Discrete Prices and the Incidence and Efficiency of Excise Taxes

Christopher T. Conlon* Nirupama S. Rao†

(Click Here for Current Version)

November 27, 2016‡

Abstract

This paper uses detailed UPC-level data from Nielsen to examine the relationship between excise taxes, retail prices, and consumer welfare in the market for distilled spirits. Empirically, we document the presence of a nominal rigidity in retail prices that arises because firms largely choose prices that end in ninety-nine cents and change prices in whole-dollar increments. Theoretically, we show that this rigidity can rationalize both highly incomplete and excessive pass-through estimates without restrictions on the underlying demand curve. A correctly specified model, such as an (ordered) logit, takes this discreteness into account when predicting the effects of alternative tax changes. We show that explicitly accounting for discrete pricing has a substantial impact both on estimates of tax incidence and the excess burden cost of tax revenue. Quantitatively, we document substantial non-monotonicities in both of these quantities, expanding the potential scope of what policymakers should consider when raising excise taxes.

Keywords: Excise Tax, Incidence, Market Power, Price Adjustment, Nominal Rigidities.

JEL Classification Numbers: H21, H22, H71.

---

*cconlon@stern.nyu.edu, New York University Stern School of Business. 44 W 4th St New York, NY.

†nirupama.rao@nyu.edu, New York University Wagner School of Public Service. 295 Lafayette St New York, NY.

‡The authors would like to acknowledge valuable input and advice from: Wojciech Kopczuk, Bernard Salanie, Jon Vogel, Kate Ho, Juan Carlos Suŕez Serrato, Sebastien Bradley and seminar participants at Columbia University, Wharton BEPP, Duke University, NYU Stern, the National Tax Association 2014 Annual Conference, IIOC 2015, the 2016 NASM of the Econometric Society. The data were provided by the Kilts Center for Marketing at the University of Chicago Booth School of Busines. Any remaining errors are our own.

1 Introduction

The extent to which cost shocks are transmitted to prices is an important question that cuts across disciplines within economics. For example, scholars in international trade study the extent to which exchange rate fluctuations are reflected in prices of final goods. Macroeconomists study the extent to which firms amplify or dampen monetary shocks through the monetary transmission mechanism. Public finance scholars study the extent to which taxes are passed through to prices, in order to determine whether the tax burden falls on consumers or on various firms in the supply chain. Where these literatures diverge is that in macroeconomics and international trade much of the focus is on *nominal rigidities* or *sticky prices*—where seemingly small rigidities, for example menu costs, can have substantial welfare and policy impacts such as monetary non-neutrality.

Nominal rigidities have been notably absent in the study of commodity taxation. Instead, the literature typically assumes that taxes are smoothly passed on to consumers, and exploits envelope conditions to link changes in taxes and the responsiveness of aggregate quantity to welfare; both as far back as Harberger (1964), and in the more recent sufficient statistics approach of Chetty (2009). The well-known theoretical result is that excess burden increases quadratically in the size of the tax, so that deadweight burden per unit of government revenue is a linearly increasing function of the tax. This leads to the policy prescription for commodity taxation of "lots of little taxes" on goods of similar elasticity rather than concentrated taxes on one such good. Similarly, the standard approach links the pass-through rate obtained from a regression of changes in prices on changes in taxes to the incidence of the tax burden. The convention is to assume that the pass-through rate is approximately constant, such that the distribution of the tax burden would be roughly the same for a somewhat larger (or smaller) tax.

We study the effects of recent increases in the excise tax for distilled spirits in Connecticut and Illinois using high-quality scanner data from Nielsen. Distilled spirits are one of the most heavily taxed commodities in the United States, with the combined state and federal tax burden comprising as much as 30-40% of the retail price.[1] Empirically, we document that when taxes rise, more than 60% of prices do not change; but that when prices do change 77% of price changes are in whole-dollar increments, most commonly $1. The source of these rigidities appears to be that 90% of distilled spirits prices in Connecticut (and over 80% in Illinois) end in ninety-nine cents. Similar pricing patterns have been documented for wide range of products (Levy, Lee, Chen, Kauffman, and Bergen, 2011). Theoretically, we show that discrete price changes can generate either incomplete or excessive pass-through for the same underlying demand. In response to tax changes, we document very little heterogeneity in the magnitude of price changes when prices changed, but substantial heterogeneity in the identity of firms and products that change prices.

[1] New York is the highest tax state for both cigarettes and gasoline. The combined state and federal tax burden in New York state for gasoline is 60.72 cents per gallon of gasoline ($2.43 per gallon average retail price); $5.85 in taxes on a pack of cigarettes ($14 average retail price); and $7.95 in taxes on a 1.75L bottle of 80 Proof Vodka (retail price as low as $14.99).

We are more likely to observe a price increase when: (1) products have higher prices, so that a \$1 price change is smaller in percentage terms (2) a particular store sells the same product for lower prices than its competitors (3) wholesale prices have increased substantially since the last retail price change. Quantitatively, these facts suggest that price changes are discrete, and that the relationship between price changes and tax changes is likely to be nonlinear. A correctly specified model of price responsiveness to tax changes such as an (ordered) logit or probit would account for discreteness, nonlinearity, and heterogeneity of response.

In order to understand the welfare effects of discrete price changes, consider the following illustrative example. A small increase in taxes of one or two cents might generate no \$1 price increases, would be borne entirely by firms, and would generate no deadweight loss. A medium-sized tax increase of 50 cents might generate a large number of \$1 price increases, and if more than half of firms increased their prices, imply a pass-through rate in excess of 100%. It would appear as if consumers bear the entirety of the tax burden, and the excess burden of taxation might be very large. If all prices responded to a 50 cent tax increase with a \$1 price increase, then a further increase in the tax to 90 cents might generate no additional price increases, would begin to shift the tax burden back onto firms, and would increase the revenue collected by the tax by 80% without increasing the excess burden. By comparison, the conventional approach would apply a pass-through rate of 200% to all other tax increases. This would imply that a 90 cent tax increase would raise prices by \$1.80, thus more than doubling the excess burden.

Accounting for these rigidities has important implications for welfare and public policy. First, it suggests that the relationship between tax changes and price changes is best described by a (series of) S-shaped curves. Second, it suggests that the slope of the curve, the pass-through rate (and incidence) is a nonlinear and non-monotonic function of the tax change. Third, it suggests that the excess burden costs of raising government revenue are not linearly increasing in the size of the tax, but are a non-monotonic (U-shaped) function of the tax change. Taken together, this implies a much broader practical scope for optimal commodity taxation than "lots of little taxes" would suggest. There potentially exist levels of taxes for which government revenue is increasing more quickly than excess burden. Because only changes in consumer prices have the potential to generate deadweight loss, a policy maker can reduce the *average* social cost of taxation by choosing tax levels that shift more of the burden onto firms and less onto consumers. The practical implication is that the policy maker should choose tax increases that are either just to the left or well to the right of the steepest part of the "S"-curve.

Our paper is organized as follows: after reviewing the relevant literature, we provide an overview of the distilled spirits industry in Section 2. In Section 3, we introduce a simple theoretical framework and compare expressions for pass-through, incidence, and welfare when price changes are discrete with those when price changes are continuous. In Section 4 we provide descriptive information about the data used in our analysis. Section 5 consists of two main parts: the first reproduces the common OLS estimates of the pass-through rate, and like the previous literature finds a pass-through rate in excess of 100%, and conditional on a price change a pass-through rate as high as 300% of the tax increase. Our second set of results focuses on properly accounting for the discreteness of the price change and shows that both incidence and the excess burden cost of taxation vary with the size of the tax in a non-monotonic way. Finally, we conclude by discussing the implications these non-monotonicities have for commodity tax policy.

Relation to Literature

A relatively new literature in public finance exploits *notches* or discontinuities in the labor income tax schedule, which lead to bunching on one side of the notch and gaps on the other. For example, Kleven and Waseem (2013) examine a tax notch where for some individuals, an additional rupee of income triggers an additional tax payment of 12,500 rupees, leading to a lower post-tax income. Despite the stark incentives created by notches, some agents locate in the dominated regions of their choice sets. Such behavior is rationalized via *optimization frictions*. The notches are used as an exogenous source of variation in order to recover structural elasticities ((Chetty, 2012), (Slemrod, 2013) and (Kopczuk and Munroe, 2015)). We take a different, though complementary, approach from the literature on optimization frictions. Instead of exploiting discontinuities in the tax schedule as a source of *exogenous* variation to recover a frictionless, long-run structural elasticity, we explicitly model the *endogenous* but discontinuous pricing strategies of firms in response to taxes. Our goal is not to recover the structural relationship between prices and taxes that would arise in a frictionless world, but rather to model how pricing rigidities would respond to alternative tax policies, and to understand the welfare implications in such a world.

In spirit our paper is most related to Nakamura and Zerom (2010) and Goldberg and Hellerstein (2013). Both of those papers allow for nominal rigidities and examine the transmission of cost shocks into prices using scanner data, though neither paper studies the impact of taxes. Nakamura and Zerom (2010) examine the pass-through of cost shocks in the coffee industry and find only 23% of commodity price shocks are reflected in retail coffee prices. Further, they find that pass-through is sluggish, with the delay almost entirely attributable to pricing behavior at the wholesale rather than retail level. Goldberg and Hellerstein (2013) find that just 7% of exchange rates shocks are passed on to retail beer prices and decompose the sources of incomplete pass-through. In contrast with these findings, we find that pass-through of excise taxes on distilled spirits is nearly immediate (within a month) and is often over-shifted (in excess of 100%).[2] One possible discrepancy is that exchange rate or commodity price fluctuations are likely to be perceived as *transitory*, while excise tax changes are likely to be perceived as *permanent*.

[2] These pass-through results are not directly comparable. Because we measure the impact of a per-unit excise tax and are worried about $1 price changes, we measure pass through of taxes in dollars to retail prices in dollars or *dollar for dollar* pass-through. The aforementioned studies measure pass through by examining percentage changes in retail prices as caused by percentage changes in input prices or exchange rates.

Unlike these two papers, ours does not follow an explicit structural approach, though it has a structural interpretation which we provide in section 3.3. The first step of the Bajari, Benkard, and Levin (2007) estimator is to recover the agents' policy functions. In our context, this amounts to modeling the discrete price change as a function of potential state variables such as the tax change or previous input price changes. If we wanted to measure the relative impact of menu costs and whole dollar pricing increments, we would need to place more structure on the problem in order to separately identify both potential sources of nominal rigidities. Because our counterfactuals do not require explicit estimates of the dynamic parameters (such as the menu cost), we show that we can work directly with the policy functions.

There is a large literature that looks to measure both the elasticity of distilled spirits consumption with respect to prices, $\frac{\partial Q}{\partial P}$, and with respect to taxes, $\frac{\partial Q}{\partial \tau}$. This literature is summarized in recent meta-analyses by Wagenaar, Salois, and Komro (2009) and Nelson (2013), which report an average elasticity between $(-0.8, -0.5)$. The majority of the papers in those meta-analyses use aggregate state-level data on distilled spirits shipments; most do not observe prices directly, but rather use taxes as a proxy for prices and assume that taxes are fully passed on to consumers i.e.: $\frac{\partial P}{\partial \tau} = 1$ so that $\frac{\partial Q}{\partial P} \approx \frac{\partial Q}{\partial \tau}$.

Consistent with our descriptive evidence, previous studies using a small number of brands have found that alcohol taxes are over-shifted (pass-through $\frac{\partial P}{\partial \tau} > 1$). Cook (1981) used price data for leading brands to calculate average yearly prices for each state. He found that the median ratio of price change to tax change for the 39 state-years that had tax changes was roughly 1.2. Young and Bielinska-Kwapisz (2002) followed the prices of seven specific alcoholic beverage products and estimated pass-through rates ranging from 1.6 to 2.1. When Alaska more than doubled its alcohol taxes in 2002, Kenkel (2005) reported that the associated pass-through was between 1.40 to 4.09 for all alcoholic beverages, and between 1.47 to 2.1 for distilled spirits. All three studies report substantial product-level heterogeneity in the degree of pass-through.

The over-shifting of taxes is not limited to taxes on distilled spirits, nor are all excise taxes over-shifted. Poterba (1996) found that retail prices of clothing and personal care items rise by approximately the amount of sales taxes. Besley and Rosen (1999) could not reject that prices increased by roughly the amount of the sales tax for some goods (including Big Macs, eggs, Kleenex, and the board-game Monopoly), but found evidence of over-shifting for more than half of the goods they studied. In gasoline, where price increments are very small (often one cent) relative to tax changes, Marion and Muehlegger (2011) found that gasoline and diesel taxes are on average fully passed through to consumers, but found evidence of incomplete pass-through when supply was inelastic or inventories were high. Doyle Jr. and Samphantharak (2008) showed that while pass-through of a gasoline tax holiday happens within a week, the tax reduction is only partially passed on to consumers. Harding, Leibtag, and Lovenheim (2012) found that cigarette taxes were less than fully passed through to consumers, while DeCicca, Kenkel, and Liu (2013) could not reject full pass-through of cigarette taxes on average, but found that consumers willing to shop faced significantly less pass-through. In another retail setting where discrete pricing increments may be important, Besanko, Dubé, and Gupta (2005) found that 14% of wholesale price-promotions were passed on at more than 100% into retail prices.

Estimated pass-through rates exceeding unity, like those we estimate, have been theoretically justified by the presence of imperfect competition among suppliers and curvature restrictions on demand.[3] Most of those results employ a single-product homogenous good framework. Katz and Rosen (1985), Seade (1985), and Stern (1987) all rely on Cournot competition with conjectural variations. Besley (1989) and Delipalla and Keen (1992) employ a Cournot model with free entry and exit. This early literature is summarized by Fullerton and Metcalf (2002). Because Cournot competition may not be a realistic assumption for many taxed goods including distilled spirits, Anderson, De Palma, and Kreider (2001) develop similar results under differentiated Bertrand competition.[4] The common thread is that demand must be sufficiently (log)-convex in order to generate overshifting. However, as Anderson, De Palma, and Kreider (2001) point out, as demand becomes too convex, the marginal revenue curves may no longer be downward sloping. Fabinger and Weyl (2012) categorize the pass-through rate and marginal revenue properties of several well-known demand systems, and show that satisfying both properties is difficult but possible under certain forms of Frechet and almost ideal demand systems (AIDS) (Deaton and Muellbauer, 1980). In a recent and notable departure from this theoretical literature, Hamilton (2009) finds that excise taxes can lead to overshifting when demand is sufficiently *concave* rather than convex and consumers purchase multiple (taxed) goods at the same time. The result depends on two things: higher taxes must lead to reduced variety of product offerings; and strategic complementarities between prices and variety.

We observe that excises taxes are over-shifted nearly immediately, and do not observe exit of brands in response to the tax; thus these long-run explanations of over-shifting are not well matched to our data. Instead, we show that in the presence of a nominal rigidity (such as whole-dollar price changes) we can generate over-shifting or under-shifting of taxes without restrictions on the log-convexity of the underlying demand curve.

[3] For example, Besley and Rosen (1999) state: "An important implication of this literature is that in an imperfectly competitive market, varying degrees of shifting are possible in the long run. Indeed, even over-shifting is a distinct possibility; i.e., the price of the taxed commodity can increase by more than the amount of the tax. These results contrast markedly with those that emerge from a competitive model. With competition, after-tax prices increase by just the amount of the tax if the long-run supply curve is horizontal, and by less than the amount of the tax if the supply curve is upward sloping."

[4] Homogenous Bertrand competition leads to marginal cost pricing and hence $\rho = 1$, which is not particularly interesting.

