase frequency:[92]

$$\mathbb{P}(k \in \mathcal{J}_{\text{Vodka}} \setminus \{j\} \mid j \in \mathcal{J}_{\text{Vodka}}) = \frac{1}{4} \cdot 0.75 + \frac{3}{4} \cdot 0.5 = 0.5625.$$

We construct these moments by pooling across all households and years. We estimate the repurchase rates as: Vodka (0.51), Gin (0.60), Rum (0.20), NA Whiskey (0.26), UK Whiskey (0.33), which are significantly higher than the unconditional market shares of the corresponding categories, and indicate a value of ρ that is significantly greater than zero.[93]

B.2. Calculating Ownership κ_{jk}

The PH first-order condition in (6) depends on κ_{jk} the "profit-weight" which captures how in equilibrium, the lowest opportunity cost wholesaler trades off \$1 of profit from j against \$1 of

[92] Trivially the repurchase rate for Gin is zero because only one gin is purchased.

[93] We omit the Tequila category because we don't see enough purchases by Connecticut households to estimate a repurchase rate.

profit from k. The identity of the marginal wholesaler is determined by (5). We have data on the volume of shipments from each manufacturer to each wholesaler, which we aggregate by calendar year. As an example, we observe how many units of Johnnie Walker Black (1750mL) are shipped from Diageo to: Eder Bros, Hartley Parker, Alan S. Goodman, and Brescome Barton. We take the annual shipments to each wholesaler in that calendar year and divide them by the total shipments of that product, which becomes γ_j^f. We repeat this for all products and wholesalers. Because not every product is sold by every wholesaler, many elements of γ_j^f are zero. Given some preliminary demand estimates $\widehat{\theta}$ we can compute the diversion ratios $D_{j\to k}(\widehat{\theta})$ and find the minimum in (5) the assumption that $mc_{jt} = p_{jt}^m + \tau_{vt} + \mathrm{w}_j$. Once we've identified the lowest opportunity cost firm we can simply evaluate $\kappa_{jk} = \frac{\gamma_k^f}{\gamma_j^f}$ for that firm.

Once we have calculated κ_{jk} for each pair of products, we can construct the "ownership matrix" $\mathcal{H}(\kappa)$ with elements κ_{jk} and use that to evaluate the first-order conditions and calculate the implied markups $\eta_{jt}(\theta)$ using the linear system (A.1). In practice, we repeat this exercise to ensure that κ_{jk} does not change as we update our estimate of $\widehat{\theta}$, which turns out to be not an issue.

We report some summary statistics for κ_{jk} below. We can see that for 78% of cases $\kappa_{jk} = 0$, which means that the lowest opportunity cost wholesaler for product j does not sell any units of product k in that calendar year. Likewise, $\kappa_{jk} = 1$ for 14.3% of cases which usually means that a single wholesaler sells products (j, k) (often different sizes or flavor of the same brand). For the remaining products several have $0 < \kappa_{jk} < 1$ which means that the marginal wholesaler treats a unit of profit from k as worth less than a unit of profit from j (but more than zero). In several cases $\kappa_{jk} > 1$. This arises when firm f captures a small share of sales from product j like $\frac{1}{4}$ and a larger share of sales from k (such as being the sole seller). In this case $\kappa_{jk} = 4$, and this increases the opportunity cost for the wholesaler. Theoretically, this creates the possibility that if a new wholesaler starts to sell product j, it can paradoxically lead to higher prices, depending on how γ and κ respond. (We see very few changes in the products that are distributed by each wholesaler, and do not model this in the paper).

	Share of κ_{jk}
Zero	78.1
(0.001, 0.5]	1.4
(0.5, 1)	2.3
1	14.3
(1.0, 1.5]	1.2
(1.5, 2.0]	1.2
(2.0, 3.0]	0.8
(3.0, 4.0]	0.3
(4.0, ∞)	0.4

Table B.2: Distribution of κ_{jk}

C. Additional Comparisons of PH and Planner [Online Only]

C.1. Comparisons Across Products

Note: All of the discussion around product-level pricing is at the wholesale level.

We provide some further details on the comparisons between PH and our tax policy alternatives. All of these focus on Q2 2013, the final period in our sample, and the one we use in our counterfactual results. In Figure 5 we showed that the (additive) wholesale markup η_j in $p_j^w = p_j^m + \tau_v + \eta_j$ varies by product and that the markup is generally increasing in p_j^m. Eliminating PH eliminates this markup and replaces it with a tax on ethanol (or volume) $p_j^* = p_j^m + \tau_e \cdot e_j$. This means products with the largest markups will see the largest price reductions when we switch from PH to a competitive market and a tax, while the products with the lowest markups under PH will likely see prices increase. We compare the prices under PH to counterfactual taxes (that hold aggregate ethanol constant) in Figure 8, and the general trend is that prices increase for the least expensive products $p_j^w < \$10/L$ and decrease for more expensive products (particularly those where wholesale prices exceed \$30/L). There are of course exceptions to the general trend, and we provide detailed product level evidence below.

In Table C.1, we provide a more detailed breakdown of the top 15 products that gain (lose) the most sales when we replace PH with an ethanol tax (which keeps aggregate consumption of ethanol constant). One advantage of this exercise in our data is that we observe the prices under PH for wholesalers p^{old} and manufacturers p_{old}^m and thus do not need to estimate them.

The largest beneficiary of the elimination of PH is Smirnoff Vodka in a 1.75L bottle. This is the best-selling product in Connecticut (and nationwide). The product is an outlier in the sense that it has a relatively low price under PH $p^w = \$11.85/L$, but a very large wholesale markup ($\eta_j = \$6.10/L$). Replacing that markup η_j and the existing volumetric tax $\tau_v = \$1.42/L$ with a new tax of $\tau_v = \tau_e \cdot e_j = \$5.48/L$ leads to a new price of \$9.81 which is more than $\$2.00/L$ lower. Allowing upstream manufacturers to adjust prices leads to only minor changes. We see a similar pattern for Jack Daniels (750mL), the most popular American Whiskey brand, wholesale prices under PH are $p_j^w = \$29.21/L$ and the wholesaler markup is $\eta_j = \$9.73/L$. Replacing the markup (and the existing volumetric tax) with a tax of $\$5.48/L$ leads to significantly lower prices and much higher sales.

Similarly, we see that the products with the lowest markups under PH see prices rise. The case of Dubra Vodka is instructive. It is the least expensive source of ethanol under the PH system at $p_j^w = \$5.88/L$ with a manufacturer price of only $\$3.98/L$ (and a tax $\tau_v = \$1.42/L$) which implies a tiny wholesaler markup of only $\eta_j = \$0.48/L$. Eliminating this markup and replacing it with a $\$5.48/L$ tax significantly increases the wholesale price to $\$9.46/L$. At this price, it is only $\$0.35/L$ less than Smirnoff, and essentially all consumers switch to Smirnoff. Indeed, most of the largest losers are the least expensive Vodka brands (in 1.75L bottles).

The other case that is instructive is the 1.75L bottle of Captain Morgan Spiced Rum. Under PH, the wholesale price was $\$15.85/L$ and the markup was $\eta_j = \$2.74/L$. Once again, eliminating this markup and replacing it with a $5.48/L$ tax will lead to higher prices and lower sales. However, in this case, much of the substitution is captured by *other sizes of Captain Morgan's Rum* found in the table of "Winners" above. Part of what we are doing by eliminating the wholesaler market power under PH, is limiting some of the second-degree price discrimination that wholesalers engage in (setting higher unit prices on smaller bottles). Manufacturer differences in unit prices for Captain Morgan are significantly smaller (\$11.69 vs \$14.70 per liter) when compared to the wholesale differences under PH (\$15.85 vs. \$23.44 per liter).

Product	p^{old}	p^{new}	p^{new}_{endog}	p^{m}_{old}	p^{m}_{endog}	mc^m	$D^e_{j\to 0}$
Biggest Winners							
Smirnoff 1.75 L	11.85	9.81	9.56	4.33	4.53	2.00	19.50
Tlmr Dw I-W 12Y 0.75 L	45.21	22.01	17.62	16.53	12.59	8.47	18.08
Jack Daniel Black Label 1.0 L	27.08	22.89	23.00	17.42	17.97	12.26	16.28
Jack Daniel Black Label 0.75 L	29.21	23.53	23.33	18.06	18.31	12.50	16.28
Malibu 1.0 L	20.85	16.76	16.55	13.88	13.92	10.17	-2.28
Capt Morgan Spiced 1.0 L	21.79	17.89	17.77	13.10	13.37	8.84	11.99
Svedka 1.75 L	13.09	12.14	11.89	6.66	6.87	4.54	18.14
Grey Goose 0.75 L	39.88	33.36	33.52	27.88	28.50	20.50	18.28
Jack Daniel Black Label 1.75 L	21.85	20.33	20.55	14.86	15.53	10.99	16.46
Smirnoff 0.75 L	19.88	17.01	17.08	11.53	12.06	7.60	16.01
Bacardi Superior Lt Dry Rum 0.75 L	18.55	16.24	16.17	10.76	11.15	7.76	16.98
Cuervo Gold 1.0 L	21.32	18.98	19.00	13.50	13.97	9.13	16.38
Makers Mark 1.0 L	32.79	28.32	28.69	22.16	23.04	15.93	21.81
Absolut Vodka 1.0 L	24.91	22.09	22.46	16.61	17.44	11.90	16.38
Capt Morgan Spiced 0.75 L	23.44	19.49	19.57	14.70	15.17	10.09	12.02
Biggest Losers							
Product	p^{old}	p^{new}	p^{new}_{endog}	p^{m}_{old}	p^{m}_{endog}	mc^m	$D^e_{j\to 0}$
Bacardi Superior Lt Dry Rum 1.75 L	12.52	14.22	14.37	8.74	9.35	6.65	20.62
Absolut Vodka 1.75 L	15.94	17.70	18.52	12.22	13.50	9.48	17.58
Popov Vodka 1.75 L	7.66	11.12	12.26	5.64	7.23	4.16	22.22
Grays Peak Vdk Dom 1.75 L	9.16	12.66	13.25	7.18	8.22	5.71	20.57
Dubra Vdk Dom 80P 1.75 L	5.88	9.46	9.73	3.98	4.71	2.89	24.49
Smirnoff Raspberry Vodka 1.75 L	10.23	13.25	14.47	8.46	10.07	6.54	14.63
Sobieski Poland 1.75 L	9.09	11.10	11.29	5.62	6.27	4.13	20.65
Skyy Vdk Dom 1.75 L	12.52	14.50	15.08	9.02	10.06	7.12	18.41
Tanqueray 1.75 L	17.09	19.58	20.56	13.10	14.62	9.82	24.53
Canadian 1.75 L	10.23	12.84	13.25	7.36	8.23	5.55	19.99
Seagrams Vo 1.75 L	11.57	15.53	16.42	9.64	11.02	7.60	22.00
Black Velvet Canadian Whiskey 1.75 L	8.52	10.52	10.49	5.04	5.47	3.61	21.50
Capt Morgan Spiced 1.75 L	15.85	16.48	17.12	11.69	12.73	8.79	13.22
Pinnacle Vodka 1.75 L	9.95	11.67	12.03	6.19	7.00	4.65	19.78
Bacardi Dark Rum 1.75 L	12.52	14.22	14.60	8.74	9.57	6.91	19.32