2 Alcohol Taxation and Industry Background

Compared with other commodities, alcoholic beverages (along with cigarettes and gasoline) are subjected to unusually high taxes.[5] We focus primarily on distilled spirits, in part because they bear a substantially larger tax burden than beer or wine. In 2010, total tax revenues on distilled spirits contributed \$15.5 billion in tax revenue for an industry in which production, distribution and retailing amount to roughly \$120 billion in revenue.

As a consequence of the 21st Amendment, states are free to levy their own taxes on spirits, as well as regulate the market structure in other ways. There are 18 *control states*, where the state has a monopoly on either the wholesale distribution or retailing of alcohol beverages (or both).[6] Connecticut, Illinois and the other thirty states are *license states*. License states follow a three-tier system where vertically separated firms engage in the manufacture, wholesale distribution, and retailing of alcohol beverages. Almost all license states have restrictions that prevent distillers from owning wholesale distributors, or prevent wholesale distributors from owning bars or liquor stores. In Connecticut and Illinois wholesalers and retailers are fully distinct.[7]

The Connecticut state regulator forbids wholesalers from engaging in temporary sales, price promotions or giveaways. The retail price data reveal few if any temporary sales; ignoring the first week of the month, there is virtually no within product-store-month price variation in Connecticut. The lack of temporary sales makes the distilled spirits market an attractive market in which to study pass-through since we do not need to distinguish between general price changes and short-term markdowns.[8]

Since 1991, the federal government has taxed distilled spirits at \$13.50 per proof-gallon, or \$4.99 for a 1.75L bottle of vodka at 80 proof.[9] The statutory incidence of *federal* excise taxes falls on the *producers* of distilled spirits or is due upon import into the United States. Our paper focuses on *state excise taxes* which are *remitted by wholesalers* and usually levied on volume rather than

[5] The taxation of alcoholic beverages has a long history in the United States. The first taxes on domestic alcoholic beverage production were collected in 1791 in order to pay off the debts of the Revolutionary War.

[6] The monopoly applies to all alcohol beverages in some states, and in others to distilled spirits but not wine or beer. Control states can adjust markups or taxes to raise revenue. A few control states, such as Maine and Vermont, nominally control the distribution and sales of spirits but contract with private firms which set prices. Control states have been the subject of recent empirical work examining the entry patterns of state-run alcohol monopolies (Seim and Waldfogel, 2013) and the effects of uniform markup rules (Miravete, Seim, and Thurk, 2014).

[7] States have other restrictions on the number of retail licenses available, or the number of licenses a single chain retailer can own. States also differ on which types of alcoholic beverages, if any, can be sold in supermarkets and convenience stores. Prior work on license states has examined the welfare effects of exclusivity arrangements in the beer industry (Asker, 2005).

[8] How one handles temporary sales is one of the principal challenges in the empirical macroeconomics literature on estimating menu costs. See Levy, Bergen, Dutta, and Venable (1997), Slade (1998), Kehoe and Midrigan (2007), Nakamura and Steinsson (2008), Eichenbaum, Jaimovich, and Rebelo (2011), Eichenbaum, Jaimovich, Rebelo, and Smith (2014).

[9] Taxes are stated in customary units of gallons, though products are sold internationally in standardized metric units of 750mL, 1L, and 1.75L bottles. A proof-gallon is 50% alcohol by volume (100 Proof) at 60 degrees Fahrenheit.

ethanol content .[10] In many states, including the two we study, retailers are expected to remit sales taxes on alcoholic beverage purchases. In some states, there is an additional sales tax that applies only to alcoholic beverages, while in others alcoholic beverages are exempt from the general sales tax.

All of these taxes are of course levied in part to address the negative health and public safety externalities of alcohol. However, governments also tax alcohol for the explicit purpose of raising revenue.[11] Few states changed their alcohol taxes over the prior decade, but following the onset of the Great Recession seven states passed legislation between 2007 and 2013 affecting alcohol taxes. We report those tax changes along with the detail of the tax change in Connecticut in Table 1. Prior to July 1, 2011 the state of Connecticut levied a (wholesale) tax on the volume of distilled spirits (independent of proof) of \$4.50 per gallon, which worked out \$2.08 per 1.75L bottle.[12] After July 1, 2011 the tax increased to \$5.40 per gallon or \$2.50 on a 1.75L bottle. The tax increase amounted to \$0.178 for 750mL products, \$0.238 for 1L products and \$0.416 for 1.75L products, which provides valuable cross-sectional variation in the tax change. In September of 2009, Illinois increased its excise tax from \$4.50 per gallon to \$8.55 per gallon, for an additional \$0.80 on 750mL bottle of vodka and \$1.87 on a 1.75L bottle of vodka.

An interesting provision of the Connecticut tax increase ensured that the tax was uniform on all units sold after July 1, 2011: retailers (and wholesalers) were subjected to a *floor tax* on unsold inventory as of July 1, 2011. This floor tax helps us measure the tax incidence as it makes the tax increase immediate on all units.[13] It should be noted that Connecticut also increased its sales tax from 6% to 6.35% at the same time as the alcohol excise tax increase. Our empirical analysis examines the impact of the specific tax increase on sales tax-exclusive retail and wholesale prices. As the sales tax is levied at the time of retail sale, any pass-through of the 0.35% increase would lead to lower retail, and perhaps wholesale, prices. Thus the pass-through rates we report for Connecticut, potentially under-estimate the true pass-through rates. Chetty, Looney, and Kroft (2009) suggest sales tax increases (in alcoholic beverages) may be substantially less salient to consumers than taxes which are included in the posted-price (as much as 30% salient and as little as 0.6%). The small magnitude of the sales tax increase and the fact that estimates of the retail pass-through rate using sales tax inclusive prices are mechanically larger but statistically indistinguishable from the results

[10] Distillers, wholesalers and retailers are also subject to federal and state corporate income taxes.

[11] For example, in 2015 Governor Sam Brownback of Kansas proposed raising alcohol and tobacco taxes to help close the state's \$648 million budget shortfall. For more details see `http://www.kansas.com/news/politics-government/article6952787.html`. In 2016, Governor Jon Bel Edwards of Louisiana proposed a similar tax increase which would raise \$27 million, as part of reducing a \$900 million deficit, see `http://www.nola.com/politics/index.ssf/2016/03/house_passes_new_alcohol_tax_h.html`.

[12] Many states levy lower tax on lower proof ready-to-drink products, or lower proof schnapps and liquers. Products less than 7% A.B.V. in Connecticut are subjected to a lower tax rate.

[13] The floor tax meant that any product not in the hands of consumers would be subjected to the new tax rate rather than the old tax rate, and prevented retailers from evading the tax by placing large orders in advance of the tax increase. It did not, however, prevent consumers from stockpiling alcoholic beverages in advance of the tax increase, though as Table 2 shows, there is no evidence of an anticipatory price effect.

presented here make this simplification less concerning.[14]

3 Conceptual Framework

In evaluating a potential change in excise taxes, we are interested in two quantities: *incidence* which compares the relative losses in consumer and producer surplus, and the *cost of public funds* which compares the increased deadweight loss to the government revenue raised. A key input in measuring both quantities empirically is the pass-through rate, though the precise relationship depends on the nature of competition.

Consider the well known example of an introduction of a per-unit tax τ into a perfectly competitive market. After the introduction of the tax, consumers pay p_c; firms remit the tax and receive $p_c - \tau = p_r$.[15] In equilibrium, quantity demanded will equal quantity supplied: $D(p_c) = S(p_c - \tau)$. The market generates consumer surplus $CS(p) = \int_p^\infty D(x)\,\mathrm{d}x$ and producer surplus $PS(p-\tau) = \int_0^{p-\tau} S(x)\,\mathrm{d}x$. An infinitesimal tax increase reduces consumer and producer surplus: $\frac{dCS}{d\tau} = -\frac{dp}{d\tau}Q$ and $\frac{dPS}{d\tau} = -(1-\frac{dp}{d\tau})Q$, where Q is the equilibrium quantity. Using ρ to describe the pass-through rate, $\frac{dp}{d\tau}$, the incidence of an infinitesimal tax change will be:

$$I = \frac{\frac{dCS}{d\tau}}{\frac{dPS}{d\tau}} = \frac{\rho}{(1-\rho)} \tag{1}$$

Fabinger and Weyl (2013) extend this well-known incidence formula to incorporate varying market structures $I = \frac{\rho}{1-(1-\theta)\rho}$, where θ functions as a conduct parameter with $\theta = 0$ indicating perfect competition ($I = \frac{\rho}{1-\rho}$); and $\theta = 1$ indicating monopoly ($I = \rho$).[16] They also provide an extension for a finite (rather than infinitesmal) tax increase which allows the pass-through rate to vary with the tax:

$$\overline{I}(\Delta\tau) = \frac{\overline{\rho}}{1-\overline{\rho}+\overline{\theta\rho}}, \quad \overline{\rho} = \frac{\int_{t_0}^{t_0+\Delta\tau}\rho(t)Q(t)dt}{\int_{t_0}^{t_0+\Delta\tau}Q(t)dt}, \quad \overline{\theta\rho} = \frac{\int_{t_0}^{t_0+\Delta\tau}\theta(t)\rho(t)Q(t)dt}{\int_{t_0}^{t_0+\Delta\tau}Q(t)dt} \tag{2}$$

Taking the expression in (2) to data would require recovering the pass-through rate $\rho(t)$, demand $Q(t)$, and even conduct $\theta(t)$ as a function of potential tax changes and then integrating up. As we show later, when price changes are discrete, $\rho(t)$ will take on a (specific) nonlinear form.

In most empirical studies of tax pass through, $\hat{\rho}(\cdot)$ is estimated via regression of changes in

[14] For example, over the one-month horizon the sales tax-inclusive retail pass-through rate is 1.856 (0.283), which is higher but statistically indistinguishable from the pass-through rate reported in column 1 of Table 5, 1.533 (0.271).

[15] The statutory incidence of the tax could of course be on consumers instead in which case the price charged by retailers could potentially be lower than the no-tax equilibrium price to reflect pass-through from consumers to producers. Because all state and federal alcohol taxes are remitted by suppliers, we focus on the case where the firms bear the statutory burden.

[16] A well known challenge in the conduct parameter approach is that intermediate cases $\theta \in (0, 1)$ can be difficult to interpret, except for a few examples such as symmetric Cournot competition.

prices, Δp_{jt}, on changes in corresponding taxes, $\Delta\tau_{jt}$, where j indexes products and t measures time, with various fixed effects and other controls x_{jt}:

$$\Delta p_{jt} = \rho_{jt}(\mathbf{X}, \Delta\tau) \cdot \Delta\tau_{jt} + \beta\Delta x_{jt} + \gamma_j + \gamma_t + \epsilon_{jt} \tag{3}$$

Here we have written the pass-through rate $\rho_{jt}(X, \Delta\tau)$ as a general function that might depend on (j, t) as well as other covariates $\mathbf{X}$, or the size of the tax change itself, $\Delta\tau$. Most of the literature assumes that for "small" tax changes, pass-through is approximately constant and avoid the integral in (2). For example, Besley and Rosen (1999) and Harding, Leibtag, and Lovenheim (2012) assume: a single pass-through rate $\rho_{jt}(X, \Delta\tau) = \rho$, or product specific pass-through $\rho_{jt}(X, \Delta\tau) = \rho_j$ respectively. A series estimator would provide a flexible semi-parametric estimate of the pass-through rate: $\rho_{jt}(X, \Delta\tau) = \alpha_0 + \alpha_1\Delta\tau_{jt} + \alpha_2\Delta\tau_{jt}^2 + \cdots$. In the limit, the series estimator is sufficiently flexible to capture the nonlinear relationship between tax changes and discrete price changes.

In practice, identification of $\rho_{jt}(X, \Delta\tau)$ can be challenging, especially when we observe a small number of points in the support of $\Delta\tau_{jt}$. For a single tax change, the only variation in $\Delta\tau_{jt}$ comes from cross-sectional variation in the package size (750mL, 1L, 1.75L). We could exploit multiple tax changes (either across states or over time). However, when we pool across states we need to impose additional assumptions on the other covariates $\mathbf{X}$ (such as the market-structure or pre-existing cost variation). With limited variation in $\Delta\tau_{jt}$, rather than assuming constant pass-through: $\rho_{jt}(X, \Delta\tau) = \rho$ or attempting to recover the function $\rho_{jt}(X, \Delta\tau)$ at all values of $\Delta\tau$, we suggest exploiting the discrete nature of price changes. We suggest modelling the *probability of a discrete change* in Δp_{jt} using an (ordered) logit or probit, which would impose a particular nonlinear structure on $\rho_{jt}(\mathbf{X}, \Delta\tau)$.[17]

3.1 Pass-Through with Discrete Price Changes

In the simplest example of discrete pricing, the firm can either increase its price by a single unit, or can keep the same price. In practice it is easier to model the change in price Δp_{jt} rather than the level:

$$\Delta p_{jt} = \begin{cases} 1 & \text{if } \Delta\tau_{jt} \geqslant \overline{\tau}_{jt}(\mathbf{X}) \\ 0 & \text{if } \Delta\tau_{jt} < \overline{\tau}_{jt}(\mathbf{X}) \end{cases} \tag{4}$$

For a large enough tax increase, the firm will always increase its price, and for a small enough tax increase the firm will keep its existing price. For each product there is a threshold level of the tax increase, $\overline{\Delta\tau}_{jt}$, beyond which the price is increased. This would imply that the true function

[17] Further, a discrete choice model that reflects the underlying pricing behavior also allows us to directly employ controls that adjust for the fact that different products are at different distances from their *(s,S)* boundary; even with infinite support in $\Delta\tau_{jt}$ approximating the discrete choice model with a high-order polynomial will also require interacting each of the polynomial terms with the covariates to capture these differences.

$\rho_{jt}(X, \Delta\tau) = \delta_{\overline{\tau}_{jt}(\mathbf{X})}(\Delta\tau_{jt})$ where $\delta_z(\cdot)$ is the Dirac delta function with point mass at z. Then for each product we can compute it's product specific pass-through rate as $\hat{\rho}_j = \frac{1}{\Delta\tau_{jt}} \int_0^{\Delta\tau_{jt}} \delta_{\overline{\tau}(X)}(\Delta\tau_{jt})$. This pass-through rate takes on only two values for each j: $\frac{1}{\Delta\tau_{jt}}$ or 0. Ignoring other covariates, the OLS estimates of the pass-through rate from (3) are a weighted average of product level pass-through $\hat{\rho} = \sum_{jt} w_{jt}\hat{\rho}_j$.

This demonstrates how it is possible to generate either incomplete pass-through or over-shifting. For example, if $\Delta\tau = 0.25$ and products are equally weighted, $w_j = \frac{1}{J}$, then $\hat{\rho}$, our reduced-form estimate of pass-through, would be a weighted average of $\rho_j = 4$ and $\rho_j = 0$. As long as more than $\frac{1}{4}$ products increase their price in response to the tax, it is possible to estimate $\hat{\rho} \geqslant 1$ without imposing log-convexity on the demand function, while if fewer products increase their price we will find incomplete pass-through ($\hat{\rho} < 1$).

In general we do not expect the econometrician to observe $\overline{\Delta\tau}_{jt}(\mathbf{X})$ directly, and instead it must be estimated. If we allow for some econometric error in $\overline{\Delta\tau}_{jt}(\mathbf{X})$ that is IID and type I extreme value: then this suggests that the correct estimator for $\hat{\rho}_{jt}(X, \Delta\tau)$ is just the predicted probability from a logit (divided by the tax increase): $\frac{Pr(\Delta p=1|\mathbf{X},\Delta\tau)}{\Delta\tau}$. Likewise, we could extend this model to an ordered logit and allow for larger price increases (or price decreases). In our main empirical specification we allow $\Delta p_{jt} \in \{-1, 0, 1, 2, 3\}$. This requires estimating the "cut points" of the ordered choice model, which we can translate into units of the tax: $\overline{\Delta\tau}_{-1,jt}(\mathbf{X})$, $\overline{\Delta\tau}_{1,jt}(\mathbf{X})$, $\overline{\Delta\tau}_{2,jt}(\mathbf{X})$ and $\overline{\Delta\tau}_{3,jt}(\mathbf{X})$.

Figure 1 illustrates the price response with a binary logit. Because the x-axis represents the tax change, and the y-axis represents the price change, the OLS estimate of the pass-through rate is the slope of a ray intersecting the sinusoidal curve, $\hat{\rho} = \frac{\Delta p}{\Delta\tau}$. We draw the complete pass-through $\rho = 1$ line in yellow for reference. For a small tax increase (red line), it might be that very few products change prices so that the estimated pass-through rate is small. For a very large tax increase (blue line) it may be that most products increase their prices by \$1, but this might be smaller (or larger) than the denominator, $\Delta\tau$. For some intermediate value of the tax increase, it might be that most products increase their prices, but that $\Delta\tau$ is not so large, leading to a higher estimate of the pass-through rate. If we plot the slope of each ray to the same S-shaped curve, we would find that the implied pass-through rate is U-shaped: rising during the steep part of the S-curve and falling over the flat parts. If we expanded the support of potential tax changes (and the potential outcomes of our price change model), price changes would follow a increasing series of S-shaped curves, and pass-through would follow a series of U-shaped curves.

With enough variation in the support of $\Delta\tau$, the linear model from (3) can trace out any relationship between tax changes and price changes $\rho_{jt}(X, \Delta\tau)$, including a nonlinear relationship with discrete price changes such as an (ordered) logit. By exploiting our knowledge that price changes are discrete, we can impose some nonlinearity on the relationship between price changes and tax changes *ex ante*, and obtain more reasonable estimates with limited support for $\Delta\tau$. This is especially important when we want to forecast for tax changes not observed in the data. If we estimate a large pass-through rate for an observed tax increase, the linear model would apply that pass-through rate to a larger increase; the nonlinear model might interpret a large pass-through rate as the top of the U-shaped curve and anticipate less pass-through for larger tax changes. This is crucial for the next section, in which develop the welfare implications, because declining pass-through rates mean not only that taxes fall more on firms than on consumers, but also that they generate less deadweight loss per dollar of government revenue.

3.2 Incidence and Excess Burden of Discrete Changes

We focus on two key welfare measures: the incidence, which measures the extent to which the tax burden is borne by consumers or firms; and the social cost of taxation, which measures how much deadweight loss is generated per dollar of government revenue. In the standard framework where prices continuously respond to infinitesimal taxes, the incidence is determined by the pass-through rate and assumed to be constant, while the social cost of taxation is linearly increasing in the size of the tax. When pricing is discrete, we show that both incidence and social cost of taxation can be increasing or decreasing in the size of the tax.

We let (P_0, P_1) and (Q_0, Q_1) be the price and quantity before and after a tax increase of $\Delta\tau$ respectively, and let $\Delta Q = Q_1 - Q_0$ and $\Delta P = P_1 - P_0$. We use the traditional linear approximations and constant marginal costs as illustrated in Figure 2, where supply is characterized by monopoly, to derive the following expressions. These expressions are approximations in the sense that the demand curve need not be linear, and thus the deadweight loss or excess burden of taxation is not exactly a triangle. The incidence of the tax is given by:

$$I(\Delta\tau) = \frac{\Delta CS(\Delta\tau)}{\Delta PS(\Delta\tau)} \approx \frac{\Delta P \cdot Q_1 + \frac{1}{2}\Delta P \cdot \Delta Q}{(P_0 - MC) \cdot \Delta Q + (\Delta P - \Delta\tau) \cdot Q_1} \tag{5}$$

And the social cost of the additional tax revenue is given by the following expression:

$$SC(\Delta\tau) = \frac{\Delta DWL(\Delta\tau)}{\Delta GR(\Delta\tau)} \approx \frac{\frac{1}{2}\Delta P \cdot \Delta Q + (P_0 - MC) \cdot \Delta Q}{\Delta\tau \cdot Q_1} \tag{6}$$

There are a number of options to take (5) and (6) to the data. For (5), we can follow our incidence expression in (2): $\frac{\overline{\rho}}{1-\overline{\rho}+\overline{\theta\rho}}$ and estimate the pass-through rate $\rho(t)$ along with some assumption about the conduct θ such as perfect competition $\theta = 0$. Alternatively, under perfect compeititon we can write the pass-through rate as a function of the elasticities of demand (ϵ_D) and supply (ϵ_S): $\rho = \frac{\epsilon_S}{\epsilon_S - \epsilon_D}$, and then estimate (or calibrate) those quantities from data. Suppose that we estimate $\hat{\rho} = 1.5$, as we do for alcohol taxes. Under perfect competition this would imply that: $I = \frac{\rho}{1-\rho} = -3$, or that $\varepsilon_S < 0$. This would imply negative incidence (whatever that means) and that firms supply less as prices increase, neither of which makes any sense. We don't fare much better under monopoly incidence: $I = \rho = 1.5$, which implies that consumers bear 150% of the welfare burden of taxation.

Likewise for the expression in (6) under perfect competition and smooth prices we can write the deadweight loss as: $\Delta DWL = \frac{1}{2} \cdot \Delta Q \Delta \tau$. Excess burden is about the *allocation* of goods to consumers, so it is about capturing the quantity response to the tax. When all adjustments are smooth $\Delta Q \approx \frac{\partial Q}{\partial P} \cdot \frac{\partial P}{\partial \tau} \Delta \tau$. If we assume perfect competition, we can re-write the pass-through rate using elasticities ($\rho = \frac{\epsilon_S}{\epsilon_S - \epsilon_D}$) to obtain: $\frac{\Delta DWL}{\Delta GR} = \frac{1}{2} \cdot \frac{\epsilon_S \, \epsilon_D}{\epsilon_S - \epsilon_D} \cdot \frac{\Delta \tau}{P_0} = \frac{1}{2 \cdot P_0} \epsilon_S \cdot \rho \cdot \Delta \tau$. This provides the well known result that the excess burden cost of raising a dollar of government revenue is increasing in both the size of the tax ($\Delta \tau$) and the pass-through rate ρ. It suggests that the cost of raising additional revenue is increasing, and that when the pass-through rate is high, the excess burden costs of revenue are rising quickly. One approach would be to estimate the pass-through rate directly and estimate (or calibrate) elasticity of supply. The second approach is to follow the suggestion of Harberger (1964) or Chetty (2009) and note that $\Delta DWL = \frac{1}{2} \frac{dQ}{d\tau} (\Delta \tau)^2$. Again, if we divide this by $GR = \Delta \tau \cdot Q_1$ we have that the social cost of taxation grows linearly in the size of the tax. This approach requires that we estimate the quantity response to the tax: $\frac{dQ}{d\tau}$. When price changes are discrete, different sized tax changes will imply different estimates for the relationship between quantity and taxes: $\frac{dQ}{d\tau}(\Delta \tau) \neq \frac{dQ}{d\tau}(\Delta \tau')$ when $\Delta \tau \neq \Delta \tau'$. If there is no price change in response to the tax, then there should not be a quantity response either. Conversely, when there is a \$1 price response to a small tax, the measured quantity response may be large (particularly if the prices of close substitutes do not adjust). We would not want to apply either estimate to predict the welfare effects of a differently sized tax change.

While smooth elasticity-based formulas are likely to be misleading for discrete price changes, all is not lost. Instead we can focus directly on the expressions in (5) and (6). We can obtain $\widehat{\Delta P}(\Delta \tau)$ directly from our (ordered) logit model. This leaves only the post-tax quantity $Q_1(\Delta P(\Delta \tau))$ to estimate:

$$\widehat{Q_1}(\Delta \tau) = Q_0 + \frac{P_0}{Q_0} \cdot \epsilon_D \cdot \widehat{\Delta P}(\Delta \tau) \tag{7}$$

Given an initial quantity Q_0 and the predicted price change as a function of the the tax $\Delta P(\Delta \tau)$, all that remains to evaluate (5) and (6) is an estimate of the demand elasticity ϵ_D. As demand becomes more elastic, the quantity response will be larger and thus we expect to find more deadweight loss per dollar of tax revenue raised. Likewise, as demand becomes more elastic, we expect firms rather than consumers to bear a larger share of the tax burden.

In our empirical exercise, we consider a wide range of demand elasticities and report results for $\epsilon_D \in \{-0.5, -1.0, -1.5\}$ which span the elasticities reported in the meta-analysis by Wagenaar, Salois, and Komro (2009), and are consistent with those we obtain from the Nielsen data. The non-monotonicity and nonlinearity of $I(\Delta \tau)$ and $SC(\Delta \tau)$ come from the discreteness in the price response $\Delta P(\Delta \tau)$, not the magnitude of the quantity response ϵ_D. The elasticity only scales the two welfare measures $I(\Delta\tau)$ and $SC(\Delta\tau)$ up or down proportionally. Our main focus is the *shape* of $I(\Delta\tau)$ and $SC(\Delta\tau)$ with respect to $\Delta\tau$.

It is important to state four caveats. The first is that the welfare expressions we have presented are for a single product, and thus assume monopolistic competition at the retailer-product level. In general, own price elasticities tend to be much larger than category level elasticities when consumers can substitute within a category. If two products are nearly perfect substitutes, there might be no welfare loss if the price of A increased by \$1.00 while the price of B remained unchanged. However, if there were no close substitute for A, then a \$1.00 price change could be met with a substantial welfare reduction. In order to fully capture these differences in the presence of discrete price changes, we would need a fully specified multi-product demand system such as Berry, Levinsohn, and Pakes (1995) or Deaton and Muellbauer (1980).[18] In our other paper (Conlon and Rao, 2015), we estimate a structural multi-product demand system and using those results, obtain qualitatively similar welfare results to those we present. Again, the discreteness of $\Delta P(\Delta\tau)$ drives the results, not the magnitude of the quantity response.

Our second caveat is that (5) and (6) represent the familiar textbook approximation to welfare under linear demands, which hold for *small price changes*. However, we consider *discrete price changes* of \$1.00, \$2.00, and \$3.00, so we may worry about the quality of these approximations. In the special case where the consumers respond only to the leading digit of the price, demand is the same for all prices between $(\$X.01, \$X.99)$ and demand is a step function. Inverse demand, is not a function at all. Because the quantity demanded does not vary on the interval $(\$X.01, \$X.99)$, the "triangle" from the consumer surplus disappears: $\frac{1}{2}\Delta P \Delta Q$, and instead we have a series of rectangles of width \$1.00.[19] Thus the two formulas become:

$$\begin{aligned} I(\Delta\tau) &= \frac{\Delta P \cdot Q_1}{(P_0 - MC)\cdot \Delta Q + (\Delta P - \Delta\tau)\cdot Q_1} \\ SC(\Delta\tau) &= \frac{(P_0 - MC)\cdot \Delta Q}{\Delta\tau \cdot Q_1} \end{aligned}$$

The main difference is that now these formulas are no longer a linear approximation, but rather they hold exactly. Because $(P_0 - MC)$ is usually much larger than ΔP, omitting the consumer triangle does not substantially effect our estimates of $SC(\Delta\tau)$. Likewise because ΔQ is smaller than Q_1, omitting the the triangle does not substantially change ΔCS or pass-through (though it leads to a smaller pass-through estimate for all $\Delta\tau$). In our empirical estimates we use the (linear) expressions from (5) and (6), but we obtain qualitatively similar results with the step-function

[18] If taxes are smoothly and continuously passed-through into prices, then envelope theorem results such as Harberger (1964) or Chetty (2009) hold and $\frac{dQ}{d\tau}$ at the aggregate level is a sufficient statistic for welfare. Again, the challenge is that $\frac{dQ}{d\tau}(\Delta\tau)$ is likely to be a poor estimate of $\frac{dQ}{d\tau}(\Delta\tau')$ when price changes are discrete.

[19] We have to be a bit careful because our Econ 101 intuition suggests that we perform Riemann Integration on the inverse demand function, which is not defined. The corresponding Lebesgue Integral of the actual demand curve $Q(P)$ is trivial (it is just the area of a rectangle for a \$1.00 increase!.)

based demand function. Again, our focus is on the shape of $I(\cdot)$ and $SC(\cdot)$ with respect to $\Delta\tau$, not the necessarily the levels.

Our third caveat is that our expression for producer surplus represents only the per-period flow profits to producers, and does not include any menu-cost or price adjustment cost. That is not to say that menu costs or other rigidities do not play a role; they play a crucial role in determining $\Delta P(\Delta\tau)$, just that any additional cost incurred in order to change prices is not included in our measure or producer welfare.

A final caveat —important for the taxation of distilled spirits —is that we are ignoring the externality associated with alcohol consumption. This is not the same as assuming that there is no externality associated with alcohol consumption. Implicitly we assume that the pre-existing excise tax is at the Pigouvian level and that further taxes are purely about raising revenue as efficiently as possible, rather than correcting behavior. There is good evidence that this was the motivation in both Connecticut and Illinois, which raised taxes in light of budget shortfalls rather than evidence that the social cost of alcohol consumption had increased. We could not find any evidence in the literature that the social cost of alcohol consumption has increased over the past decade.

3.3 Menu Costs and Dynamic Structural Models

In this section we relate our simple statistical ordered logit model to the estimation of a more complicated dynamic structural model of price setting firms. We begin by considering the (static) model of price adjustment in Goldberg and Hellerstein (2013) where retailers maximize:

$$\Pi_t^r = \sum_j (p_{jt}^r - p_{jt}^w - mc_{jt}^r) \cdot q_{jt}^r(p_{jt}^r) - A_{jt}^r \cdot I[p_{jt}^r \neq p_{j,t-1}^r]$$

Retailer r sells product j in period t at price p_{jt}^r, pays a wholesale price p_{jt}^w and has a marginal cost of selling mc_{jt}^r (sometimes referred to as a "non-tradeable cost" in the exchange rate literature). The main addition is the adjustment cost A_{jt}^r that the retailer pays when the changing the price.[20] We can consider the retailer's dynamic problem as choosing a sequence of price vectors:

$$\max_{\mathbf{p_t^r}} \sum_{t=0}^{\infty} \beta^t \Pi_t^r(\mathbf{p_t^r}, \mathbf{p_t^{-r}}, \mathbf{p_t^w}, \mathbf{mc_t^r})$$

which we can write in recursive form:

$$V^r(\mathbf{p_t^r}, \mathbf{p_t^{-r}}, \mathbf{p_t^w}, \mathbf{mc_t^r}) = \max_{\mathbf{p_t^r}} \quad \Pi_t^r(\mathbf{p_t^r}, \mathbf{p_t^{-r}}, \mathbf{p_t^w}, \mathbf{mc^r}) + \beta E_{t+1}[V^r(\mathbf{p_{t+1}^r}, \mathbf{p_{t+1}^{-r}}, \mathbf{p_{t+1}^w}, \mathbf{mc_{t+1}^r})]$$

[20] Unlike here where our objective to to understand how taxes are passed through to prices, the goal of Goldberg and Hellerstein (2013) was to use revealed preference arguments to bound A_{jt}^r and then to recover the non-traded mc_{jt}^r.