Table C.1: Top Winners and Losers: PH vs Ethanol Tax

This table provides product-level analysis of prices from the scenario that replaces PH with an ethanol tax and holds aggregate ethanol consumption fixed in Table 7. The Biggest Winners see the largest sales increases while the Biggest Losers see the largest sales declines.
p^{old}, p^{m}_{old} denote the prices under PH of wholesalers and manufacturers respectively, and mc^m denotes the estimated marginal costs of the manufacturer.
p^{new} denotes the wholesale prices under the alternative ethanol tax, and $p^{new}_{endog}, p^{m}_{endog}$ denotes the wholesale and manufacturer prices under an ethanol tax where manufacturers endogenously respond.

The second exercise examining variation across products here is to consider the potential for "tagging" or whether the wholesaler market power under PH leads to higher prices on the "right products" or the "wrong products". To understand this, consider our planner's problem and how PH deviates from that problem in (A.7):

$$\frac{p_j(\lambda_e,\lambda_r)-p_j^{PH}}{\mu_j(\theta)} = \underbrace{\Big(1-\mu_j^{PH}/\mu_j(\theta)\Big)(mc_j+UPP_j)+\frac{\mu_j^{PH}}{\mu_j(\theta)}\cdot\Delta UPP_j^{PH}}_{\text{PH Distortion}(\theta)} + \lambda_e^*\cdot D^e_{j\to 0}$$

If the collection of terms labeled *PH Distortion* were perfectly negatively correlated with the external damage term $\lambda_e^* \cdot D^e_{j\to 0}$, then these terms would effectively cancel out and the prices under PH p_j^{PH} would closely line up with the planner's ideal prices $p_j(\lambda_e, \lambda_r)$. We plot the PH distortion for all products against observed wholesale prices in Figure C.1. We notice that it is positive (PH sets prices too high) for $\theta \leq 0.5$ and that it rises with the wholesale price/marginal cost as we predict in Section 6.4 because there is a proportional markup μ_j applied to marginal cost mc_j. For larger values of θ (the planner cares more about revenue) Figure C.1 indicates that PH prices too low compared to the planner. This is in part because the planner corrects for the externality, but also because the UPP_j term for the planner exceeds the $UPP_j(\kappa)$ term from the PH game (the wholesalers do not fully incorporate cross-price effects). The bigger problem here is that as the PH distortion becomes negative it becomes impossible to correct for the external damage term, and we move *further away* from the planner's preferred prices.

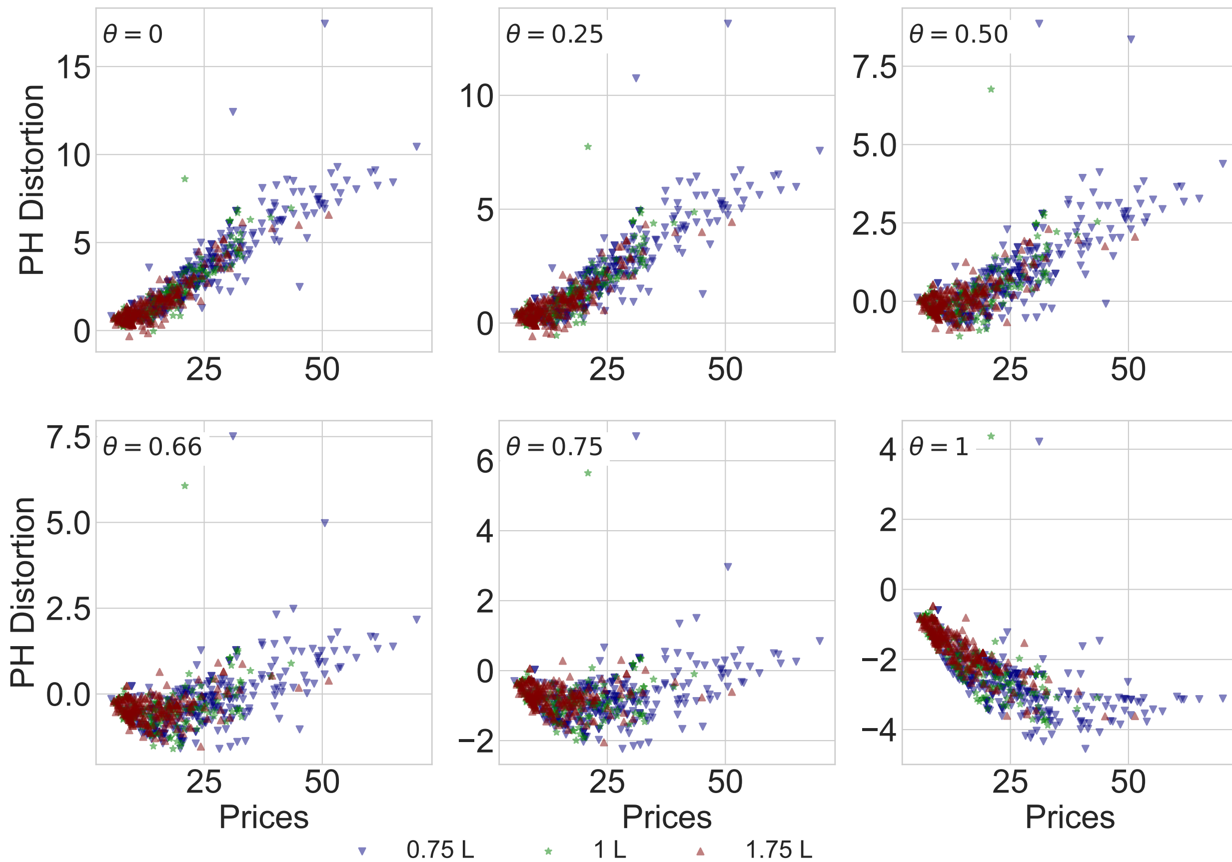
The problem is that we do not know the planner's weights (λ_r, λ_e). What we can instead do is compute the term of PH distortion at different values of θ and check the correlation (or R^2) between PH distortion and diversion away from ethanol $D^e_{j\to 0}$ (which we can compute from our demand estimates). This still has the potential pitfall that even if these terms were correlated, without knowing λ_e^*, we might get the magnitude wrong and be far from the planner's optimal prices. In Table C.2, we do this and find that we never find an $R^2 > .02$ for any value of θ, suggesting that the PH distortion is not helpful in identifying "which products to tax" at all. In contrast, we find (not surprisingly) that $D^e_{j\to 0}$ is highly correlated with the ethanol content e_j of product j, and that a regression of our external damage term on: own ethanol content e_j and diversion to the outside good $D_{j\to 0}$ fits nearly perfectly.[94] That is, the planner might not *exactly* want to tax ethanol e_j, but that ethanol alone has an $R^2 = 0.869$ of the the object the planner would want to levy a tax on.

Example:

As an example the average $Distortion_j(\theta)$ at $\theta = 0$ is \$2.56/L while the average marginal reduction in ethanol consumption for someone not purchasing j is $\frac{1}{J}\sum_j D^e_{j\to 0} = 0.18$ Liters. If the external damage from drinking was $14.22/L$ (of pure ethanol) or about \$5.68/L of Vodka at 40% ABV, then *on average* the higher prices under PH could match the average level of marginal external damage. Of course, we could use (A.12) and set $(1+\tau_r)\cdot \tau_v \approx \frac{\lambda_e^*}{J}\sum_j D^e_{j\to 0}$.

[94]We regressed $D_{j\to 0} \sim \beta_1 e_j + \beta_2 D_{j\to 0} + \varepsilon_j$ and obtained an $R^2 > 0.99$ for products within a single market and across all products and markets.

Figure C.1: Price Differences Between PH Price and Planner Price at Different θ



Note: The chart above shows the share of vodka consumption by national price per liter category. A product's national price category is determined using the average price per liter across all NielsenIQ markets outside of Connecticut-designated market areas. For products only sold in Connecticut or Massachusetts, the state price is used in place of the national price to calculate the price per liter.

	Mean	StDev	R^2
$D^e_{j\to 0}$	0.180	0.050	
$D_{j\to 0}$	0.440	0.050	0.091
$Proof$	79.970	10.480	0.869
$Proof + D_{j\to 0}$			0.993
Elasticity	-4.840	0.360	0.003
UPP	2.270	0.740	0.010
UPP^{PH}	0.540	0.470	0.002
$UPP - UPP^{PH}$	1.720	0.700	0.008
PH Distortion $\theta = 0$	2.560	2.150	0.000
PH Distortion $\theta = 0.25$	1.580	1.630	0.001
PH Distortion $\theta = 0.50$	0.470	1.080	0.003
PH Distortion $\theta = 0.66$	-0.310	0.770	0.009
PH Distortion $\theta = 0.75$	-0.770	0.660	0.015
PH Distortion $\theta = 1$	-2.190	0.910	0.008

Table C.2: Prediction of Diversion Away from Ethanol: $D^e_{j\to 0}$

Note: All regressions predict $D^e_{0\to 0,t}$ for each product and quarter (15,285 observations) as a function of the listed regressors. The interaction $Proof_j \times D_{j\to 0,t}$ includes the base terms $D_{j\to 0,t}$ and $Proof_j$.