This dynamic setup more closely resembles the setup of Nakamura and Zerom (2010).[21]

The general solution to problems of this type, is that for each period there is an optimal price, $p^*(\Omega_t)$ that the firm would set absent any frictions. This optimal price can be written as a function of the state variables $\Omega_t = \{\mathbf{p_t^{-r}}, \mathbf{p_t^w}, \mathbf{mc_t^r}, X_t\}$. As the marginal costs or demand shocks accumulate over time, the gap between p_t^r and p^* grows until the firm pays the cost A_{jt}^r and updates its price. In other words, prices follow an (s, S) rule.

The main distinction between our setup and the setup in Nakamura and Zerom (2010) or Goldberg and Hellerstein (2013) is rather than choosing a continuous p^* our firms chose among a discrete grid of points so that $\Delta p_{jt} = p_{jt} - p_{j,t-1} \in \{\ldots, -2, -1, 0, 1, 2, \ldots\}$.[22] Unlike static Bertrand Nash competition with continuous prices, when prices themselves are discrete, it is likely that there are large number of potential equilibria (and computing a single equilibrium is computationally demanding). The second difference is that we have over 30 retailers and over 500 products. It would be computationally infeasible to keep track of all prices of all competitors $\mathbf{p_t^{-r}}$, or even all products at the same firm $\mathbf{p_t^r}$.

Instead consider the two-step approach of Bajari, Benkard, and Levin (2007) where the goal is to recover the parameters of the payoff function. Implicitly they assume that an MPE exists, and only one equilibrium is played (even though there could be multiple equilibria). The first stage of the approach is to estimate the policy functions of the agents: $\hat{\sigma}_r(\Delta p_{jt}, \Omega_t)$, and the transition densities of the exogenous variables: $\hat{f}(\Omega_{t+1}|\Omega_t)$. The second stage considers deviations from the policy functions to recover the parameters of the payoff function $\Pi^r(\theta)$.

Our empirical approach is to model the policy function $\hat{\sigma}_r(\Delta p_{jt}, \Omega_t)$ as an ordered logit over possible discrete price changes. The preferred policy function would be a non- or semi-parametric estimate of the $Pr(\Delta p_{jt}^r|\Omega_t)$, which is exactly what our flexible ordered logit estimates. The policy function may contain a number of (potentially endogenous) equilibrium objects such as interactions among firms, as well as implicitly allowing for dynamic objects such as menu-costs or other nominal rigidities.

We do not attempt to recover the structural parameters of $\Pi^r(\theta)$ including any potential menu-costs. In order to separately identify rigidities that come from menu-costs and rigidities that come from whole-dollar pricing increments we would need to impose additional structure on the profit function $\Pi(\theta)$, as well as the transition densities of the exogenous variables $\hat{f}(\Omega_{t+1}|\Omega_t)$.[23]

[21] Again, in that paper the objective was to recover the adjustment costs A_{jt}^r, and to understand how the frequency and size of price adjustment responded to changes in the underling stochastic process of cost shocks. They employed the Pakes and McGuire (1994) algorithm to compute the Markov-Perfect Equilibrium (MPE) of the above game.

[22] In contemporaneous work, Ellison, Snyder, and Zhang (2015) employ an approach in the spirit of Bajari, Benkard, and Levin (2007) to estimate a model of discrete price changes with menu costs in order to quantify managerial inattention.

[23] There are several quantities which could help separately identify the magnitudes of menu-costs from whole-dollar pricing rigidity. We could measure the frequency of price adjustment for high and low priced goods, as well as for high and low revenue goods; with the idea being that menu costs make firms more likely to adjust prices on high revenue items rather than merely high-priced items. We have estimated menu-costs using calibrated toy versions of our setting. In those simple simulations we find that implied menu costs are often 6-8x larger when one does not

The key is that for the counterfactuals we are interested in (e.g: incidence of taxation at alternative levels of $\Delta\tau$) we only need a predictive estimate of how prices respond to alternative taxes, we do not need to separate structural parameters within $\Pi(\theta)$. Therefore, we do not need to perturb agents beliefs about the future transition of the exogenous state variables $Pr(\Omega_{t+1}|\Omega_t)$, nor the contents of the per-period profit function $\Pi^r(\theta)$. Instead we only need to consider the policy functions of firms at different state variables.[24] The point is that we do not need to solve the full dynamic game to understand how incidence varies with tax changes. Instead we need only the estimated policy functions evaluated at a different state variable $\hat{\sigma}_r(\Delta p_{jt}, \Omega'_t)$.[25] *Moreover, solving the full dynamic game would not provide any additional information beyond the estimated policy functions.*

This approach has some caveats. Namely, we assume that a larger tax increase would not effect the stationary Markovian strategies of firms, nor would it affect future transition rules $f(\Omega_{t+1}|\Omega_t)$. While the first might be reasonable, the second is more likely to be problematic. In the long-run, if manufacturers bear some of the incidence we might expect the rate of future upstream price increases to decline in response to a larger tax.[26] This is problematic even in the structural dynamic approach as the problem would become non-stationary.

4 Data Description

Our primary data source is the Kilts Center Nielsen Homescan Scanner dataset. The Nielsen data are a substantial improvement over previous price data in the alcohol tax literature; much of the prior literature relies on the ACCRA Cost of Living Index data, which survey a small number of products and stores in each state. Nielsen provides weekly scanner data, which track prices and sales at the UPC (universal product code) level for a (non-random) sample of stores in all fifty states, though in practice we only have sufficient data from 34 states.[27] These weekly data are available from 2006-2013, and include data from both standalone liquor stores as well as from supermarkets and convenience stores. Participation in the Nielsen dataset is voluntary, and not all stores participate. Supermarkets are much more likely to be included in the Homescan dataset than stand alone liquor stores, and larger chain stores are more likely to participate than smaller mom-and-pop stores. This leads to there being better coverage for states where spirits are sold

account for whole-dollar pricing increments. Without more data on different sized tax increases, we do not believe we can accurately estimate menu-costs without putting substantial additional structure on the problem.

[24] There is a growing literature which formally explores identification of counterfactuals dynamic discrete choice models when the full model is under-identified Kalouptsidi, Scott, and Souza-Rodrigues (2015).

[25] The analog in the single agent case would be if we had estimated the bus replacement model of Rust (1987) but were only interested in how the replacement probabilities would look if all buses had 10,000 extra miles.

[26] The upside is that short and medium-run measures of incidence are more likely to overstate the relative burden on consumers vis a vis firms which is in the opposite direction of the one we are most worried about.

[27] We lack sufficient data from 16 states, many of which are control states (in bold): **Alabama**, Alaska, Hawaii, **Idaho**, Kansas, **Montana, New Hampshire**, New Jersey, **North Carolina**, Oklahoma, Oregon, **Pennsylvania**, Rhode Island, Tennessee, **Utah, Vermont, Virginia**.

in supermarkets. We examine the effect of the July 2011 tax-increase in the state of Connecticut, where we observe 34 (mostly larger) stand-alone liquor stores. Because spirits are also available in supermarkets in Illinois, we have data from 713 stores there. We lack sufficient data to study the 2013 excise tax increase (and sales tax decrease) in Rhode Island or the 2009 excise tax increase in New Jersey.

Our second dataset tracks the wholesale prices for the state of Connecticut only. Wholesale prices are a key predictor of retail price changes. These prices were scraped by us from the Connecticut Department of Consumer Protection (DCP) from August 2007 to August 2013. These data are available because Connecticut requires that all licensed wholesalers post prices. Wholesalers agree to charge retailers these prices for the entire month, and are legally not allowed to provide quantity discounts or price discriminate.[28] Only 18 wholesale firms have ever sold brands of distilled spirits that we observe in the Nielsen dataset, and more than 80% of sales come from just six major wholesalers. Because Illinois does not require that wholesalers publicly post prices, we do not have wholesale pricing information there.

A third piece of data are the *estimated wholesaler marginal costs* from our previous paper. We use these estimated marginal costs only to compute producer surplus, and never as explanatory variables in our regression estimates. In Conlon and Rao (2015), we estimate the prices paid by wholesale firms to importers and distillers at the product level. Those estimated marginal costs are the output of structural model of demand using the approach of Berry, Levinsohn, and Pakes (2004). In Conlon and Rao (2015), we find that the average wholesale markup in Connecticut is approximately 30-40% for most products, which is larger than markups in other states, especially for more expensive products. We have wholesale marginal cost estimates for over 80% of the Nielsen data by quantity.

There were two important considerations in the construction of our dataset. The first was is products across competing wholesalers, or from wholesalers to Nielsen UPC's. Here we consolidate products so that a product is defined as brand-flavor-size such as *Smirnoff Orange Vodka 750mL*.[29] Sometimes a "product" may aggregate over several UPC's, as changes in packaging can result in a new UPC.[30] Second, a "product" may have one UPC for 2007 and 2008, but a different UPC in 2009 and 2010 if the packaging was redesigned. The third most common occurrence is that the same product may be available in both glass and plastic bottles at the same time. We rarely observe

[28] Connecticut is one of 12 states with a set of regulations known as *Post and Hold*, which mandates that all wholesalers post the prices they plan to charge retailers for the following month. Wholesalers must commit to charging these prices for the entire month (after a look-back period when wholesalers can view one another's initially posted prices and adjust their prices downwards without beating the lowest price for the product). For a detailed analysis of these regulations please see (Conlon and Rao, 2015).

[29] We emphasize the 750mL bottles which comprise 58% of products and 37% of sales, and the 1750mL products which comprise one third of products, and 55% of sales; we also include 1L bottles in our study though they are substantially less popular (about 8% of products and sales). Other sizes such as 375mL "pint" and 50mL "nip/airplane" are excluded from our analysis.

[30] UPC changes most commonly arise with special promotional packaging such as a commemorative bottle, or a holiday gift set. At other times, the change in UPC may be purely temporal in nature.

price differences for glass and plastic packaging within a product-month, so we also consolidate these UPCs. In total, these consolidations help us to construct a more balanced panel of products over time, and avoid gaps during holiday periods, or products going missing when packaging changes. This is especially important when our goal is to capture *changes* in prices within a product-store over time.

The second consideration relates to the time component of our dataset. Connecticut's tax changes took effect July 1, 2011 and we observe (fixed) wholesale prices for each calendar month. Another feature of Connecticut law is that temporary retail sales are not allowed, and retail "sales" must be registered with the DCP in advance. This provides a substantial advantage over other states or product categories in that weekly prices are more likely to accurately describe the prices consumers face. The Nielsen scanner data are recorded weekly, and some weeks span two months. Because wholesale prices vary at the monthly level, we aggregate our retailer data to the store-month. We allocate weeks to months based on the month of the end date but drop the first week of the month since we cannot completely assign weeks that span two months to one month or the next. We aggregate to monthly store data by taking the sales weighted median price for each product-store; in practice there is only a single price for 99% of store-month-product observations.

We construct price changes at both the retail and the wholesale level over different time horizons. For example, we compute 1 month, 2 month, 3 month, 6 month, and 12 month price changes. This lets us measure potential pass-through effects over different time horizons, and allows for the fact that pass-through may not happen instantaneously.

5 Empirical Results

We first provide descriptive evidence of the immediate and large price response to Connecticut's July 2011 tax increase, and then document the pricing behaviors that underlie this response. To match the prior literature we next estimate pass-through rates using linear regression and find evidence of over pass-through, though estimates vary over different horizons. In light of the pricing behaviors we document and the rigidities they imply we then estimate discrete choice models to better approximate the pricing patterns we see in the data. Finally, we use these estimates to predict how prices will change in response to tax increases of different magnitudes and simulate the resulting incidence and social cost of tax revenue.

5.1 Descriptive Evidence

Before running any regressions, we summarize observed price changes in Connecticut by month and year. We highlight July 2011, when state alcohol taxes on spirits increased by $0.178 for 750mL products, $0.238 for 1L products and $0.416 for 1750mL products. Table 2 reports average sales-weighted monthly retail and wholesale price changes and demonstrates three facts. First, there is a regular, seasonal component to price changes with prices increasing in some months like January and July and decreasing in others like February. Second, wholesale and retail prices immediately and sharply increased in reaction to the tax hike with a mean wholesale price change of $1.462 and a mean retail price change of $0.422, which were substantially larger than other July price increases, and larger than the tax increase. These seasonal patterns suggest that Connecticut raised taxes in a month when wholesalers and retailers were likely to raise prices anyway, potentially facilitating the immediacy of impact of the tax hike on prices. Finally, though the tax increase passed in May 2011, prices were not increased until the taxes took effect in July.

We can decompose the average price changes into two components: the magnitude of price adjustment, and the frequency of price adjustment. Figures 3 and 4 plot the sales-weighted fraction of wholesale and retail products experiencing price changes (positive or negative) and price increases (positive only) each month. As is evident from the two figures, there is a substantial spike in the frequency of price adjustment (and especially price increases) commensurate with the July 2011 tax increase. Wholesalers changed 90% of prices in July 2011, increasing 68.1% of prices. Retailers changed somewhat fewer prices, 43.8%, increasing 36.1%. These figures suggest most of the response to the tax is immediate, rather than taking place over many months.

We also provide evidence on the magnitude of price changes in Table 3. At the retail level, even during the period of the tax change, the majority of products (3,566 out of 5,605 or 63.6%) experience no price change, and in other months the fraction of unchanged prices is even higher (83.6%). Conditional on observing a price change, more than 75% of retail prices change in increments of $1.00. Compared to retail prices in other periods, we see that price changes are both more frequent, and more likely to be positive during the period of the tax change. Among products that do experience price changes during July 2011: 7.5% of products decrease their price by $1.00, 38.6% increase their price by $1.00, 12.4% increase their price by $2.00, and 3.5% increase their price by $3.00. Later, when we model the price change decision as an ordered logit, we restrict the set of discrete price changes to $\{-1, 0, +1, +2, +3\}$.

Our explanation for why retail prices change in $1.00 increments is provided in Table 4. We observe that almost 91% of retail prices in Connecticut end in ninety-nine cents. An additional 3.6% of products (mostly selling for less than $15) end in forty-nine cents. In Illinois, we see that 80.7% of products end in ninety-nine cents, and 4.7% end in forty-nine cents. However, we also observe a few chains where products end in ninety-seven cents (5.3%) and ninety-eight cents (1.7%). Additionally, nearly three-quarters of price changes are from ninety-nine to ninety-nine cent prices.

We document substantially less rigidity in the wholesale prices. In Table 3, we see that 43.8% of wholesale prices are adjusted in increments of $1.00. However, during the period of the tax change we also observe wholesale price adjustments of twenty-five cents (5.5%) and fifty cents (5.0%) in addition to $1.00 (11.3%). These make sense because the tax change on 750mL bottles is $0.178, while the tax change on 1.75L bottles is $0.416, and the price increments would represent slightly more than full pass-through of the tax. Meanwhile, Table 4 suggests that the modal wholesale price ending is ninety-one cents (50.4%).[31] In total, we see both less rigidity and more frequent adjustment of wholesale prices than of retail prices in Connecticut, as well as less "focus" around ninety-nine cents.

This focus on prices which end in ninety-nine cents is not unique to our setting. A literature which documents the presence of "convenient prices" as a source of nominal rigidities in macroeconomics includes: Kashyap (1995), Knotek (2010) or Levy, Lee, Chen, Kauffman, and Bergen (2011).[32] Our distilled spirits data support previous findings in macroeconomics which suggest that retail prices are less volatile than wholesale prices (at least after one ignores temporary sales). More broadly, tabulations of the Nielsen data demonstrate that prices are concentrated at a handful of price points in many product market. Approximately 40.1% of all prices are set at at one of the four most common price points for each product category. At the product category level, in more than 49% of the 1,113 product categories of the Nielsen data, at least half of all prices in each category are set at one of the four most common price points per category. The discrete pricing patterns that we observe are not unique to spirits; retailers price discretely in many product markets.

Though measured tax pass-through rate is large ($\geqslant 1$), Tables 3 and 4 indicate that a majority of retail prices are not changed at all. Instead, the measured retail pass through rate is generated by a small number of large, whole-dollar price increases. These pricing patterns show that retailers, and to a lesser degree wholesalers, do not smoothly pass on cost shocks but rather adjust prices in $1.00 increments.

5.2 Linear Model: Pass-Through of Alcohol Taxes into Spirits Prices

To start, we ignore price rigidities and employ a linear regression of first differenced prices on first differenced taxes in order to estimate pass-through rates described by (3). We provide two main sets of specifications, the first is a constant pass-thorugh rate $\rho_{jt}(\mathbf{X}, \Delta\tau) = \rho$ and the second is the "semi-parametric" (in $\Delta\tau$) estimate where we estimate a container size specific pass-through rate. An estimated pass-through rate of one implies full or 100% pass-through, while a rate exceeding one implies over-shifting of the tax burden and less than one indicates incomplete pass-through. An observation corresponds to a store-month-upc price. We allow for month of year fixed effects and year fixed effects, but cannot allow for month-year fixed effects and still identify the impact of

[31] The 91 cent wholesale price mirrors the 99 cent retail price due to an unusual feature of Connecticut law governing alcohol sales. Retailers are legally prevented from selling below cost, and cost is interpreted as the wholesale per unit price plus 8 cents. Additionally, wholesale prices are generally quoted in cases rather than a per-bottle equivalent pricing.

[32] A deeper question is why we observe ninety-nine cent prices. One potential explanation is that consumers exhibit "left digit bias" and do not fully process information. This idea is explored in Lacetera, Pope, and Sydnor (2012). Another explanation might be that firms consider only a smaller number of discrete price points for cost or information processing reasons, or may simply lack the technology to display prices that do not end in $0.99 (perhaps because the last two digits are permanently printed). While interesting, these "why" explanations are beyond the scope of our paper.

the tax change. The hope is that time fixed effects control for seasonality in price adjustment (such as the fact that prices are often adjusted in July). We estimate the differenced model over one-, three- and six-month differences to assess how pass-through rates vary over time. We also allow for product fixed effects α_j, which in the differenced model have the interpretation of a product-specific (common across stores) time trend.

Tables 5 and 6 report our estimates of the pass through of taxes into Connecticut retail prices and Illinois retail prices, respectively. Later, we repeat this exercise and report the wholesale pass-through for Connecticut only (because we do not observe wholesale prices in Illinois) in Table 8. The first row of each table reports average pass-through rates across all products while the subsequent rows report pass through rates separately for each bottle size. For all three tables the first three columns report average pass-through rates using all observations, while the last three columns limit the samples only to products where we have observed a change in price. Because they comprise only about 8% of the overall sample, the results for 1L bottles can be noisier than those for 750mL or 1.75L bottles.

As reported in Table 5, average retail prices reacted strongly to the tax increase with a one-month pass-through rate of 1.533 (0.271) that declines to 1.013 (0.264) at the six-month horizon, implying that taxes were first over-shifted and then fully passed through to retail prices, though the estimates are statistically indistinguishable. As the size-specific estimates show, smaller products experience higher pass-through rates with six-month pass-through rates of 2.084 (0.503), 1.586 (0.470) and 1.009 (0.263) for 750mL, 1L and 1750mL products, respectively. We report pass-through rates over differences up to 6 month but the pattern holds over longer horizons as well (though the standard errors increase). Over a 24-month horizon, for example, the estimated coefficients for 750mL,1L and 1750mL products are 1.826 (0.612), 0.915 (0.591) and 0.757 (0.396), respectively. In order to highlight the discrete nature of the price response, we estimate much larger pass-through rates (often as large as 200-300%) once we only consider products which experience a a price change in Columns (4) through (6). These higher conditional pass-through rates are also declining in product size, with a pass-through rate almost as high 400% for the 18 cent tax change on 750mL bottles, and a pass-through rate of approximately 200% for the 42 cent tax increase on 1.75L bottles. We cannot reject the hypothesis of the pass-through rate implied by $\Delta p_j = 1$ for any of our conditional specifications.

We also examine the impact of Illinois' September 2009 \$4.05 per gallon (\$1.07/L) tax increase on retail prices which we report in Table 6. The major difference is that this constitutes a much larger tax increase, one that is much closer to \$1 per bottle than the Connecticut increase. In contrast with Connecticut, the pass-through rate rises as we expand the time horizon from one month to six months. We attribute this to the fact that in Connecticut the tax increase was concomitant with the seasonal price adjustment, while in Illinois the tax increase was four months before the seasonal price adjustment in January.[33] After six months, we find evidence that pass-through estimates for 750mL and 1L products, which experienced tax changes of \$0.80 and \$1.07 respectively, were statistically greater than 100%, while we cannot reject complete pass-through (100%) for 1.75L bottles which experienced a \$1.87 tax change. We also observe that in Illinois the discrepancy between pass-through and pass-through conditional on a price change (at the six month horizon) is much smaller than in Connecticut. We attribute this to the larger size of the tax increase (which ranged from \$0.80 to \$1.87 instead of \$0.18 to \$0.41); the frequency of price adjustment is higher in response to the larger tax increase, and the denominator of the pass-through rate $\Delta\tau_{jt} \approx \$1$ for 750mL and 1L sizes.

We also pool data across Connecticut and Illinois. We report those results for just the 3 month window in Table 7. These results largely mimic those in Tables 5 and 6, and when we fully interact all of the variables with state indicators (except for the product fixed effects) as we do in Column (5), we get nearly identical results. These pooled regressions are helpful to demonstrate that the measured pass-through rate is in general declining in the size of the tax increase $\Delta\tau$. The only tax increase that appears to be under-shifted is the tax increase of Illinois of \$1.872 on 1.75L bottles (coefficient of 0.830), meanwhile the tax increase of \$1.07 on 1L bottles by Illinois appears to be fully passed-through (coefficient of 1.061). These are much smaller than the estimated pass-through rate of a \$0.17 tax increase on 750mL bottles in Connecticut (coefficient of 1.936). Here, we believe that the relationship between the size of the tax and the \$1 price increase is not coincidental. We attempt to recover the nonlinear structure of $\rho_{jt}(\mathbf{X}, \Delta\tau)$ by allowing $\Delta\tau_{jt}$ to enter the regression for Δp_{jt} as a polynomial function. Rather than report those estimates as a table, we display them in Figure 5. We plot the estimated pass-through rates from Column (5) with their standard errors, along with polynomial functions of $\Delta\tau_{jt}$. Note that the pass-through rate itself is a rate of change; in the context of Figure 5 where the x-axis marks tax changes, a linear pass-through rate would be a horizontal line. The dashed black line of Figure 5 represents a regression of Δp_{jt} on $\Delta\tau_{jt}, (\Delta\tau_{jt})^2$, while the dotted line also includes $(\Delta\tau_{jt})^3$.[34]

For Connecticut, we are also able to measure the pass-through rate of tax changes into wholesale prices. We report those estimates in Table 8. The patterns are roughly similar to those in Table 5. The main differences are that for 750mL bottles we measure higher pass-through rates (in excess of 300%) and somewhat lower but not statistically different estimates for 1.75L bottles. It is important to note that the frequency of price adjustment (as indicated in Figure 3) is much higher at the wholesale level than at the retail level, such that nearly all wholesale prices are adjusted after 6 months. We present estimates both for all wholesalers (between one and six firms selling the same product) or for the lowest priced wholesaler. We find these estimates are highly similar (in part because there is very little price dispersion at the wholesale level). Beyond the scope of our

[33] Our ability to delve deeper into this explanation is limited by the lack of wholesale pricing data for Illinois.

[34] These polynomial regressions tend to put substantially more weight on Illinois than on Connecticut because we have much more data and much smaller standard errors in Illinois.

data is whether wholesale price adjustments are driven by price adjustments among the upstream manufacturer/distillers.

Even within a product, there is substantial heterogeneity across stores in the response to the tax, which we report in Table 9. We find that stores that sell products at relatively lower prices are more likely to experience price changes in response to the tax change, than those stores with relatively higher prices. We use two different discrete measures of high and low-price stores by product and month: the first column constructs dummies for prices above or below the median price, the second column constructs dummies for the highest and lowest price retailer (allowing for ties) selling the same product. For both measures we find that at low-price stores the tax is passed on at a rate of roughly 290% to 300% while at high-price stores the tax is passed on a rate of only 26% to 31%. We also employ a continuous measure of relative price in the third column and again find lower relative prices are correlated with larger pass-through rates.[35]

These patterns suggest that when tax increases are concomitant with seasonal price changes as in Connecticut, the tax increase is quickly passed through to retail prices but that when a tax increase follows a seasonal price adjustment as in Illinois it may take several months for enough prices to be adjusted to reflect the tax increase. It also suggests that there is substantial cross-store heterogeneity within a particular product in the price adjustment process driven primarily by relative prices before the introduction of the tax.

5.3 Discrete Price Changes

In our main specification, we directly model the change in price of product j by retail firm f in month t as an ordinal logit:

$$\begin{aligned} \Delta p_{fjt} &= k \text{ if } Y^*_{fjt} \in [\alpha_k, \alpha_{k+1}] \\ Y^*_{fjt} &= f(\Delta\tau_{jt}, \theta_1) + g(p^w_{jt}, p_{fjt}, \theta_2) + h(p_{fjt}, p_{-f,jt}, \theta_3) + \beta X_{fjt} + \gamma_t + \varepsilon_{jft} \end{aligned} \tag{8}$$

The key is that we restrict Δp_{fjt} to a set of discrete values; in our empirical example these values are $\Delta p_{fjt} \in \{-1, 0, 1, 2, 3\}$. For observed price changes outside this range we assign them to the nearest discrete price change. The relationship between Δp_{fjt} and covariates is informed both by the dynamic optimization problem that firms solve in Section 3.3 and the regression results of Section 5.2. To be more precise, we allow the tax change to flexibly influence Δp_{fjt} through $f(\cdot)$. We also allow the cumulative change in the wholesale price since the last change in the retail price, Δp^w_{jt}, to flexibly enter equation (8) in $g(\cdot)$. More specifically, we allow for both a polynomial in the (cumulative) Δp^w_{jt} and an indicator for no change in the wholesale price since the last retail price adjustment. Once we control for $g(\cdot)$, we find that additional controls for duration between

[35] In addition, higher pass-through rates at stores that initially charge lower prices effectively compresses the distribution of prices for a given spirits product and thus may limit the ability of consumers to reduce pass-through rates by shopping.

price adjustments are not significant. Motivated by Table 9, we also include $h(p_{fjt}, p_{-f,jt}, \theta_3)$, which includes indicators for being the highest or lowest priced seller of a the good among the competition, as well as a polynomial in the difference between the firm's price and the median competitor price for the same good.

We include additional covariates such as the lagged price of the same product at the same retailer (to capture price changes at "cheap" vs. "expensive" products), the annual sales of that product at that retailer (to capture which products are important for overall profitability) the aggregate annual unit sales of the retailer (to capture "large" vs. "small" retailers), and the number of firms that also sell the product (which could capture "competition" or "popularity"). We include dummy variables for month of year, and year in order to capture seasonal patterns in price adjustments. We do not include product specific fixed effects, because we worry about the incidental parameters problem in the nonlinear model.[36] We omit other covariates, even though they have strong predictive content such as the fraction of retailers which experience price changes for the same product in the same period because of endogeneity concerns, and because we worry it will affect our estimates of $f(\cdot)$.[37]

We present the results of two ordinal logit regressions in Table 10. The first uses data from only the state of Connecticut. The advantage is that for Connecticut we have matched wholesale price data and are able to include the $g(\cdot)$ term. This turns out to be very important in capturing the heterogeneity in the response to the tax change. More specifically, it helps determine how close to the boundary of the "band of inaction" of the (s, S) rule a firm-product is. The results in Table 10 indicate that products with no wholesale price change since the last retail price change are very unlikely to experience a retail price change, while the index Y^*_{fjt} is increasing at an increasing rate in cumulative wholesale price changes (since the last retail price change). The disadvantage of using only the Connecticut data is that we observe tax changes only within the narrow range of \$0.17 – \$0.41 per bottle. For this reason, we restrict $f(\Delta\tau_{jt}, \theta_1)$ to be linear so that we do not extrapolate a polynomial outside the observed domain.

Thankfully, once we include the Illinois data, we observe tax changes between \$0.17 – \$1.87 per bottle. This lets us include a more flexible cubic form for $f(\cdot)$. It also means that we can analyze much larger tax changes and still be *interpolating* rather than *extrapolating*. The disadvantage is that we must omit $g(\cdot)$, the main source of heterogeneity in the "initial conditions". Most of the other terms in Table 10, have the expected sign and are in agreement whether we use data only from Connecticut or whether we include data from Illinois as well. For both specifications, $h(\cdot)$ has the expected interpretation that higher priced firms experience smaller price adjustments on average, just as in the linear regression from Table 9. Using only the Connecticut data, prices seem more likely to increase among: lower priced products, with lower sales, at larger retailers, when fewer competitors sell the same product. While with Illinois larger price increases are associated

[36]Omitting product fixed effects from the linear models in Section 5.2 tends to affect the R^2 of the regression, but not the estimated pass-through rate.

[37]We present some of these results in Appendix B.

with smaller stores selling more popular products.

In our counterfactuals, we vary the size of the tax change, and predict price changes (and welfare) for each store-product. There are two possible methods of prediction. The first is the *smooth expected price change* which takes $E[\widehat{\Delta p_{fjt}}] = \sum_{\forall k} k \cdot Pr(\Delta p_{fjt} = k | \Delta \tau_{jt})$. The second is the *best guess* price change. The *best guess* of an ordinal model is not the most likely outcome, but rather the outcome that corresponds to a draw $u = 0.5$ from the CDF of the ordered choice model. Formally the best guess is $\Delta p_{fjt}(\Delta \tau_{jt}) = a$ where $\sum_{k' < a-1} Pr(\Delta p_{fjt} = k', \Delta \tau_{jt}) < 0.5 \leqslant \sum_{k' < a} Pr(\Delta p_{fjt} = k', \Delta \tau_{jt})$. We use the *best guess* because it provides a single discrete prediction. In contrast, the *expected price change* almost always predicts non-integer price changes like \$0.84, which would not allow us to understand the welfare implications of discrete pricing. Because we are interested in the causal impact of excise taxes, we do not report the level of the predicted price change, but the *difference* between the predicted change of a tax increase of some positive $\Delta \tau$ and the predicted price change of no tax increase $\Delta \tau = 0$.

We present the (difference of) predicted price changes under both methods in Figure 6. We also compute the implied pass-through rates. These are $\rho(\Delta \tau) = E[\frac{\widehat{\Delta P_{fjt}}(\Delta \tau_{jt})}{\Delta \tau_{jt}}]$ evaluated at different levels of $\Delta \tau_{jt}$, where the expectation averages store-product specific pass-through rates by (post tax increase) quantity similar to the Paasche price index.[38] In all graphs we use vertical lines to denote the observed tax increase of \$0.237/L in Connecticut and \$1.07/L in Illinois.