C.2. Comparisons over Time

One potential disadvantage of simple tax instruments is that they are often specified in nominal terms. The federal excise tax on spirits has remained fixed at $\$2.85/L$ for some time, while Connecticut's specific tax on spirits increased from \$4.50 per gallon (\$1.18 per liter) to \$5.40 per gallon (\$1.42 per liter) on July 1, 2011, and again to \$5.93 per gallon (\$1.56 per liter) on October 1, 2019, while the sales tax increased from 6% to 6.35% in July 2011, and has remained constant since then.[95] The idea that nominal taxes decline over time in real terms, and the welfare consequences are explored in Blanchette et al. (2020); Seim and Thurk (2023). Rather than repeat those exercises here, we will demonstrate how variation in demand for spirits over time affects our conclusions.

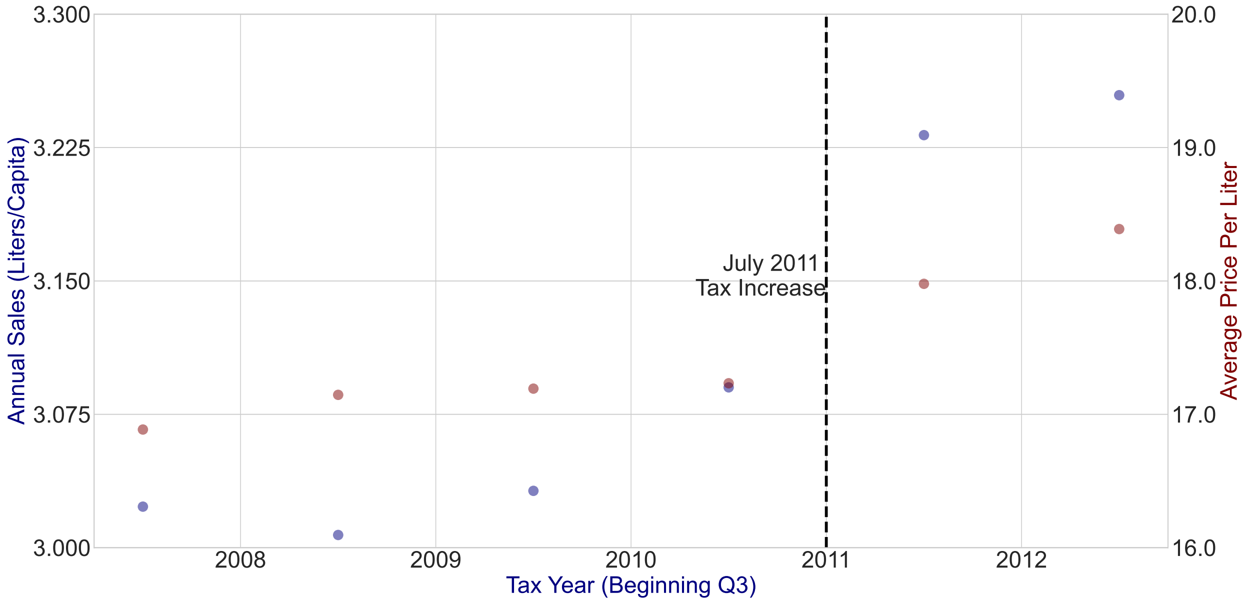
Recall that Figure 1 shows an increase in the consumption of distilled spirits both nationally and in Connecticut during the period covered by our data (July 2007 - June 2013). In Figure C.2, we show how the average price per liter increases over time from \$16.88/L to \$18.39/L (8.9%), while taxes increase only \$0.24/L over the same period. Quantity also increases in both absolute terms (up 12%) and per capita terms (up 8.28%) over the same period, despite the increase in prices and taxes. This poses the question: How do PH and simple tax instruments compare when faced with rising consumer demand over time?

One potential advantage that private firms may have is that they are free to adjust prices each month (including under PH), and that these prices can respond to changes in demand or consumer preferences. Markups depend on the elasticity of demand $\mu^{PH}_{j,t} = \frac{1}{1+\epsilon_{jj,t}(\mathbf{p_t})}$, which is necessarily time-varying. This means that while excise taxes are typically *nominal*, markups depend on both *real and nominal* components (particularly if rising incomes make households less elastic).

There are two different ways to understand how the planner in (7) approaches the external damage from ethanol. One interpretation is that the planner has a literal ethanol budget $\mathbf{e} \cdot \mathbf{q_t} \leq E$ and would like to keep the aggregate ethanol consumption below E in every period. To explore this possibility, in the upper panel of Figure C.3, we calculate the rate of a simple volumetric tax that

[95] To preserve the real power of the $\$1.18/L$ tax from July 2007 through October 2019, the tax would need to be $\$1.46/L$, suggesting that in Connecticut excise taxes have more than kept pace with inflation (albeit in a lumpy manner). Federal taxes of course, have not.

Figure C.2: Average Price and Sales Volume over Time



Average Wholesale Prices Per Liter (in Nominal USD) (CT Price Posting Data).
Our sample runs July 2007 - June 2013. We aggregate our sample "fiscal years" running from July (start of Q3) to June (end of Q2) of the subsequent year and denote as $2007.5, 2008.5, \ldots$
Source: Total Sales Volume in Liters (Author's Calculations/Harmonized Quantity Data).

would keep consumption of ethanol from spirits fixed (in per capita terms): (a) at 2007 levels; (b) at 2013 levels; (c) matching the year-by-year observed consumption. The idea is that a regulator might have found either the overall time path of consumption ideal, or the consumption level in either the first or last year in our data ideal. This represents an upper bound on how much taxes would need to increase over time, because we hold manufacturer prices and federal taxes fixed and replace the wholesale markup with a volumetric tax.

To keep consumption at the 2012-2013 level, the tax would need to increase from \$3.91/L in 2007 to \$5.66/L in 2013, while to keep per capita consumption 8.28% lower at the 2007-2008 level, this would require a higher level of taxes and an increase from \$4.43/L in 2007 to \$6.23/L in 2013. Either scenario would require taxes to increase by 40% during the sample period. Because demand increases significantly (particularly after July 2011), taxes would have to increase at least 27% over our sample to match the PH level of the ethanol consumption period by period. During this period, inflation was only around 12% (2% per year), suggesting that even indexing volumetric taxes to inflation would lead to a faster growth of ethanol consumption than under PH. In the lower panel of Figure C.3, we compare the total consumption of ethanol under PH with the volumetric tax alternative. We set the the tax to \$4.43/L to match the consumption in the first year in our sample, and then allow the tax to rise at an annual rate of 2% or 5%. With a tax increase of 5% per year, ethanol consumption remains at or below pre-existing (PH) levels, while with an increase of 2% per year, consumption would increase by 21% (compared to 29% without indexing).

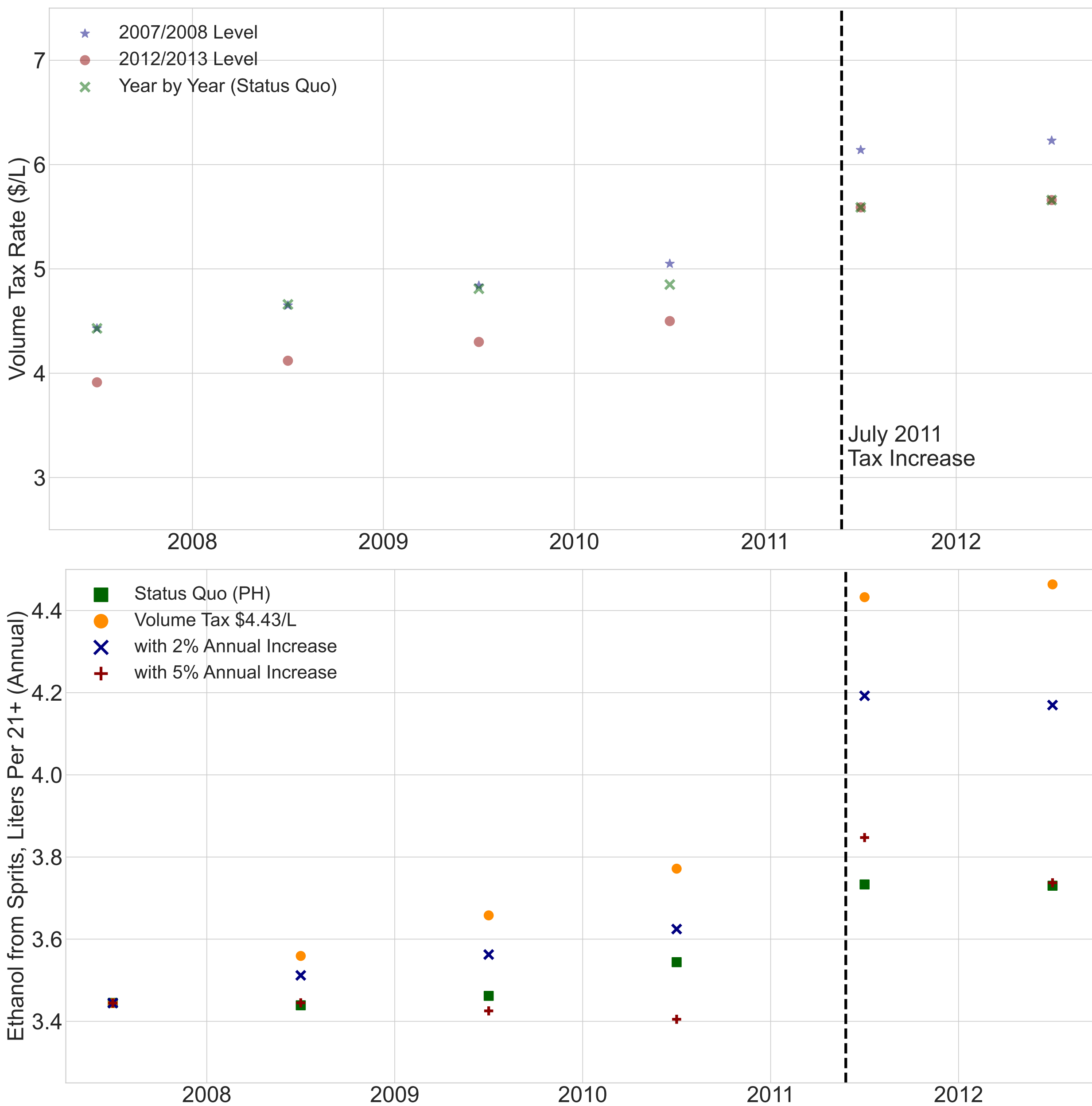
The second interpretation is that the planner does not have a fixed ethanol budget E, but rather the planner puts some Pareto weight λ_e on the external damage term. The shadow cost of externality λ_e could increase in the level of ethanol consumption $\mathbf{e} \cdot \mathbf{q_t}$ (but we do not know how quickly). In this case, what (8) and (A.5) suggest is that the planner's preferred volumetric tax rate τ_v should be higher when products are more substitutable with lower ethanol alternatives (larger average $D^e_{j\to 0}$ as in (A.12)). However, this also means that the planner should respond to a *ceteris paribus* increase in demand by letting consumption (and consumer surplus) rise (at least in part), as long as the (marginal) external damage from alcohol (λ_e) does not increase too quickly.[96] If anything, we would expect that $D^e_{j\to 0}$ will decline when demand increases (as the outside option becomes less attractive, we also expect fewer people to substitute away from spirits).[97] Indeed, this is exactly what we observe in Figure C.4, the sales-weighted average of $D^e_{j\to 0}$ declines over time by around 8% over our sample, suggesting that spirits are becoming less substitutable with the outside good (and more with one another) over time. This does not necessarily imply that the planner would want to *reduce* taxes on ethanol as λ_e may still be rising.

Much like in the rest of our article, we wish to avoid taking a stand of the magnitude of the externality and the planner's weights (λ_e, λ_r), instead we repeat the exercise of Figure 9 using the demand estimates from our very first quarter (2007 Q3), instead of our very last quarter (2013 Q2). We report the results in Figure C.6. Qualitatively the results are highly similar to those in Figure 9: volumetric taxes or ethanol taxes are able to deliver higher consumer surplus (around 12%) at similar levels of ethanol consumption to PH. Also, by replacing the highly profitable wholesale tier with competitive distribution and a tax significantly increases tax revenue (by around 270%). The only important difference is that the total revenue generated under the PH system (tax revenue plus wholesaler profit) is now slightly higher than the competitive alternative plus volumetric tax

[96]This might be because λ_e increases with $\mathbf{e} \cdot \mathbf{q_t}$, or it might be that the externality function itself $f(\mathbf{e} \cdot \mathbf{q_t})$ is *convex*. Griffith et al. (2019) calibrate a convex external damage function at the *individual* level.

[97]This is a straightforward application of Conlon and Mortimer (2021) to (A.5).

Figure C.3: Volumetric Taxes Over Time



Sample runs July 2007 - June 2013. We aggregate our sample "fiscal years" running from July (start of Q3) to June (end of Q2) of the subsequent year and denote as $2007.5, 2008.5, \ldots$

(Upper) Tax Rates are in \$/L. Taxes hold ethanol consumption fixed at either (2007Q3-2008Q2, 2012Q3-2013Q2 levels) or that hold ethanol at the existing PH level year by year.

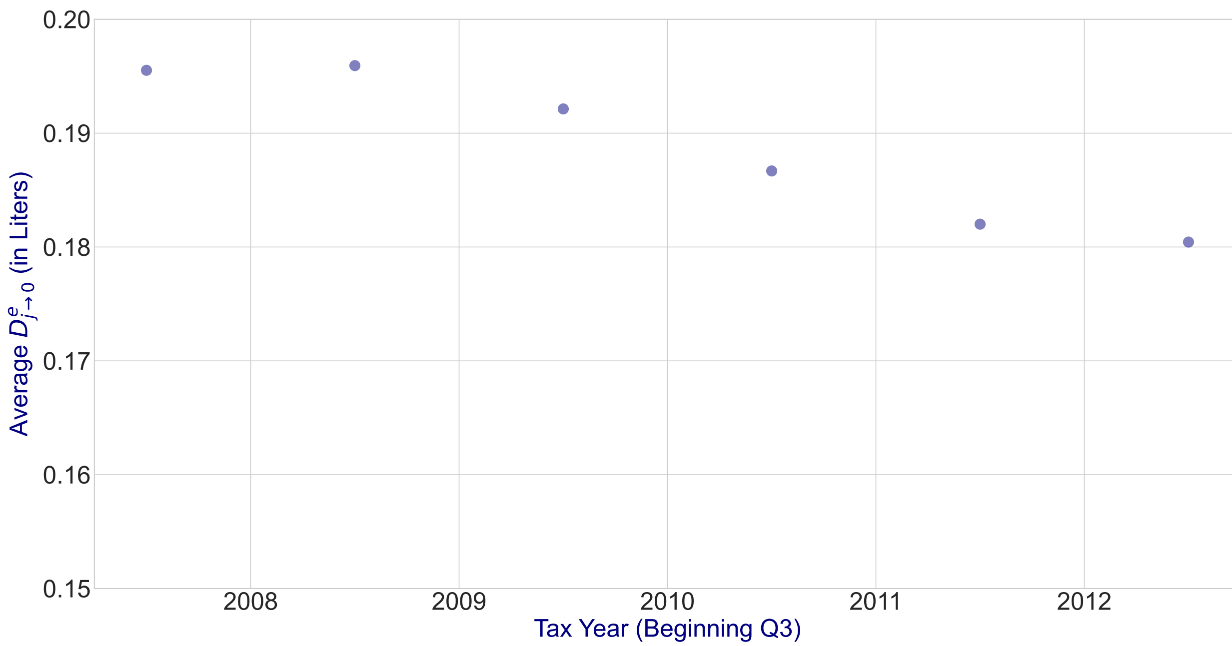
(Lower) Ethanol Consumption is Liters per capita (Adults 21+) at the tax rate of $\tau_{v,t} = \$4.43 \cdot (1+r)^t$

Vertical line denotes the \$0.24/L tax increase observed in July 2011.

Source: Authors' Calculations.

rather than slightly lower. In theory, at the same level of ethanol consumption (but lower consumer surplus), a perfectly designed lump-sum tax could extract more revenue under the PH system than a volumetric tax would raise.

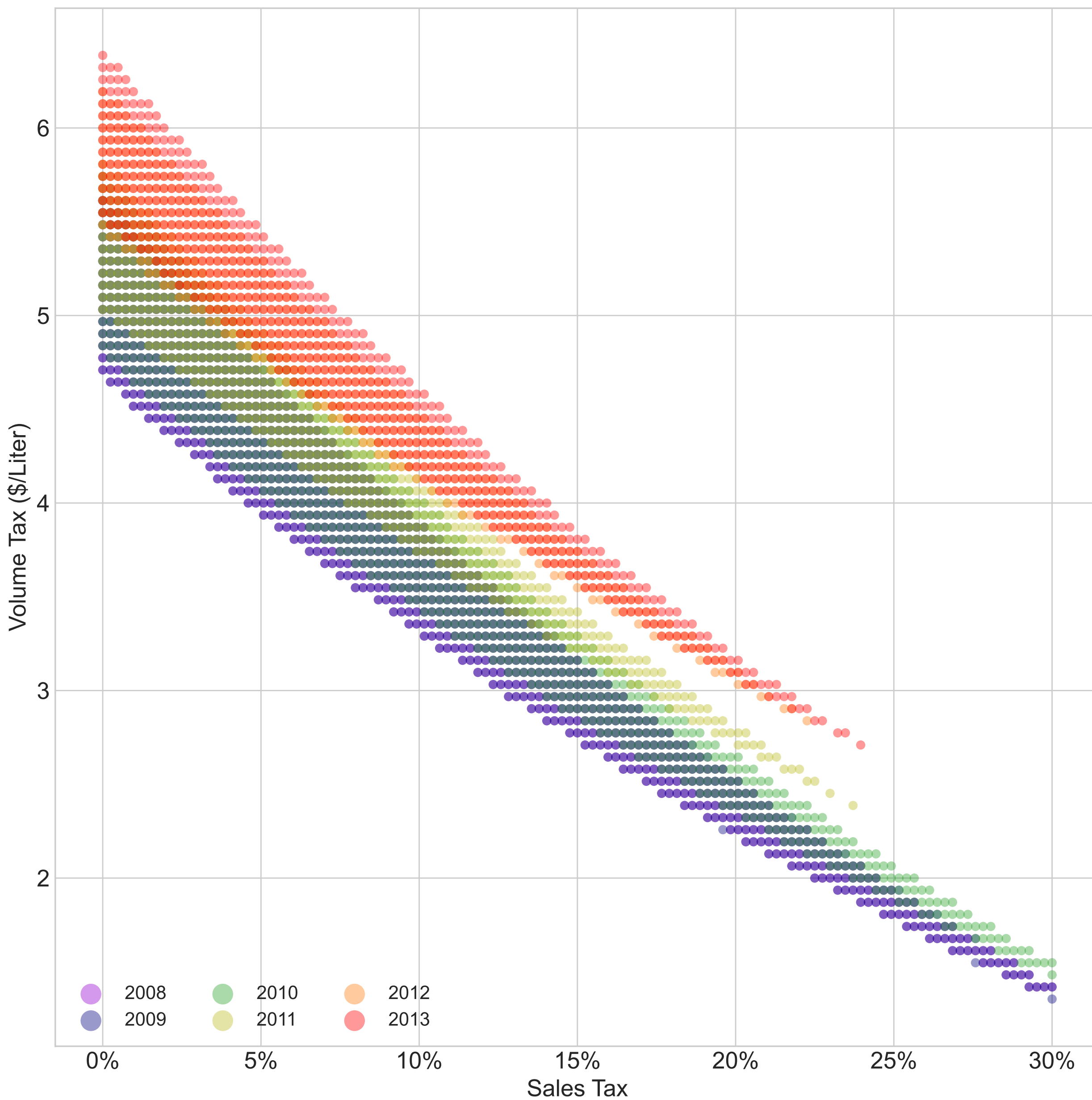
Figure C.4: Diversion Away from Ethanol



Sales weighted $D^e_{j\to 0}$ by year. Assumes outside good has zero ethanol $e_0 = 0$
Sample runs July 2007 - June 2013. We aggregate our sample "fiscal years" running from July (start of Q3) to June (end of Q2) of the subsequent year and denote as $2007.5, 2008.5, \ldots$
Source: Authors' Calculations.