The main difference in the best guess predictions between the predictions from the "Connecticut Only" model and the model which also includes data from Illinois is that the "Connecticut Only" specification predicts smaller and less concentrated price increases. In part this comes from the ability to incorporate heterogeneity in wholesale costs. Without this heterogeneity, the joint CT-IL model tends to predict highly correlated responses to tax changes such as the very large number of products which increase prices around a tax increase of 50 cents per liter. Under the best guess prediction we also report the % of observations that are correctly predicted as well as the prediction rate of the naive model that predicts the same outcome (no price change) for every observation. We report these classification rates for the period of the tax change (July 2011 in Connecticut); including data from Illinois does not improve in sample fit for Connecticut, but it does allow us to analyze larger potential tax changes.

Ideally, we would have wholesale data from Illinois and could estimate a monolithic model. Because our data are incomplete, we average the models from column (1) and (2) of Table 10. We take a 50-50 average of the two models and use this as our prediction of $\widehat{\Delta P_{fjt}}(\Delta \tau_{jt})$ under counterfactual tax rates. We also report the predictions of the averaged model in Figure 6. The hope is this allows us the to incorporate the heterogeneity in initial conditions from model (1) $g(\cdot)$,

[38] We choose to quantity weight using post-increase predicted quantities because it more closely correlates with Paasche price index. We have also constructed the same figures using the Laspeyres index and obtain imperceptibly different results. This is because the change in quantity weighting is small relative to the large and discrete nature of price changes. We have tried other weighting measures such as equal weighting across store products and obtain qualitatively similar results.

while incorporating the more flexible response to the tax of model (2) $f(\cdot)$. Averaging two consistent models should produce another consistent model. There are more sophisticated ways to select the relative contribution of each model, but they typically involve estimating each model many times on different samples.[39] All counterfactual predictions we report in the subsequent section are for the averaged model, though we provide separate predictions under model (1) and (2) for all of the results in the subsequent section in Online Appendix B. We also report results of a more restrictive binary choice model where firms can only increase by $1 or not change prices in Online Appendix A. Qualitatively, both appendices provide similar to results to our main specification.

5.4 Welfare Implications

To better understand the implications of discrete price adjustment for tax incidence and efficiency we estimate the quantities given by equations (5) and (6) over a range of potential tax increases. We use the best guess predicted price changes from the averaged ordered logit model (as reported in Figure 6).

To estimate the surplus losses to consumers and producers as well the deadweight loss and revenue raised by taxes of different magnitudes we draw on a combination of data, parameter estimates and assumptions. Our main input is the predicted price change at different tax levels $\widehat{\Delta P}(\Delta\tau)$ which we obtain from our ordered logit model in the previous section. We also rely on the observed store-product-month level price and quantity (P^0_{fjt}, Q^0_{fjt}) directly from the Nielsen data in the quarter prior to Connecticut's July 2011 tax increase (2011Q2).

In order to predict counterfactual quantities under different prices, we need an estimate of the demand elasticity ϵ_D from equation (7). Here we assume constant elasticity of demand, and consider a range of elasticities which span those obtained in meta-analysis by Wagenaar, Salois, and Komro (2009) $\{-1.5, -1.0, -0.5\}$. Our own estimates range between $[-0.5, -1.0]$ depending on the precise specification. We could obtain both counterfactual quantities and consumer surplus from a more complicated multi-product demand system such as Berry, Levinsohn, and Pakes (1995) or Deaton and Muellbauer (1980). Arguably, such a system might better capture substitution between products as prices of some brands increase while others do not. We design our welfare calculations to look as similar as possible to the traditional formulas employed in public finance, so as to highlight that discrete price points, rather than the multi-product demand system drive the results.[40]

[39] In econometrics consistency of averaged models (under misspecification) is established in Zhang (2015). Model averaging is quite popular in the machine learning literature, for a general overview please consult: Hastie, Tibshirani, and Friedman (2001).

[40] The assumption of zero cross-price elasticities will matter only to the degree that the welfare gain from switching products varies substantially with the size of the tax change. As a robustness test we compute ΔCS and ΔPS using the structural demand model from our other paper (Conlon and Rao, 2015). In general we find less incidence on consumers than in the constant elasticity framework, but the way incidence (and efficiency) vary with respect to $\Delta\tau$ remains qualitatively similar.

We also need an estimate for the product-level marginal cost MC_{jt} in equations (5) and (6). We use an estimate of the marginal cost for the *wholesaler* recovered from the structural model in our other paper (Conlon and Rao, 2015). This has some advantages and disadvantages. It is preferable to use the wholesaler's marginal cost rather than the retailer's marginal cost (wholesale price) because wholesaler surplus is substantially larger than retailer surplus and wholesalers operate within the same state. A disadvantage is that we do not include distillers or manufacturers in our surplus calculations.[41] One rationale for doing so is that distillers are often multinationals and are out-of-state businesses. The more pressing concern is that we have little to no information on the production function for distilled spirits.[42] As a robustness check, we obtain qualitatively similar results using fixed markup rules between 20-50% which we believe are in line with the markups in the industry.

It is important to point out that our measure of producer surplus: $\Delta PS = (P_0 - MC) \cdot \Delta Q + (\Delta P - \Delta\tau) \cdot Q_1)$ is the *flow profit* measure to the firm, and does not include any menu costs or other adjustment costs that might be incurred when changing prices. We do not explicitly separate out menu-costs, however, if menu-costs are present they are implicitly included in our estimates of the policy function $\Delta P(\delta\tau)$. Because such an adjustment cost would be amortized over several periods, evaluating welfare for both consumers and producers would involve more complicated dynamic expressions for both. Again, we stick to the familiar area of rectangles and triangles to more closely match the traditional public finance approach and highlight that discrete price adjustment (rather than a fully dynamic model) drives the main result. We also re-iterate that our welfare measures assume that pre-existing excise taxes already correct for the externality, so there are no social benefits in further reduction of distilled spirits consumption.

Figures 7 and 8 presents an array of plots. Each row of figures uses a different assumed elasticity $\{-0.5, -1, -1.5\}$. The columns break out the results by bottle size. We report separate results for 750mL bottles, 1.75L bottles and all three bottle sizes together (which also includes the less popular 1L bottles). Again, the vertical dotted lines denote the observed tax increases in Connecticut and Illinois that identify the ordered logit parameter estimates. Like before, the x-axis denotes counterfactual tax changes in dollars per liter; however for the 750mL and 1.75L plots we re-label the axis in dollars per bottle.

For each of the nine plots, the red line measures the average price change $E[\widehat{\Delta P}_{jt}(\Delta\tau)]$ and the blue line measures the implied pass-through rate: $E[\Delta p_{jt}/\Delta\tau_{jt}]$. It is important to note that these vary with product size, but not with the elasticity because neither involves the quantity response to the price change. The green line denotes the implied incidence $\Delta CS/\Delta PS$ for various levels of the tax, which does vary with the elasticity. As expected, when demand is unit-elastic and

[41]Ideally, we would observe the prices paid by wholesalers to distiller/importer/manufacturers. Unfortunately, these contracts are proprietary information and direct data are not available to us.

[42]To the best of our knowledge the most important inputs in distilled spirits production are: advertising, starch, and water.

marginal costs are constant the pass-through rate and the incidence track one another very closely. As demand becomes more elastic, the incidence shifts away from consumers and towards firms, and when demand becomes less elastic the incidence is shifted more onto consumers. Again, this follows the traditional public finance intuition. When the share of products with whole-dollar price increases rises sharply, so does the pass-through rate. When the share of products with whole-dollar price increases grows more slowly, the pass-through rate (and the incidence) fall.

Where our results deviate from the usual intuition is that both the pass-through rate and the incidence vary quite a bit with the tax rate and follow a (series of) U-shaped pattern. For 750mL bottles, there is very little predicted price response to tax changes of less than 20 cents. This implies that firms bear most of the incidence of the tax. At around 25 cents, we observe a large number of products increasing their price so that the average price, the pass-through rate, and the incidence all rise quickly. For tax increases of more than 25 cents, consumers bear between 75% and 80% of the tax incidence. For the 1.75L bottles, there are sharp increases in the average price (and pass-through rate) around 25 cents per bottle, and $1.00 per bottle. This creates a substantial U-shape in both the pass-through rate, and the incidence of the tax. As the taxes increase beyond $1.00 per bottle for the 1.75L bottles, most products have already adjusted prices, thus the taxes rise more quickly than prices, and incidence shifts back towards firms. When we aggregate across all three sizes (including the 1L bottles has little effect as they are only 8% of sales), again we see the same series of U-shaped patterns, though the location of the peak and the trough shift somewhat when we aggregate across sizes. In total, the level of incidence varies with the elasticity. When demand is elastic (-1.5), consumers bear between 50%-65% of the tax; when demand is unit elastic consumers bear between 60% and 100% depending on the size of the tax; and when demand is inelastic (-0.5) the incidence varies between 100% and 200% depending on the size of the tax. We can see directly on the Figures how the implied incidence can exceed 100%: as many products increase prices by $1.00, the producer surplus is flat with respect to the tax as they trade higher prices off against fewer units, while the consumer surplus becomes extremely steep as they pay substantially higher prices. Discreteness allows us to obtain $\Delta CS/\Delta PS \geqslant 1$ without imposing convexity restrictions on the demand curve.[43]

For the purpose of comparison, Figure 8 presents results from the linear model with constant pass-through using the size specific estimates from column (2) of Table 5. Those values are $\rho = (1.90, 1.833, 1.154)$ for each of the three sizes respectively.[44]We construct our estimates using the same welfare formulas as we use for 7. Under the linear model, the average price (red line) is linearly increasing for all values of the tax while the implied pass-through rate (blue line) is perfectly horizontal. The incidence (green line) is nearly perfectly horizontal and for unit elastic demand more or less follows the pass-through rate. Both the implied pass-through rate and the incidence are

[43]Recall the linear model often obtained estimates of the pass-through rate $\hat{\rho} \geqslant 1.5$. Under monopoly, this would imply incidence of 150% or more.

[44]We get nearly identical estimates using the interacted parameters from column (5) of Table 7.

strictly higher in all cases than in the nonlinear model from Figure 7. Our nonlinear model seldom implies a pass-through rate in excess of 100% while the linear model always implies excessive pass-through. As anticipated in our theoretical discussion, many of these quantities (such as incidence around 200%) are difficult to interpret. Particularly problematic is that the incidence for small tax changes in 750mL bottles with inelastic demand (-0.5) actually has an asymptote because lost producer surplus is vanishingly small relative to lost consumer surplus.

Our main takeaway from Figure 7, is that discrete price changes imply that pass-through (and incidence) can vary substantially with the size of the tax, and often take on a series of U-shapes as different levels of the tax trigger a cluster of whole-dollar price increases. The standard linear regression approach would estimate a single pass-through rate from one tax increase and use that to understand the price response to tax increases of different sizes. Once we account for the discreteness, we see that approach is likely to be problematic.

The discreteness of price adjustment also has important implications for tax efficiency, which we measure as excess burden per dollar of tax revenue raised. Under the conventional approach (with smooth adjustment), deadweight loss increases quadratically in the size of the tax, while revenue increases linearly. This implies that the ratio is linearly increasing in the size of the tax. We plot the ratio of deadweight loss per dollar of tax revenue in Figure 9 which accounts for the discreteness of price changes. Two patterns emerge. The first is that the elasticity shifts the level of deadweight loss per unit of tax revenue, with more elastic demand leading to larger excess burden (as we would expect). The second is that the relationship between the efficiency of the tax and the level of the tax is not linearly increasing, but is non-linear and non-monotonic. By magnifying the quantity response to the same price increase, the elasticity acts merely to stretch (or shrink) the deadweight loss per dollar of tax revenue, it does not change the location of the local minima or maxima. When demand is unit elastic, a tax increase of 60 cents per liter generates twice as much deadweight loss per dollar of revenue raised as a tax increase of 40 cents per liter. However a tax increase of $1.00 per liter generates 20% less deadweight loss per dollar of revenue raised than the 60 cent per liter tax increase.

With discrete price changes there are intervals where the ratio of excess burden per dollar of tax revenue actually declines —revenue increases outpace surplus losses. Specifically, following a threshold where many prices are adjusted, even as taxes rise, many fewer prices are increased leading to small average quantity responses but additional revenues. This suggests a much closer relationship between the incidence of taxes and the efficiency of taxes when both vary with the size of the tax change. Taxes are most efficient when the incidence of taxes is lowest —when they are borne mostly by firms rather than consumers; we can see this by looking at Figures 7 and 9 together. In short, taxes that do not trigger price increases: (1) are paid by firms, (2) cannot generate deadweight loss because without price changes quantity remains unchanged.

Again, for the purpose of comparison we present parallel results from the OLS estimates in

Figure 10. As predicted by theory, the deadweight loss per dollar of tax revenue raised is linearly increasing in the size of the tax. Again, because the implied pass-through in the linear model exceeds 100% all of the time, the cost of raising funds is substantially larger in all cases. For unit elastic demand, the implied cost of raising a dollar of tax revenue in the linear model (for even a small tax) is strictly above that of the nonlinear model *for any value of the tax*.

The fact that the efficiency cost of tax revenues declines over some ranges suggests that states deciding on tax changes would do better on the efficiency front if they consider how prices are changed when they set tax increases. Larger tax increases in some cases may actually entail a lower cost of public funds on average. While finding the local minimum of a curve like 9 may be difficult, avoiding the local maximum may be somewhat easier. The potential savings are potentially large. Given unit elastic demand, a 60 cent per liter tax generates an average deadweight loss per dollar of tax revenue raised of over 1.0 units, while a 35 cent tax would generate only 0.6. This means it is possible to reduce the social cost of taxation by as much as 40%.

The challenge for policymakers is estimating how far from the (s, S) boundary firms are. We think this exercise is easier than it looks, at least for a single bottle size. In our case, the tax increase to avoid is the 60 cent per liter tax, which translates to an increase of \$1.05 for the most popular 1.75L bottle size. We don't think it is coincidental that the social cost of taxation is highest when the tax change and the price increment are similar. We see that policymakers in both Connecticut and Illinois pursued very different strategies for tax efficiency, but both were successful in avoiding the local maximum. In Connecticut, policymakers implemented a relatively small tax of 24 cents per liter that triggered a relatively small number of \$1.00 price increases. In Illinois, policymakers implemented a very large tax of \$1.07 per liter that triggered a large number of \$1.00 price increases, but collected a large amount of revenue and triggered relatively few \$2.00 price increases. Where the exercise becomes more complicated, is that tax increases that are good for one package size may be bad for others; but by focusing attention on the most popular bottle sizes, it may be possible to mitigate this problem.

6 Short and Long Run Implications.

One potential criticism is that in considering a more dynamic discrete price adjustment process by firms, we have explicitly created a distinction between short-run and long-run excise tax pass-through rates and incidence that is not present in constant pass-through linear models. There is a question as to whether or not it is worthwhile for policy makers to try and exploit the short-run dynamics we document to reduce the cost of raising public funds through excise taxes. That is, there is a question as to whether these pricing frictions are meaningful in the "long-run".

There are two different ways to think about what "long-run" means here. In one sense, "long-run" might mean how the incidence of an extremely large tax (or at least one much larger than the rigidity) would be allocated.[45] While this is an interesting exercise, we don't think this is the relevant quantity for our policymaker's decision.

The second way to think about the "long-run" is over a longer time horizon. Here we have some options. The first and most basic option would be to estimate our ordered logit model of price increases over a longer time horizon. We observe qualitatively similar results when we repeat our exercise over 6-month and 1-year horizons. In a sense, this is evident from the results from the linear model reported in Table 5, where the pass-through rate is relatively stable over different time horizons. In principle (and given enough data), we could estimate our model an even longer horizon, though it would likely require a longer panel than we observe.