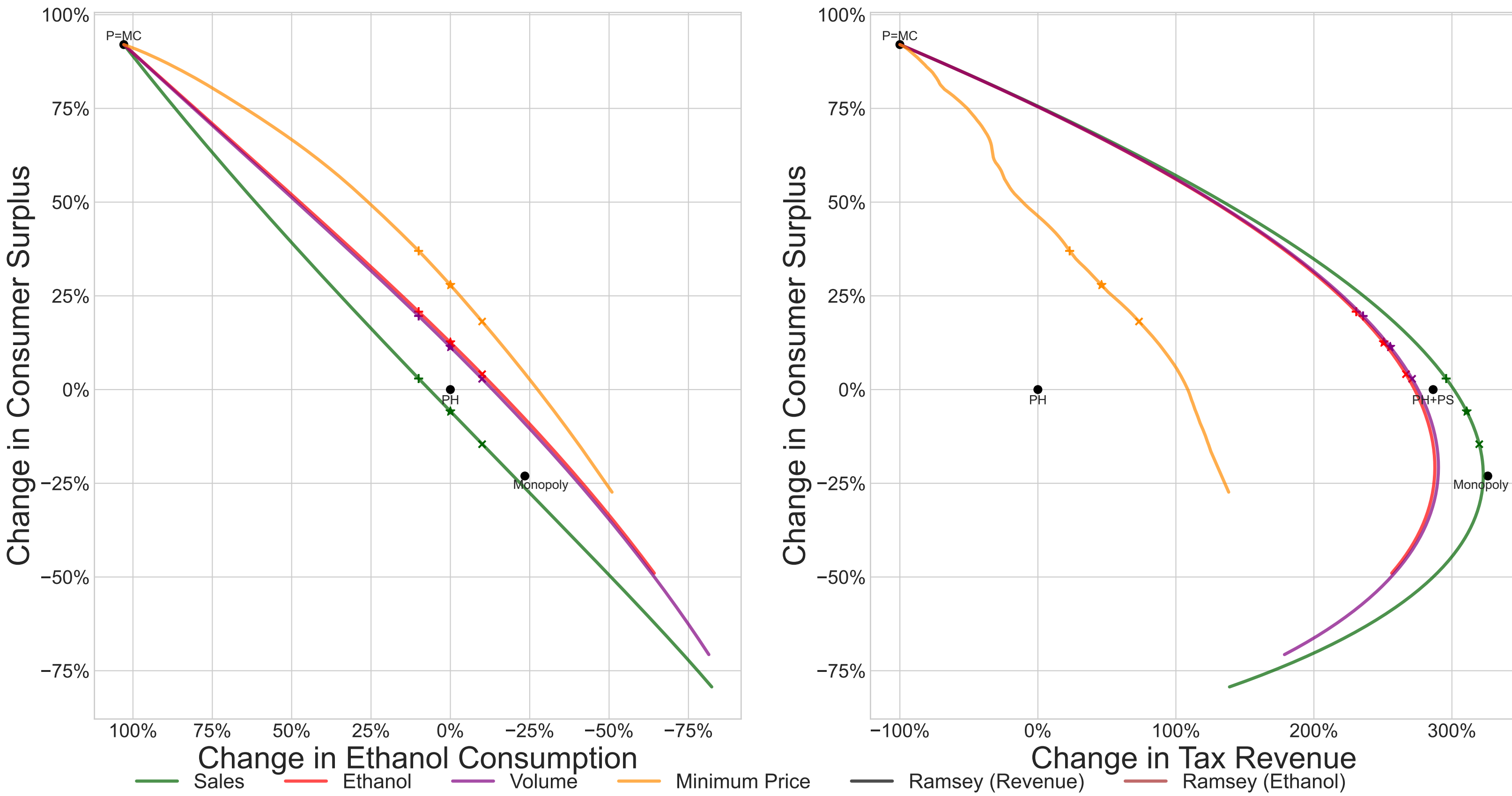
This motivates the second exercise: is there a combination of simple taxes (volumetric and sales taxes) that dominate the PH system on each of the planner's objectives: higher consumer surplus, lower ethanol consumption, and more tax revenue generated? Our goal in Figures 10 and C.5 is to characterize the set of tax rates that dominate PH for each period in our data and observe how the set of rates that outperform PH change over time. This range is somewhat wider (particularly for low levels of sales taxes) in the case of Figure C.5 where we simply require that the tax alternatives raise more revenue than the status quo (as opposed to more revenue than a lump-sum tax on wholesaler profits would raise in Figure 10). Allowing for endogenous manufacturer responses only slightly shrinks the region of taxes that dominate PH, but it tends to shift the region towards lower overall rates (when manufacturers raise prices they capture some of the additional revenue, but also make it easier to hit the ethanol target).

Figure C.5: Combinations of Simple Taxes that Dominate PH



Note: Each dot signifies a point in the space of sales and volumetric tax rates (τ_s, τ_v) that dominates the PH outcome in all three categories: greater consumer surplus, lower ethanol consumption, greater tax revenue.
Our criteria for tax revenue is that it **exceeds the amount that is raised by existing volumetric taxes** under PH. (point PH from Figure 9).
We calculate each point using the Q2 (April, May, June) estimates for that calendar year.
Source: Authors Calcuations.

Figure C.6: Consumer Surplus vs. Tax Revenue and Ethanol Consumption Under Alternative Policies For 2007 Q3



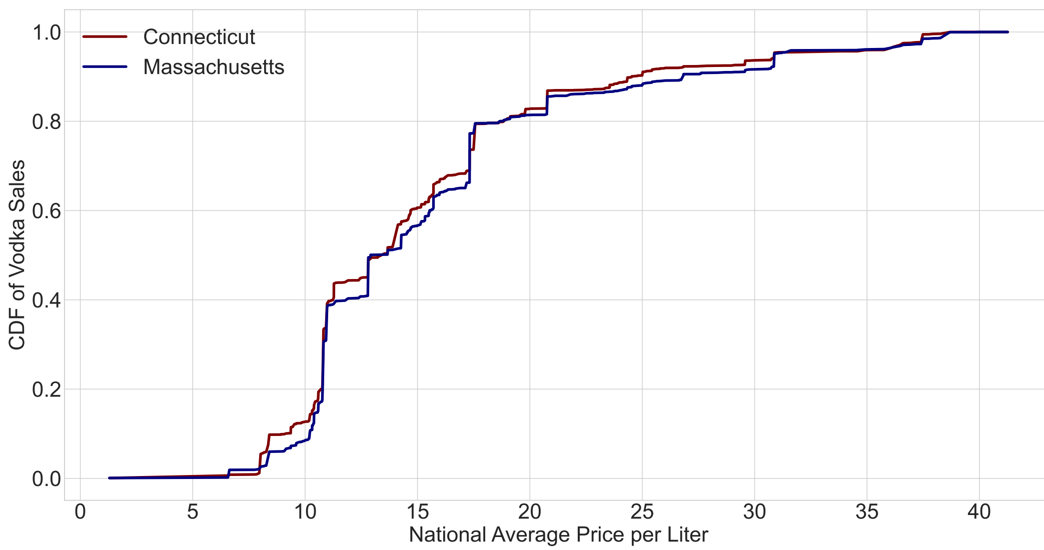
Note: The figure above plots the change in tax revenue (left panel) and ethanol consumption (right panel) against the change in consumer surplus for each of the policy alternatives to PH detailed in Table 5 that we consider. The frontiers trace the trade-off between consumer surplus and tax revenue or ethanol consumption for each policy instrument. Stars indicate an aggregate ethanol consumption level equal to total ethanol under PH, while (×) denotes 10% less and (+) denotes 10% more ethanol consumption (in the left panel higher ethanol consumption corresponds to less tax revenue). We also mark competitive prices without taxes (denoted by P = MC), and PH pricing. In the right panel, we indicate the revenue generated by existing excise taxes under PH pricing as well as the sum of tax revenue and wholesale profits generated by PH.

D. Additional Specifications and Robustness Tests [Online Only]

D.1. Alternative to Figure 3

One concern about Figure 4 is that our choice of price bins may seem arbitrary. A better but more complicated way to address this concern is to rank all vodka products by their national price per liter and compare the CDF of purchases for Connecticut and Massachusetts. If cumulative sales are larger at each national price, then we can say that Connecticut consumes an inferior bundle of vodkas. (We could repeat the exercise for all products, but that might conflate preferences for different categories: Vodka vs. Tequila or Scotch Whisky for "quality"). We plot this in Figure D.1 and show that the bundle in CT nearly FOSD the bundle in MA (except for a few ties).

Figure D.1: CDF of Vodka Consumption by National Average Price Per Liter



Note: The chart above shows the share of vodka consumption by national price per liter category. A product's national price category is determined using the average price per liter across all NielsenIQ markets outside of Connecticut-designated market areas. For products only sold in Connecticut or Massachusetts, the state price is used in place of the national price to calculate the price per liter.

D.2. Correlation Between Markup Measures

Tables 1 and 2 report summary statistics including measures of profit margins $\eta_j = p_j - mc_j$ and the Lerner index $L_j = \eta_j / p_j$ at the product level (aggregated across all 24 quarters of our data). In Table D.1 we report the correlation in markups across all products and periods. Some key features: manufacturer and wholesaler Lerner indices exhibit a strong positive correlation $\rho = 0.72$. This is consistent with wholesalers and manufacturers possessing market power and pricing in line with the elasticity of demand. Retailer Lerner markups are (weakly) negatively correlated with the manufacturer and wholesaler markups. In our prior work (Conlon and Rao, 2020), we find that retailer pricing tends not to track the elasticity of demand but rather that retailers tend to round to the next highest \$0.99 price ending and add \$1, \$2, etc.. The additive markup measures η_j are positively correlated with one another as well as with the manufacturer price p_m. This is consistent with larger markups (in dollar terms) on more expensive products as we document in Figure 5.

	Lerner (M)	Lerner (W)	Lerner (R)	η_M	η_W	η_R	p_M
Lerner (M)	1.00	0.72	-0.14	0.10	0.36	-0.08	-0.11
Lerner (W)	0.72	1.00	-0.24	0.22	0.60	-0.02	0.07
Lerner (R)	-0.14	-0.24	1.00	-0.08	-0.16	0.62	-0.05
η_M	0.10	0.22	-0.08	1.00	0.87	0.59	0.97
η_W	0.36	0.60	-0.16	0.87	1.00	0.45	0.78
η_R	-0.08	-0.02	0.62	0.59	0.45	1.00	0.61
p_M	-0.11	0.07	-0.05	0.97	0.78	0.61	1.00

Table D.1: Correlation Between Markup Measures

Additive Markup: $\eta_j = p_j - mc_j$; Lerner Markup $L_j = \eta_j / p_j$. (M)anufacturer; (W)holesaler; (R)etailer.

D.3. Sensitivity of Demand Estimates

D.3.1. Varying the Nesting Parameter

We explore the sensitivity of our parameter estimates by fixing the nesting parameter ρ at different increments between $\rho = 0$ (plain logit) and $\rho = 1$ (all substitution within the nest) and re-estimating the remaining parameters of the model. We include the demand moments, the supply moments, the micro-moments, and the second-choice category moments.

We compute a one-step GMM estimator using the same instruments and the same (2SLS) weighting matrix for each value of ρ. As indicated in Table D.2, $\widehat{\rho} = 0.242$ minimizes the GMM objective. We also report our second-stage GMM estimates (which use the approximation to the optimal instruments, and an updated weighting matrix $\widehat{W}(\widehat{\theta})$ which gives the $\widehat{\rho} = 0.269$ that we report in Table 3.

As we increase ρ in Table D.2 we see that more individuals stay within the same product category ($Vodka \rightarrow Vodka$) and fewer divert to the outside good. At our estimate $\widehat{\rho} = 0.269$ this corresponds to 46% of consumers switching to the outside good, and 69% of consumers switching from one Vodka to another if the first-choice product was unavailable (target 50.6%). For the other categories $\widehat{\rho} = 0.269$ tends to over-estimate the "same category" switching behavior (Rum 56.3% vs 20.2%, NA Whiskey 55.6% vs. 26.1%, UK Whiskey 52.5% vs. 33.2%). The exception is Gin for which the model predicts 50.0% while the target from the NielsenIQ panelist data is 60.4%. Absent these "second-choice moments" we would estimate a value of $\rho \in (0.45, 0.5)$ which would imply that 70% of consumers would substitute within the category, so that including these moments pushes us towards smaller values of ρ.

It is important to note that because we impose the supply moments, we are effectively constraining the markups to match (on average), so that each row in Table D.2 has a nearly identical average Lerner markup $(p^w_{jt} - p^m_{jt} - \tau_{jt})/p^w_{jt} = 0.238$. As we increase ρ (and fix the markup) the own- (and cross-) elasticities increase so that consumers become more elastic, but there is less substitution to the outside good, more to other products in the same category, and the overall elasticity of alcohol with respect to a 1% tax declines. This aggregate elasticity captures how quickly consumers substitute away from spirits as we raise the price, and ends up being a good barometer of the welfare impacts of the tax alternatives. It is important to note that as we adjust ρ, other parameters (particularly σ_0 and π_0, which govern overall taste for spirits) also adjust so that ρ is not the only parameter that determines the own- and cross-elasticity. The own pass-through rate is relatively unaffected by changes in ρ but is overshifted ≈ 1.3 and consistent with reduced-form estimates in Conlon and Rao (2020).

The point of Table D.3 is to show how the second-choice moments, in particular, help to identify a key parameter ρ, which governs our welfare predictions. One important caveat discussed in Appendix B.1 is that we have "pseudo second-choice" data constructed by looking within a household's purchase history over time. This is not true second-choice data because we do not know a household that previously purchased *Smirnoff Vodka* which instead purchased *Skyy Vodka* did so because Smirnoff was unavailable, it may have been because the price went up, or because they have some love of variety. Berry and Haile (2024); Conlon and Gortmaker (2025) for a more technical discussion of micro-data and second-choices, and Conlon and Mortimer (2021) for how second-choice diversion measures related to small quality changes and price changes.

Table D.2: Sensitivity to different values of nesting parameter ρ

ρ	Own	Agg	Lerner	Outside Good	Vodka	Gin	Rum	NA	UK	Tequila	Objective
0.05	-4.626	-0.622	0.239	0.577	0.476	0.162	0.272	0.250	0.212	0.169	5863.563
0.10	-4.662	-0.601	0.239	0.552	0.533	0.250	0.348	0.331	0.296	0.257	5813.226
0.15	-4.699	-0.579	0.239	0.526	0.585	0.331	0.418	0.405	0.371	0.336	5778.403
0.20	-4.738	-0.556	0.239	0.500	0.632	0.405	0.482	0.472	0.440	0.409	5758.438
0.242	-4.774	-0.535	0.238	0.477	0.668	0.463	0.532	0.524	0.494	0.466	5753.096
0.25	-4.781	-0.532	0.238	0.473	0.674	0.473	0.541	0.532	0.503	0.475	5753.305
0.269	-4.791	-0.524	0.238	0.465	0.690	0.499	0.563	0.556	0.525	0.501	5361.402
0.30	-4.826	-0.507	0.238	0.445	0.713	0.535	0.594	0.588	0.560	0.536	5763.204
0.35	-4.875	-0.480	0.238	0.417	0.748	0.591	0.644	0.638	0.612	0.592	5789.146
0.40	-4.927	-0.453	0.237	0.388	0.780	0.642	0.688	0.683	0.659	0.643	5832.463
0.45	-4.983	-0.424	0.237	0.359	0.809	0.689	0.729	0.725	0.703	0.689	5894.985
0.50	-5.043	-0.395	0.237	0.328	0.835	0.731	0.766	0.762	0.742	0.731	5978.927
0.55	-5.107	-0.363	0.236	0.298	0.859	0.770	0.800	0.796	0.778	0.770	6086.915
0.60	-5.175	-0.330	0.236	0.267	0.880	0.805	0.831	0.827	0.811	0.805	6222.142
0.65	-5.250	-0.296	0.235	0.235	0.900	0.837	0.859	0.856	0.841	0.837	6388.561
0.70	-5.332	-0.260	0.235	0.202	0.918	0.867	0.885	0.882	0.869	0.866	6591.177
0.75	-5.421	-0.222	0.234	0.170	0.935	0.894	0.908	0.906	0.895	0.893	6836.444
0.80	-5.519	-0.183	0.234	0.137	0.950	0.918	0.930	0.928	0.920	0.918	7132.864
0.85	-5.630	-0.141	0.233	0.103	0.964	0.941	0.950	0.948	0.942	0.942	7492.025
0.90	-9.526	-0.077	0.233	0.018	0.946	0.912	0.923	0.922	0.912	0.911	8602.323
0.95	-15.003	-0.076	0.233	0.011	0.974	0.957	0.962	0.962	0.957	0.956	10772.279

Note: We profile demand estimates by varying the level of ρ. This uses the (aggregate) demand moments, the (aggregate) supply moments, and micro-moments from NielsenIQ Panelist data.
Markups, own elasticity, and outside good diversion are unweighted averages over (j,t). Aggregate elasticity is the market-level reduction in purchase volume for a 1% sales tax averaged over markets. Pass-through is own $\frac{\partial p_j}{\partial c_j}$ (dollar for dollar) averaged over products in the final market.
Caution is required comparing GMM objectives across specifications since they have different weighting matrices.
Source: Authors' calculations

D.3.2. Allowing for Wholesaling Costs

We might worry that the main results are driven by our assumption that in the absence of post-and-hold policies, the wholesaler tier becomes perfectly competitive. A reasonable concern is that wholesaling is not costless, and unless wholesalers charge a markup above manufacturer prices, they may not be able to cover the costs of hiring drivers, and operating warehouses. To alleviate these concerns, we set $\mathbf{mc^w} = \mathbf{p^m} + 1$, so that the wholesaler incurs an additional cost of \$1 per liter both when estimating the demand model, and when computing the counterfactual. We think this is reasonable, as it is in line with the wholesaler margins on the lowest margin items.[98] The exercise is slightly different from Table 7 where we hold the parameter estimates fixed, and allow for a $1/L wholesale margin.

Qualitatively, the patterns in Figure 9 in the main text and Figure D.3, which allows for the $1 per liter wholesaling cost, are nearly identical. The relative ranking of various tax instruments, and most importantly, the fact that post and hold is clearly dominated by alternative taxes on a competitive market, remains the same. Quantitatively, the somewhat higher cost means that the overall level of additional tax revenue that can be generated is reduced slightly, such that we can never increase revenue by more than 250%. The resulting equilibrium prices are highly similar, the main difference being that rather than capturing all of that as additional tax revenue, some must be used to cover the wholesaler costs.

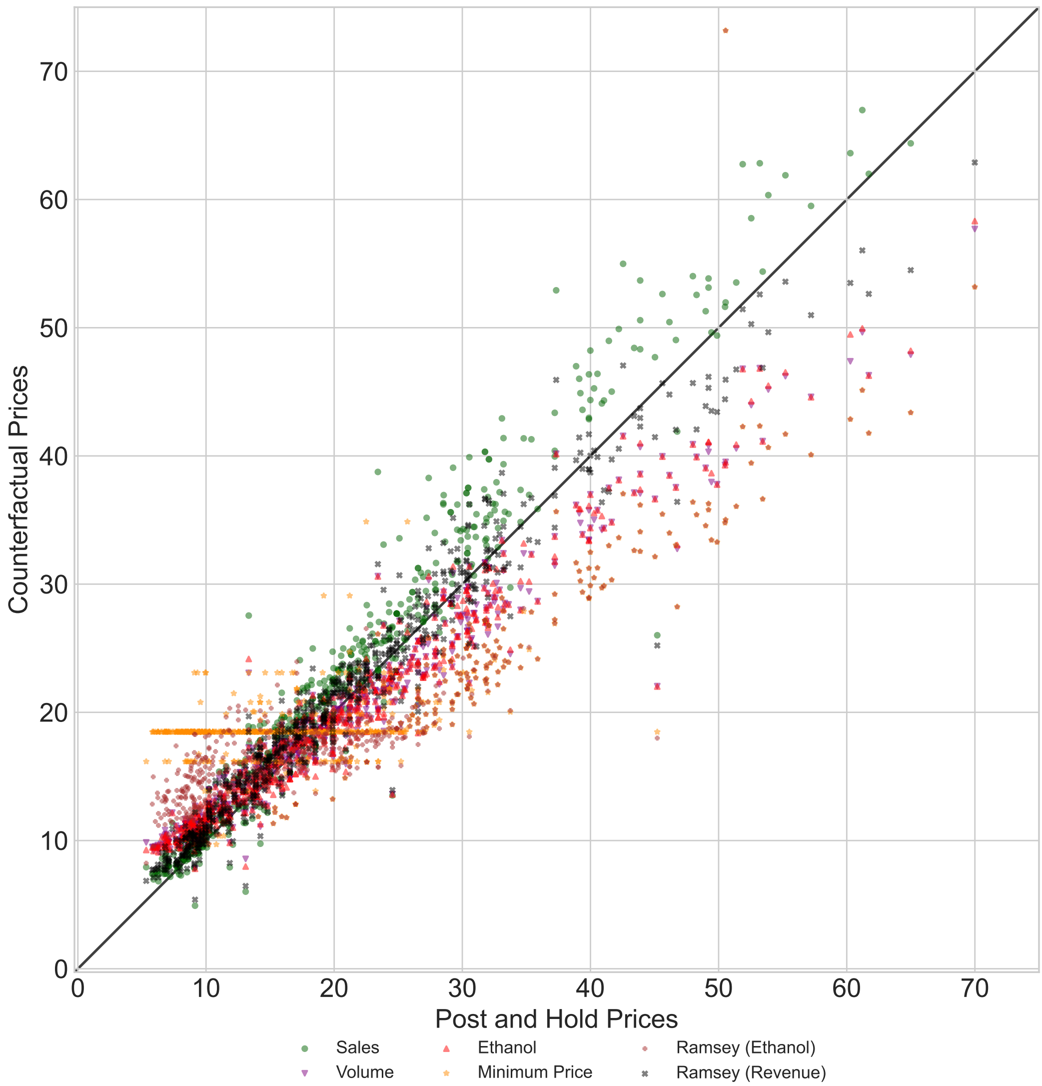
Table D.3: Distributional Impacts of Counterfactual Policies with $\mathbf{wc} = 1$