A more challenging approach would be to explicitly consider a world where retailers dynamically adjust prices when facing a stream of demand and cost shocks. Ultimately all of the price increases that occur because of the tax might have occurred eventually as costs or demand rose. In order to understand the causal impact of taxes in this world, we would likely need a fully specified macro-style model of dynamic price adjustment. This would require a large number of additional assumptions, but could provide additional insights into mechanisms and precise magnitudes, and is best left for future research.

Our perspective is that the short run may not be such a short period of time. As indicated in Figure 4, in 2014 (well after the tax increase) retail prices increase less than once per year on average, and often with a predictable seasonal pattern. It is not unreasonable to think that a well timed tax increase could have a relevant horizon of 2-3 years. When paired with the potential to reduce the social cost of excise taxation by up to 40%, this seems relevant.

7 Conclusion

Empirically, we document an important rigidity in the pricing of distilled spirits that affects how taxes are pass-though to prices. Specifically, we demonstrate that retailers set the vast majority of prices of spirits products at only a handful of price points in terms of the cents portion of the price. In Connecticut 94.2% are set at just two price points and in Illinois 90.7% are set at just three price points. Discrete prices mean that firms follow a type of (s, S)-rule when deciding whether or not to change prices; they will withstand cost shocks, including tax increases, until they are sufficiently far from their optimal price that moving to the next price point leaves them better off, resulting in infrequent but large prices changes in set increments. Consistent with prior studies, we find that prices increase by more than the amount of the tax on average. However, we also show that 60% of products experienced no price change at all, and 75% of products which did experience a price change changed their price in whole dollar increments

Theoretically, we show these pricing rigidities can rationalize both incomplete or excessive pass-

[45]This exercise is present in Fabinger and Weyl (2013) as the "tax that kills the market". Likewise, Goldberg and Hellerstein (2013) looks at how pass-through would differ in a frictionless world.

through without placing restrictions on the demand curve. Because of nominal rigidities in pricing, the pass-through rate is not a constant, but is rather a (series of) Dirac Delta function(s) in the magnitude of the tax increase. Because of this structure, using parameter estimates from a linear regression of change in price on change in tax can be misleading when evaluating alternative tax policies. We also show that a correctly specified ordered logit model can account for the discreteness of prices and be used to recover the pass-through rate, incidence, and efficiency of alternative tax policies. Such a model could be interpreted as a "policy-function" of a more complicated dynamic game or decision problem, or as a purely statistical reduced form.

Quantitatively, we demonstrate that when price changes are discrete pass-through, incidence, and tax efficiency are non-linear and non-monotonic functions of the tax and that nominal rigidities (such as pricing increments or menu costs) have potentially important implications for tax policy. We document that price increases in response to tax changes can be concentrated around certain levels of tax increases. Further increases beyond these levels may not generate further price increases, but instead come out of firm profits. By allowing both the incidence and efficiency to vary with the tax rate, this suggests a more tightly linked relationship between incidence and efficiency than suggested by the previously literature which generally treats the two separately. Taxes are most efficient when the consumer incidence is minimized, sometimes larger taxes can produce a lower average cost of public funds. This strongly contrasts with the conventional wisdom that tax efficiency is linearly decreasing in the amount of tax revenue raised.

Our simulations raise the possibility of better policy design. By considering the pricing patterns and the optimization frictions they create explicitly, policymakers can improve the efficiency or minimize the incidence on consumers of tax increases. While setting a tax increase to coincide with one of the most efficient tax increases (i.e. the bottom a U) may be difficult for a taxing authority to achieve, avoiding tax increases that trigger the greatest deadweight loss per dollar of revenue raised may be an achievable goal in markets that feature discrete price adjustment, by avoiding raising taxes in the same increments. The U-shaped pattern of pass-through rates suggests that excise tax increases close in size to the increment by which prices are typically adjusted will lead to higher pass-through rates, more incidence on consumers, and more deadweight loss per dollar of tax revenue.

References

Anderson, S. P., A. De Palma, and B. Kreider (2001): "Tax incidence in differentiated product oligopoly," *Journal of Public Economics*, 81(2), 173–192.

Asker, J. (2005): "Diagnosing Foreclosure due to Exclusive Dealing," Working Paper.

Bajari, P., C. L. Benkard, and J. Levin (2007): "Estimating Dynamic Models of Imperfect Competition," *Econometrica*, 75(5), 1331–1370.

Berry, S., J. Levinsohn, and A. Pakes (1995): "Automobile Prices in Market Equilibrium," 63(4), 841–890.

——— (2004): "Automobile Prices in Market Equilibrium," *Journal of Political Economy*, 112(1), 68–105.

Besanko, D., J.-P. Dubé, and S. Gupta (2005): "Own-brand and cross-brand retail pass-through," *Marketing Science*, 24(1), 123–137.

Besley, T. (1989): "Commodity taxation and imperfect competition," *Journal of Public Economics*, 40(3), 359 – 367.

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

Chetty, R. (2009): "Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods," *Annual Review of Economics*, 1(1), 451–488.

——— (2012): "Bounds on elasticities with optimization frictions: A synthesis of micro and macro evidence on labor supply," *Econometrica*, 80(3), 969–1018.

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

Conlon, C. T., and N. S. Rao (2015): "The Price of Liquor is Too Damn High: State Facilitated Collusion and the Implications for Taxes," Working Paper.

Cook, P. J. (1981): "The Effect of Liquor Taxes on Drinking, Cirrhosis, and Auto Fatalities," in *Alcohol and public policy: Beyond the shadow of prohibition*, ed. by M. Moore, and D. Gerstein, pp. 255–85. National Academies of Science, Washington, DC.

Deaton, A., and J. Muellbauer (1980): "An Almost Idea Demand System," *American Economic Review*, 70(3), 312–326.

DeCicca, P., D. Kenkel, and F. Liu (2013): "Who Pays Cigarette Taxes? The Impact of Consumer Price Search," *Review of Economics and Statistics*, 95(2), 516–529.

Delipalla, S., and M. Keen (1992): "The comparison between ad valorem and specific taxation under imperfect competition," *Journal of Public Economics*, 49(3), 351 – 367.

Doyle Jr., J. J., and K. Samphantharak (2008): "$2.00 Gas! Studying the Effects of a Gas Tax Moratorium," *Journal of Public Economics*, 92(3), 869–884.

Eichenbaum, M., N. Jaimovich, and S. Rebelo (2011): "Reference Prices, Costs, and Nominal Rigidities," *American Economic Review*, 101(1), 234–62.

Eichenbaum, M., N. Jaimovich, S. Rebelo, and J. Smith (2014): "How Frequent Are Small Price Changes?," *American Economic Journal: Macroeconomics*, 6(2), 137–55.

Ellison, S. F., C. Snyder, and H. Zhang (2015): "Costs of Managerial Attention and Activity as a Source of Sticky Prices: Structural Estimates from an Online Market," Working Paper.

Fabinger, M., and E. G. Weyl (2012): "Pass-Through and Demand Forms," Unpublished Mansucript.

——— (2013): "Pass-Through As An Economic Tool: Principles of Incidence Under Imperfect Competition," *Journal of Political Economy*, 121(3), 528–583.

Fullerton, D., and G. E. Metcalf (2002): *Handbook of Public Economics, Volume IV* chap. Tax Incidence. North-Holland.

Goldberg, P., and R. Hellerstein (2013): "A Structural Approach to Identifying the Sources of Local-Currency Price Stability," *Review of Economic Studies*, 80(1), 175–210.

Hamilton, S. F. (2009): "Excise Taxes with Multiproduct Transactions," *American Economic Review*, 99(1), 458–71.

Harberger, A. C. (1964): "The Measurement of Waste," *American Economic Review*, 54(3), 58–76.

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

Hastie, T., R. Tibshirani, and J. Friedman (2001): *The Elements of Statistical Learning*, Springer Series in Statistics. Springer New York Inc., New York, NY, USA.

Kalouptsidi, M., P. Scott, and E. Souza-Rodrigues (2015): "Identification of Counterfactuals and Payoffs in Dynamic Discrete Choice with an Application to Land Use," Unpublished Mansucript.

Kashyap, A. K. (1995): "Sticky Prices: New Evidence from Retail Catalogs," *The Quarterly Journal of Economics*, 110(1), 245–274.

Katz, M. L., and H. S. Rosen (1985): "Tax Analysis in an Oligopoly Model," *Public Finance Quarterly*, 13(1), 3–19.

Kehoe, P. J., and V. Midrigan (2007): "Sales and the Real Effects of Monetary Policy," Working Paper.

Kenkel, D. S. (2005): "Are Alcohol Tax Hikes Fully Passed Through to Prices? Evidence from Alaska," *American Economic Review*, 95(2), 273–277.

KLEVEN, H. J., AND M. WASEEM (2013): "Using notches to uncover optimization frictions and structural elasticities: Theory and evidence from Pakistan," *The Quarterly Journal of Economics*, 128(2), 669–723.

KNOTEK, E. S. (2010): "Convenient Prices and Price Rigidity: Cross-Sectional Evidence," *Review of Economics and Statistics*, 93(3), 1076–1086.

KOPCZUK, W., AND D. MUNROE (2015): "Mansion Tax: The Effect of Transfer Taxes on the Residential Real Estate Market," *American Economic Journal: Economic Policy*, 7(2), 214–57.

LACETERA, N., D. G. POPE, AND J. R. SYDNOR (2012): "Heuristic Thinking and Limited Attention in the Car Market," *American Economic Review*, 102(5), 2206–2236.

LEVY, D., M. BERGEN, S. DUTTA, AND R. VENABLE (1997): "The Magnitude of Menu Costs: Direct Evidence from Large U. S. Supermarket Chains," *The Quarterly Journal of Economics*, 112(3), 791–824.

LEVY, D., D. LEE, H. A. CHEN, R. J. KAUFFMAN, AND M. BERGEN (2011): "Price Points and Price Rigidity," *Review of Economics and Statistics*, 93(4), 1417–1431.

MARION, J., AND E. MUEHLEGGER (2011): "Fuel tax incidence and supply conditions," *Journal of Public Economics*, 95(9), 1202–1212.

MIRAVETE, E., K. SEIM, AND J. THURK (2014): "Complexity, Efficiency, and Fairness in Multiproduct Liquor Pricing," Working Paper.

NAKAMURA, E., AND J. STEINSSON (2008): "Five Facts about Prices: A Reevaluation of Menu Cost Models," *The Quarterly Journal of Economics*, 123(4), 1415–1464.

NAKAMURA, E., AND D. ZEROM (2010): "Accounting of Incomplete Pass-Through," *Review of Economic Studies*, 77(3), 1192 – 1230.

NELSON, J. P. (2013): "Meta-analysis of alcohol price and income elasticities –with corrections for publication bias," *Health Economics Review*, 3, 17–17.

PAKES, A., AND P. MCGUIRE (1994): "Computing Markov-Perfect Nash Equilibria: Numerical Implications of a Dynamic Differentiated Product Model," *The RAND Journal of Economics*, 25(4), 555–589.

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

RUST, J. (1987): "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica*, 55(5), 999–1033.

SEADE, J. (1985): "Profitable cost increases and the shifting of taxation: equilibrium response of markets in Oligopoly," Discussion paper, University of Warwick, Department of Economics.

SEIM, K., AND J. WALDFOGEL (2013): "Public Monopoly and Economic Efficiency: Evidence from the Pennsylvania Liquor Control Board's Entry Decisions," *American Economic Review*, Forthcoming.

Slade, M. E. (1998): "Optimal Pricing with Costly Adjustment: Evidence from Retail-Grocery Prices," *The Review of Economic Studies*, 65(1), 87–107.

Slemrod, J. (2013): "Buenas notches: lines and notches in tax system design," *eJournal of Tax Research*, 11(3), 259.

Stern, N. (1987): "The effects of taxation, price control and government contracts in oligopoly and monopolistic competition," *Journal of Public Economics*, 32(2), 133 – 158.

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*, 104(2), 179–190.

Young, D. J., and A. Bielinska-Kwapisz (2002): "Alcohol Taxes and Beverage Prices," *National Tax Journal*, 55(1), 57–73.

Zhang, X. (2015): "Consistency of model averaging estimators," *Economics Letters*, 130, 120 – 123.

Table 1: Recent Changes in Distilled Spirits Taxes

State	Old Tax	New Tax	Effective Date	Notes
Connecticut	\$4.50/gal + 6% sales tax	\$5.40/gal + 6.35% sales tax	July 1, 2011	
Illinois	\$4.50/gal + 6.25% sales tax	\$8.55/gal + 6.25% sales tax	Sept 1, 2009	additional local sales tax
Kentucky	\$1.92/gal	\$1.92/gal + 6% sales tax	April 1, 2009	additional 11% wholesale tax
Maryland	\$1.50/gal + 6% sales tax	\$1.50/gal + 9% sales tax	July 1, 2011	
Massachusetts	\$4.05/gal	\$4.05/gal + 6.25% sales tax	September 1, 2009	Ended Jan 1, 2011
New Jersey	\$4.40/gal + 7% sales tax	\$5.50/gal + 7% sales tax	August 1, 2009	sales tax was 6% before 7/1/06
Rhode Island	\$3.75/gal + 7% sales tax	\$5.40/gal + 0% sales tax	December 1, 2013	

Note: The table above describes the nature and timing of the tax changes for each of the seven states that have altered their taxation of alcohol since 2007.

Table 2: Mean Change in Retail and Wholesale Prices, CT

	Wholesale Prices						Retail Prices				
month	2008	2009	2010	2011	2012	month	2008	2009	2010	2011	2012
1	0.381	0.421	0.460	0.338	0.469	1	-0.018	0.241	0.350	0.362	0.350
2	-0.923	-1.258	-1.110	-0.899	-1.050	2	0.139	-0.221	-0.317	-0.197	-0.166
3	0.694	0.850	0.714	0.420	0.822	3	-0.042	-0.058	-0.139	-0.105	0.014
4	-0.222	-0.180	-0.220	0.011	-0.017	4	-0.029	-0.018	-0.024	-0.010	-0.073
5	0.244	0.077	0.002	-0.102	-0.023	5	-0.046	-0.129	-0.194	-0.134	-0.009
6	-0.836	-0.927	-1.030	-0.760	-0.707	6	0.080	-0.146	-0.152	-0.121	-0.090
7	0.878	1.089	1.348	**1.462**	0.991	7	0.097	-0.042	0.070	**0.422**	0.020
8	0.124	-0.100	-0.188	-0.340	-0.165	8	0.048	0.124	-0.112	-0.106	0.160
9	-0.108	-0.124	-0.119	0.173	0.189	9	0.004	-0.199	-0.062	0.000	-0.019
10	-0.745	-0.785	-0.563	-0.790	-0.836	10	0.003	-0.023	-0.112	-0.098	-0.044
11	0.204	0.102	-0.270	0.271	0.465	11	-0.040	-0.323	-0.317	-0.303	-0.175
12	0.710	0.480	0.750	0.207	0.121	12	0.009	-0.113	-0.075	-0.097	-0.095

Note: The table above reports the average monthly change in Connecticut retail and wholesale prices for each month between 2008 and 2012, weighted by retail units.