		% Change in CS					
No Change in Ethanol	% Total Revenue	% Overall	Below $25k	$25k-$45k	$45k-$70k	$70k-$100k	Above $100k
Ramsey (Ethanol)	7.1	27.8	5.1	4.3	0.2	15.6	40.7
Minimum Price	10.4	27.8	5.0	5.1	2.5	16.5	40.4
Ethanol	211.2	10.9	0.8	0.3	-0.2	4.8	16.6
Volume	214.3	10.1	-0.7	-2.0	-2.0	2.9	16.3
Sales+Volume	260.0	-0.9	-1.4	-2.1	-2.9	-2.7	-0.2
Ramsey (Revenue)	266.6	-3.7	-2.3	-2.6	-3.4	-4.8	-4.1
Sales	271.3	-8.8	-2.1	-1.7	-3.4	-6.6	-12.2
-10% Ethanol							
Ramsey (Ethanol)	33.6	17.9	-6.4	-13.0	-15.8	-1.5	33.2
Minimum Price	37.5	17.8	-6.5	-12.2	-13.9	-0.7	32.8
Ethanol	228.3	2.3	-9.3	-14.9	-14.4	-9.1	10.1
Volume	231.3	1.4	-11.0	-17.2	-16.3	-11.1	9.7
Sales+Volume	275.6	-10.4	-11.0	-16.4	-16.3	-16.0	-8.4
Ramsey (Revenue)	280.0	-13.0	-12.2	-17.2	-17.0	-18.0	-12.0
Sales	280.4	-17.7	-11.3	-15.5	-16.3	-18.9	-19.7
+10% Ethanol							
Minimum Price	-11.1	37.1	16.0	24.0	20.5	34.7	46.4
Ramsey (Ethanol)	-6.9	36.9	16.1	23.0	18.1	33.5	46.7
Ethanol	190.9	19.3	10.5	16.2	14.6	19.2	22.6
Volume	193.9	18.5	9.1	14.0	12.8	17.3	22.3
Sales+Volume	238.5	8.7	8.1	13.2	11.2	11.4	7.9
Ramsey (Revenue)	254.9	2.3	4.4	8.0	6.4	4.6	0.5
Sales	256.4	0.1	7.1	13.0	10.1	6.5	-4.7

Note: The table above reports estimates of the impacts of the counterfactual policy alternatives described in Table 5 on tax revenue collected, overall consumer surplus, and the distribution of consumer surplus across the five income bins. All effects are reported as percentage changes relative to the PH baseline. The top panel describes the impact of alternative policies that limit ethanol consumption to the same aggregate level as under PH while panels B and C report the effects of alternative policies that reduce and increase ethanol consumption by 10%, respectively. Revenue is calculated as the additional tax revenue raised by the state compared to the existing excise tax collections.
Source: Authors' calculations

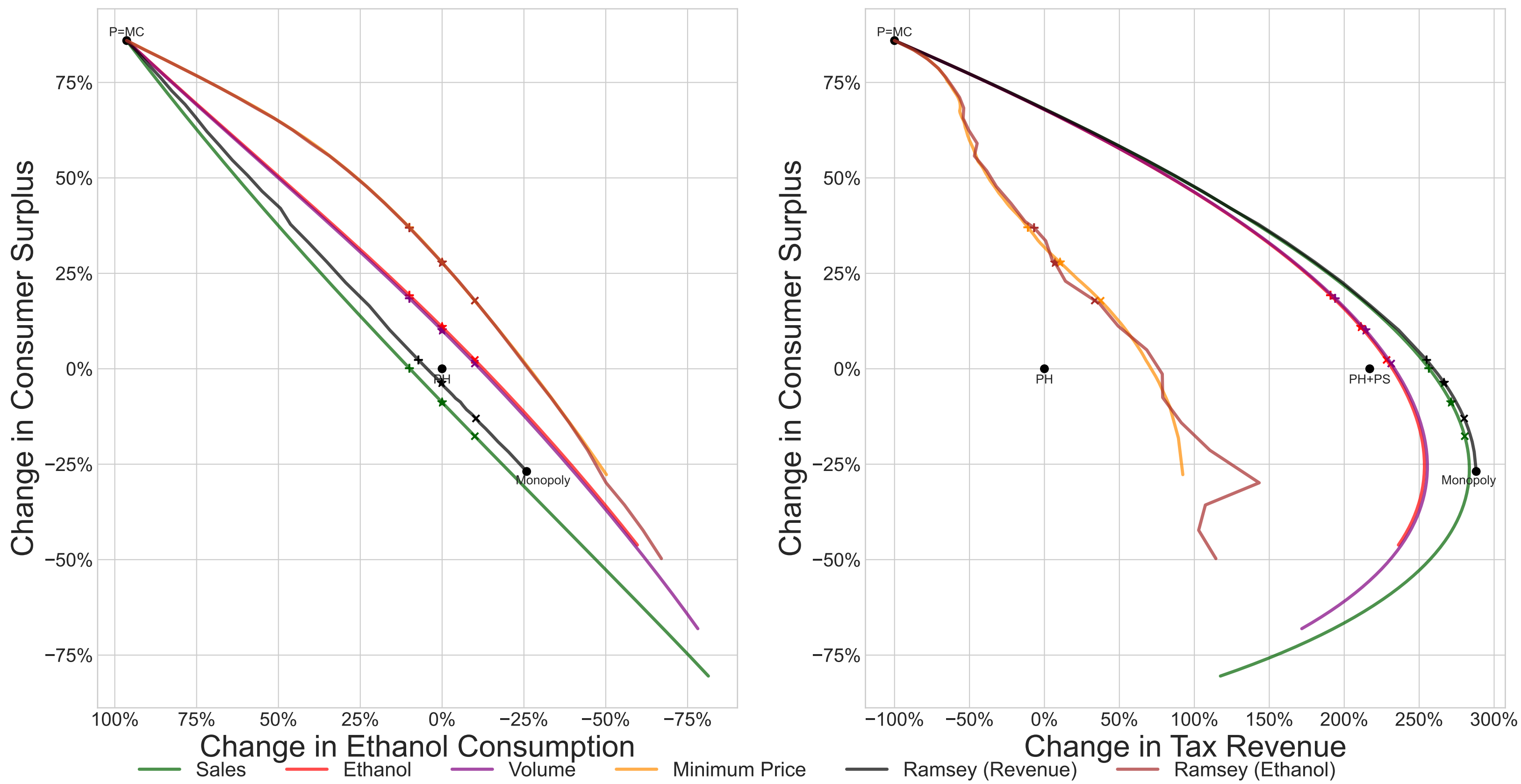
[98] We obtain similar results if we consider larger wholesaling costs of $\mathbf{mc^w} = \mathbf{p^m} + 2$ or $\mathbf{mc^w} = \mathbf{p^m} + 3$.

Figure D.2: Prices Under PH vs. Other Policy Alternatives with $\mathbf{wc} = 1$



Note: The figure above plots product prices under PH against prices under our counterfactual policy alternatives. In each of our counterfactual scenarios, we consider a tax rate that would keep the overall level of ethanol fixed at the status quo. Our taxes follow the definitions in Table 5, and are levied on a competitive market with a \$1/L additional wholesaling cost. The solid black 45-degree line illustrates prices unchanged from PH.

Figure D.3: Consumer Surplus vs. Tax Revenue and Ethanol Consumption Under Alternative Policies with $\mathbf{wc} = 1$



Note: The figure above plots the change in tax revenue (left panel) and ethanol consumption (right panel) against the change in consumer surplus for each of the policy alternatives to PH detailed in Table 5 that we consider. The frontiers trace the trade-off between consumer surplus and tax revenue or ethanol consumption for each policy instrument. Stars indicate an aggregate ethanol consumption level equal to total ethanol under PH, while ($\times$) denotes 10% less and (+) denotes 10% more ethanol consumption (in the left panel higher ethanol consumption corresponds to less tax revenue). We also mark competitive prices without taxes (denoted by P = MC), and PH pricing. In the left panel, we indicate the revenue generated by existing excise taxes under PH pricing as well as the sum of tax revenue and wholesale profits generated by PH.

E. Additional Cross State Evidence [Online Only]

E.1. Cross-state Evidence on Consumption Effects of States Ending PH

Theory suggests that PH leads to higher markups, which is supported by the price comparisons detailed in Section 4.1. As such it is natural to expect that these higher prices may reduce aggregate alcohol consumption at the state level, which may be a policy objective.

To assess the impact of PH laws on aggregate alcohol consumption, we assemble a panel of annual state data measuring wine, beer, and spirits consumption, as well as demographic characteristics. These data are drawn from the National Institute on Alcohol Abuse and Alcoholism (NIAAA) *U.S. Apparent Consumption of Alcoholic Beverages*, which tracks annual consumption of alcoholic beverages for each state. We use the timing of when different states terminated PH laws (often as the result of lawsuits) to measure the association between regulation and alcohol consumption. Table E.1 reports PH termination dates. This table matches Cooper and Wright (2012), who also run a similar panel regression to the one we describe below (and obtain similar results):[99]

Table E.1: States with Post and Hold Laws

	Wine	Beer	Spirits
Connecticut	Y	Y	Y
Delaware	End 1999	End 1999	End 1999
Georgia	N	Y	Y
Idaho	Y	Y	N
Maine	Y	Y	N
Maryland	End 2004	End 2004	End 2004
Massachusetts	End 1998	End 1998	End 1998
Michigan	Y	Y	Y
Missouri	Y	N	Y
Nebraska	End 1984	N	End 1984
New Jersey	Y	Y	Y
New York	Y	Y	Y
Oklahoma	End 1990	End 1990	Y
Pennsylvania	N	End 1990	N
South Dakota	Y	N	Y
Tennessee	N	Y	N
Washington	End 2008	End 2008	N
West Virginia	N	N	Y

Note: The table above lists all states that have or have repealed PH regulations and details the types of alcoholic beverages covered by PH rules. Y denotes a state and beverage category with PH provisions. N denotes a state and beverage category was never subject to PH laws. The year of repeal is denoted for states that ended their PH regulations. No state adopted PH after the start of sample period, 1983. This table is a reproduction of Table 1 of Cooper and Wright (2012).

These state panel regressions are similar to those of Cooper and Wright (2012) and have the form:

$$Y_{it} = \alpha + \beta PH_{it} + X_{it}\gamma + \delta_t + \eta_i + \epsilon_{it} \tag{E.1}$$

[99] In contrast, Saffer and Gehrsitz (2016) find a null effect of PH on prices, but rely on ACCRA data which tracks the price of only one brand each for: beer (Budweiser 6-pack), wine (Gallo Sauvignon Blanc) and distilled spirits (J&B Scotch).

The dependent variable is the log of apparent consumption per capita, where consumption is in ethanol-equivalent gallons and the relevant population is state residents age 14 and older. PH_{it} is a dummy variable equal to one if state i has a PH law in place at time t; X_{it} is a vector of control variables; and δ_t and η_i are time and state fixed effects, respectively. The coefficient of interest, β, describes the reduction in alcohol consumption associated with PH laws.

We report the results in Table E.2. The specification of column 1 includes only time and state fixed effects while column 2 adds state-specific linear time trends. Accounting for state differences in underlying consumption trends attenuates the wine coefficient, rendering it statistically insignificant, but increases the magnitude and precision of beer and spirits coefficients and makes them statistically significant.

The identifying variation comes from the handful of states ending their PH requirement. There are a number of reasons we should remain cautious about taking the regression estimates too seriously. The first is that we don't know why states terminate PH, though in several cases it was the result of losing a lawsuit rather than through the legislative process. The bigger issue is that when states eliminate PH, they tend to also change tax rates, and liberalize other laws regarding the distribution and sale of alcoholic beverages. We may wrongly attribute other factors (ending prohibitions on Sunday sales, etc.) to eliminating PH.

Table E.2: Post and Hold Laws and State Alcohol Consumption

	(All)	(All)	(All)	(PH only)	(PH NE)
Wine					
PH	-0.0545***	-0.0215	-0.0197	-0.0277	-0.00360
	(0.0183)	(0.0192)	(0.0192)	(0.0182)	(0.0356)
R^2	0.965	0.984	0.984	0.985	0.988
Beer					
PH	-0.0155	-0.0218**	-0.0207**	-0.0192**	-0.0297**
	(0.0113)	(0.00968)	(0.00959)	(0.00859)	(0.0134)
R^2	0.891	0.968	0.968	0.954	0.980
Spirits					
PH	-0.00702	-0.0731***	-0.0725***	-0.0665***	-0.0851***
	(0.0175)	(0.0183)	(0.0181)	(0.0175)	(0.0279)
R^2	0.950	0.982	0.982	0.976	0.984
Year FE	Y	Y	Y	Y	Y
State FE	Y	Y	Y	Y	Y
State Time Trends	N	Y	Y	Y	Y
Demog. Controls	N	N	Y	Y	Y
PH States	N	N	N	Y	Y
NE States	N	N	N	N	Y
Observations	1,428	1,428	1,428	532	168

Note: The table above presents coefficients from regression equation E.1. The outcome of interest is the log of apparent consumption per capita, where consumption is in ethanol equivalent gallons and the relevant population is state residents age 14 and older. Column 1 only includes state and time fixed effects. Column 2 adds state-specific time trends while column 3 also includes state demographic controls. Column 4 limits the sample to states that have had PH laws. Column 5 restricts the sample further to only northeastern states that once had PH laws. The alcohol consumption data are from the National Institute on Alcohol Abuse and Alcoholism, which is part of the National Institutes of Health; the demographic information comes from the Census Bureau's intercensal estimates. Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

E.2. Cross-state Evidence on Employment and Establishment Effects of States Ending PH

Advocates for PH argue that the regulation benefits small retailers by ensuring that they pay the same wholesale prices as large retailers such as Costco or Total Wine and More.[100] If PH does indeed protect small retailers, PH states like Connecticut should be home to more small-scale retail establishments. The impact of PH on employment and the total number of establishments, however, is less clear. While under PH small retailers enjoy uniform pricing, these uniform prices are the higher prices that result from non-competitive wholesaler pricing behavior. Having more small retailers in a retail sector that faces lower margins due to high wholesale prices could lead to either more or fewer establishments that overall employ more or fewer workers.

Table E.3 provides some empirical evidence regarding these questions. The regressions presented in Table E.3 are of the same form as the estimation equation above, and describe the impact of PH spirits regulations on three different outcomes: share of small retail establishments, log employment in the liquor retail sector, and log liquor stores per capita.[101]

The uppermost panel of Table E.3 examines the impact of PH regulations on the prevalence of small liquor retailers (that is, establishments with between one and four employees). Column one uses only only data from 2010 and includes demographic controls—state population and median income—and finds a marginally significant positive relationship between PH and share of small liquor retail establishments. Columns two through four use the full panel from 1986 through 2010. Adding state and year fixed effects does not yield a significant coefficient, as shown by column two. Column three adds state-specific time trends, which control for changes in spirits consumption that vary by state. Adding these additional controls reveals that states with PH regulations do in fact have a larger share—4.8 percentage points larger—of small retail establishments. Dropping all states outside of the northeast does not substantively affect the coefficient but increases the precision of the estimate.

The middle panel examines the impact of PH regulations on employment in the alcohol retail sector. The dependent variable is the log of employment in the liquor retail sector per capita age 14 years and older. Looking at data from only 2010 does not suggest a statistically significant relationship between employment and PH laws. Adding year and state fixed effects as shown in column 2 reveals that states with PH laws actually have lower per-capita liquor retail employment. Including state time trends reduces the magnitude and precision of the coefficient from -1.762 (0.198) to -0.497 (0.239). Focusing on northeastern states (column 4) does not have an appreciable further impact on the estimates, though the estimate is less precise.

The bottom panel assesses how the number of establishments per capita is affected by PH regulations. As in the employment panel, examining the 2010 data alone does not suggest a statistically significant relationship between number of retailers and PH laws. Column two uses the full panel with state and time fixed effects, yielding a significant and negative coefficient. Controlling for state time trends reduces the coefficient to -0.608 (0.0914). As in the other panels, examining only northeastern states doesn't appreciably change the coefficient.

[100] For examples of complaints by small retailers, see `https://www.thewesterlysun.com/wire_news/connecticut-s-liquor-law-faces-challenge/article_36891777-e489-56c4-b4f1-c761a30e0059.html`

[101] Panel data describing state liquor retail establishment counts and employment come from the Census County Business Patterns for 1986 through 2010.

Table E.3: Post and Hold Laws and Alcohol Retailing

	2010 Only	All	All	Northeast
Share of 1-4 Employee Retailers	0.0728*	0.0339	0.0477*	0.0472**
	(0.0432)	(0.0209)	(0.0262)	(0.0227)
R-Squared	0.144	0.867	0.940	0.962
Log(Alcohol Employment/Pop 14+)	0.452	-1.762***	-0.497**	-0.422*
	(0.336)	(0.198)	(0.239)	(0.223)
R-Squared	0.064	0.467	0.740	0.821
Log(Liquor Stores Per Capita)	0.344*	-1.335***	-0.608***	-0.515***
	(0.204)	(0.0866)	(0.0914)	(0.103)
R-Squared	0.128	0.855	0.954	0.963
Obs	51	1,275	1,275	300
Demographic Controls	Y	Y	Y	Y
State FE	N	Y	Y	Y
Year FE	N	Y	Y	Y
State Specific Trends	N	N	Y	Y

Note: The table presents coefficients from regression equation E.1 where the outcome of interest is the share of retailers with 1-4 employees in the uppermost panel, the log of employment in the liquor retail sector per capita in the middle panel, and log of liquor stores per capita in the bottom panel. The reported coefficients correspond to a binary variable that is equal to one when spirits are subject to PH regulations. Column 1 uses only data from 2010 and includes demographic controls. Columns 2 through 4 use the full 1986 - 2010 panel. Column 2 adds state and year fixed effects. Column 3 adds state specific time trends and column 4 limits the sample to only northeastern states. Standard errors in parentheses.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

E.3. Comparing Connecticut and Massachusetts

Section 4 reports descriptive evidence comparing prices and consumption patterns in Connecticut and Massachusetts. In addition to being adjacent states that share media markets, the demographic compositions of the states are very similar. Appendix Table E.4 below reports summary statistics from the 2010 American Community Survey. The states have identical gender shares and nearly identical racial composition, with approximately 80% identifying as white. Household income and educational attainment are also closely aligned, with only modest differences in average income and years of schooling. Average age is nearly identical, and the average age among adults is very similar, though Connecticut adults are roughly 10 months older on average.

The demographic similarity of the two states makes Massachusetts a useful non-PH state to compare with Connecticut.

Table E.4: Demographic Characteristics in Connecticut and Massachusetts

	Connecticut		Massachusetts	
Variable	Mean	SD	Mean	SD
Female (1 = Yes)	0.52	0.50	0.52	0.50
White (1 = Yes)	0.78	0.41	0.81	0.39
Age (All)	38.87	22.95	38.61	22.71
Age (18+)	47.70	18.36	46.84	18.36
Household Has Children (1 = Yes)	0.39	0.49	0.37	0.48
Household Income	89,496	102,066	83,246	86,509
Years of Schooling (18+)	13.29	2.87	13.41	3.02

Note: The table presents state means and standard deviations of demographic variables from the 2010 American Community Survey. Individual variables are weighted by person weights while household variables are weighted by household weights.