Table 3: Retail and Wholesale Monthly Price Change Increments

Retail Prices			Retail Prices, 07/2011			Wholesale Prices			Wholesale Prices, 07/2011		
0	384,889	-	0	3,566	-	0	24,013	-	0	124	-
1	15,154	20.1%	1	788	38.6%	-2	2,738	6.8%	1	125	11.3%
2	7,372	9.8%	2	253	12.4%	2	2,695	6.7%	0.25	61	5.5%
-1	7,141	9.5%	-1	152	7.5%	1	1,681	4.2%	0.5	55	5.0%
-2	5,849	7.8%	4	80	3.9%	-1	1,680	4.2%	4	50	4.5%
3	4,121	5.5%	-2	77	3.8%	-4	1,621	4.0%	2.5	47	4.2%
-3	3,638	4.8%	3	72	3.5%	4	1,619	4.0%	0.67	40	3.6%
4	2,665	3.5%	0.6	54	2.6%	3	1,514	3.8%	3	29	2.6%
-4	2,357	3.1%	0.5	50	2.5%	-3	1,427	3.5%	5	29	2.6%
0.5	1,602	2.1%	1.5	41	2.0%	-1.5	1,308	3.2%	-1	28	2.5%
5	1,313	1.7%	-3	33	1.6%	1.5	1,213	3.0%	0.3	20	1.8%
Other	24,015	31.9%	Other	439	21.5%	Other	22,787	56.6%	Other	627	56.4%
Number $\neq 0$	75,227	100.0%	Number $\neq 0$	2,039	100.0%	Number $\neq 0$	40,283	100.0%	Number $\neq 0$	1,111	100.0%
Whole Dollar	56,580	75.2%	Whole Dollar	1,577	77.3%	Whole Dollar	17,647	43.8%	Whole Dollar	336	30.2%
.99 to .99	55,963	74.4%	.99 to .99	1,565	76.8%	.91 to .91	13,295	33.0%	.91 to .91	287	25.8%
Total	460,116		Total	5,605		Total	64,296		Total	1,235	

Note: The table above reports the frequency of retail and wholesale price change increments for Connecticut overall and in July 2011. The percentages reflect the percent of the number of non-zero price changes (number $\neq 0$). These counts are unweighted.

Table 4: Frequency of Cents Portion of Retail and Wholesale Prices

Retail Prices: CT			Wholesale Prices: CT			Retail Prices: IL		
99	416,736	90.6%	91	32,435	50.4%	99	4,223,452	80.7%
49	16,419	3.6%	41	6,893	10.7%	97	277,847	5.3%
59	7,666	1.7%	58	2,194	3.4%	49	246,216	4.7%
89	7,447	1.6%	16	2,074	3.2%	98	89,935	1.7%
93	2,064	0.4%	24	1,992	3.1%	79	53,477	1.0%
69	1,531	0.3%	74	1,925	3.0%	0	26,298	0.5%
95	1,089	0.2%	79	1,537	2.4%	29	25,029	0.5%
79	822	0.2%	8	1,402	2.2%	48	23,263	0.4%
Other	6,342	1.4%	Other	13,844	21.5%	Other	269,321	5.1%
Total	460,116		Total	64,296		Total	5,234,838	

Note: The table above reports the frequency of different price points in terms of the cents portion of the price in Connecticut and Illinois. Connecticut's minimum pricing law requires all 750mL, 1L, and 1.75L bottles be priced at least 8 cents above the wholesale price. Thus a wholesale price ending in 91 cents reflects a minimum retail price ending in 99 cents.

Table 5: Pass-Through: Taxes to Retail Prices (Connecticut)

	All Retailers			Δ Retail Price≠ 0		
	Connecticut July 1, 2011 Tax Increase of $0.24/L					
Δ Retail Price	1m	3m	6m	1m	3m	6m
	(1)	(2)	(3)	(4)	(5)	(6)
Δ Tax	1.533***	1.257***	1.013***	3.096***	2.301***	2.016***
	(0.271)	(0.202)	(0.264)	(0.706)	(0.479)	(0.553)
Δ Tax*I[size=750mL]	1.168***	1.900***	2.084***	3.191**	3.822***	4.072***
	(0.432)	(0.387)	(0.503)	(1.577)	(0.899)	(1.144)
Δ Tax*I[size=1L]	2.146***	1.833***	1.586***	5.550***	3.376***	3.553***
	(0.650)	(0.383)	(0.470)	(1.663)	(0.920)	(1.132)
Δ Tax *I[size=1.75L]	1.520***	1.154***	1.009***	2.985***	2.191***	2.027***
	(0.309)	(0.227)	(0.263)	(0.718)	(0.502)	(0.570)
Observations	460,116	437,057	410,288	75,227	113,098	142,220
Product FE	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Note: The table above reports OLS estimates of the pass-through of taxes into retail prices in Connecticut. All regressions are weighted by 2011 Nielsen units. Standard errors are clustered at the UPC level.
*** Significant at the 1 percent level, **Significant at the 5 percent level and *Significant at the 10 percent level.

Table 6: Pass-Through: Taxes to Retail Prices (Illinois)

	All Retailers			Δ Retail Price≠ 0		
	Illinois Sept 1, 2009 Tax Increase of $1.07/L					
Δ Retail Price	1m	3m	6m	1m	3m	6m
	(1)	(2)	(3)	(4)	(5)	(6)
Δ Tax	0.487***	0.965***	1.267***	1.092***	1.251***	1.493***
	(0.044)	(0.045)	(0.052)	(0.107)	(0.069)	(0.065)
Δ Tax*I[size=750mL]	0.575***	1.363***	1.828***	0.977***	1.780***	2.239***
	(0.082)	(0.083)	(0.079)	(0.175)	(0.134)	(0.095)
Δ Tax*I[size=1L]	0.904***	1.046***	1.307***	1.145***	1.188***	1.443***
	(0.144)	(0.134)	(0.125)	(0.123)	(0.091)	(0.102)
Δ Tax *I[size=1.75L]	0.442***	0.826***	1.062***	1.157***	1.091***	1.256***
	(0.050)	(0.049)	(0.059)	(0.108)	(0.069)	(0.064)
Observations	5,234,838	4,974,620	4,687,529	2,206,133	2,907,497	3,200,085
Product FE	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Note: The table above reports OLS estimates of the pass-through of taxes into retail prices in Illinois. All regressions are weighted by 2009 Nielsen units. Standard errors are clustered at the UPC level.
*** Significant at the 1 percent level, **Significant at the 5 percent level and *Significant at the 10 percent level.

Table 7: Three Month Pass-through Using Pooled Data (CT,IL)

	Δ Price				
	(1)	(2)	(3)	(4)	(5)
$\Delta\tau$	1.048***		1.023***		
	(0.044)		(0.043)		
$\Delta\tau \cdot I[750]$		1.400***		1.399***	
		(0.062)		(0.070)	
$\Delta\tau \cdot I[1000]$		1.514***		1.365***	
		(0.167)		(0.161)	
$\Delta\tau \cdot I[1750]$		0.875***		0.864***	
		(0.050)		(0.049)	
$\Delta\tau \cdot I[750] \cdot CT$					1.936***
$\Delta\tau = .1783$					(0.386)
$\Delta\tau \cdot I[1000] \cdot CT$					1.825***
$\Delta\tau = .2377$					(0.382)
$\Delta\tau \cdot I[1750] \cdot CT$					1.141***
$\Delta\tau = .4161$					(0.226)
$\Delta\tau \cdot I[750] \cdot IL$					1.359***
$\Delta\tau = .8024$					(0.083)
$\Delta\tau \cdot I[1000] \cdot IL$					1.061***
$\Delta\tau = 1.070$					(0.127)
$\Delta\tau \cdot I[1750] \cdot IL$					0.830***
$\Delta\tau = 1.872$					(0.049)
Constant	0.023***	0.021***			
	(0.005)	(0.005)			
UPC FE	No	No	Yes	Yes	Yes
Month FE	No	No	Yes	Yes	x State
Year FE	No	No	Yes	Yes	x State
N	5,411,677	5,411,677	5,411,677	5,411,677	5,411,677
R^2	0.011	0.011	0.026	0.026	0.029

Note: The table above reports OLS estimates of the pass-through of taxes into retail prices over three-month differences using data from Connecticut and Illinois. All reported standard errors are clustered at the UPC level and weighted by state-level shares (rather than units).
*** Significant at the 1 percent level, **Significant at the 5 percent level and *Significant at the 10 percent level.

Table 8: Pass-Through: Taxes to Wholesale Prices (Connecticut)

	All Wholesalers			Lowest Price Wholesaler		
Δ Wholesale Price	1m	3m	6m	1m	3m	6m
	(1)	(2)	(3)	(4)	(5)	(6)
Δ Tax	1.302***	0.960***	0.805***	1.616***	1.290***	1.135***
	(0.368)	(0.278)	(0.255)	(0.465)	(0.297)	(0.252)
Δ Tax*I[size=750mL]	3.598	3.265***	3.446***	2.932	3.330***	3.486***
	(2.562)	(1.035)	(0.579)	(2.333)	(0.976)	(0.502)
Δ Tax*I[size=1L]	-2.295	0.047	1.664***	-0.990	0.927	2.217***
	(1.946)	(0.899)	(0.544)	(1.957)	(0.865)	(0.516)
Δ Tax *I[size=1.75L]	1.274***	0.870***	0.816***	1.609***	1.189***	1.127***
	(0.487)	(0.298)	(0.252)	(0.585)	(0.317)	(0.251)
Observations	64,296	60,841	56,798	42,988	41,080	38,538
Product FE	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Note: The table above reports OLS estimates of the pass-through of taxes into wholesale prices in Connecticut. All regressions are weighted by 2011 Nielsen units. Standard errors are clustered at the product level.

Table 9: Pass-Through: Taxes to Retail Prices Relative to Other Stores

	Above/Below Median	Min/Max	Continuous
	(1)	(2)	(3)
Δ Tax	1.532***	1.516***	1.522***
	(0.280)	(0.270)	(0.248)
Δ Tax * High	-1.218***	-1.252***	
	(0.366)	(0.380)	
Δ Tax * Low	1.357***	1.460***	
	(0.500)	(0.496)	
Δ Tax * Relative			-0.174***
			-0.026
High Price	-0.423***	-0.382***	
	(0.043)	(0.039)	
Low Price	0.198***	0.287***	
	(0.042)	(0.040)	
Relative to Median			-0.042***
			(0.004)
Observations	460,116	460,116	460,116
Product FE	Yes	Yes	Yes
Month+Year FE	Yes	Yes	Yes
High Measure	Above Median	Maximum	Continuous
Low Measure	Below Median	Minimum	% Deviation

Note: The table above reports OLS estimates of the pass-through of taxes into retail prices in Connecticut for low and high-priced stores. All regressions weighted by 2011 Nielsen units and reported standard errors are clustered at the UPC level. Relative prices are from PRIOR month. Columns (1)+(2) use indicator variables. Column (3) uses percentage deviation from median price.

Table 10: Quarterly Price Change $-1 to $3, Ordered Estimates

	Ordered Logit Index	
	CT Only	CT and IL
	(1)	(2)
$\Delta\tau$	1.156***	−0.303**
	(0.112)	(0.154)
$\Delta\tau^2$		4.670***
		(0.278)
$\Delta\tau^3$		−2.058***
		(0.105)
Highest Price	−0.347***	−0.127***
	(0.015)	(0.003)
Lowest Price	0.187***	0.373***
	(0.014)	(0.003)
Relative Price	−142.858***	−726.550***
	(0.411)	(0.0001)
Relative Price^2	60.923***	165.217***
	(2.624)	(0.0001)
Relative Price^3	−26.817***	−68.847***
	(2.579)	(0.0001)
$\log(p_{j,t-1})$	−0.062***	−0.048***
	(0.012)	(0.003)
log(2010 Sales)	−0.053***	0.008***
	(0.003)	(0.001)
log(Store Size)	0.016***	−0.037***
	(0.003)	(0.001)
log(# Competitors)	−0.020**	−0.010***
	(0.009)	(0.001)
$\Delta p_w = 0$	−2.206***	
	(0.025)	
Δp_w	31.608***	
	(1.834)	
Δp_w^2	22.303***	
	(2.680)	
Δp_w^3	−5.765**	
	(2.251)	
Year FE	Yes	Yes
Quarter FE	Yes	Yes
Classification Accuracy	0.639	0.564
Null Accuracy	0.572	0.572
Log Likelihood	349,103	6,714,827
N	224,276	2,693,076

Note: Polynomial order selected via BIC.
*** Significant at the 1 percent level, **Significant at the 5 percent level and *Significant at the 10 percent level.

Figure 1: Probability of $1 Tax increase at different tax sizes

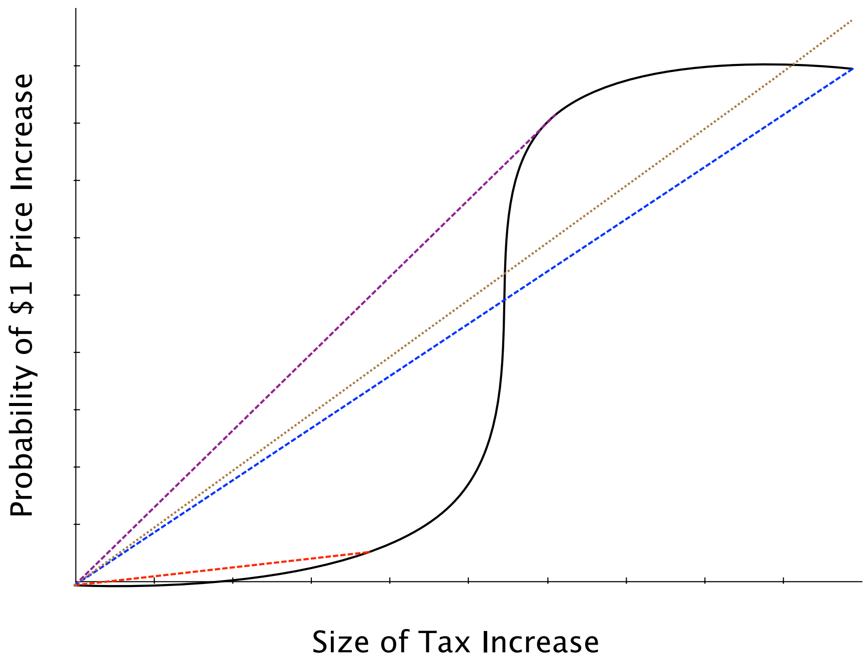


Figure 2: Change in Surplus When Price is Changed

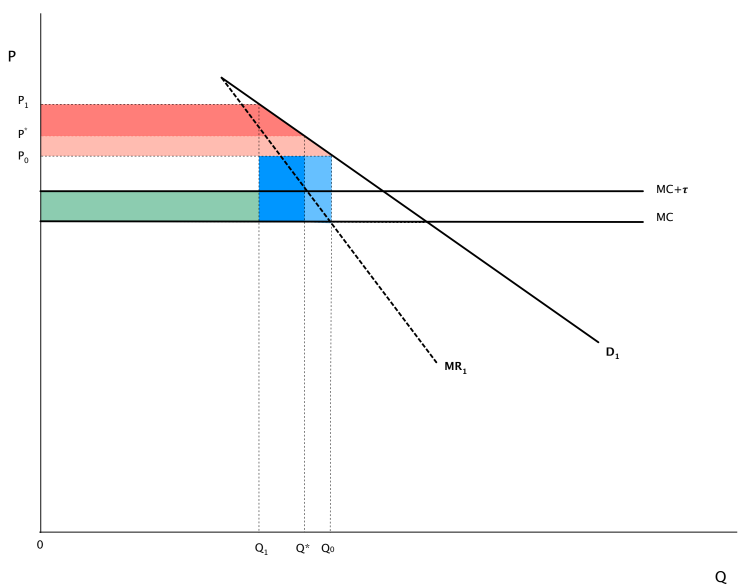
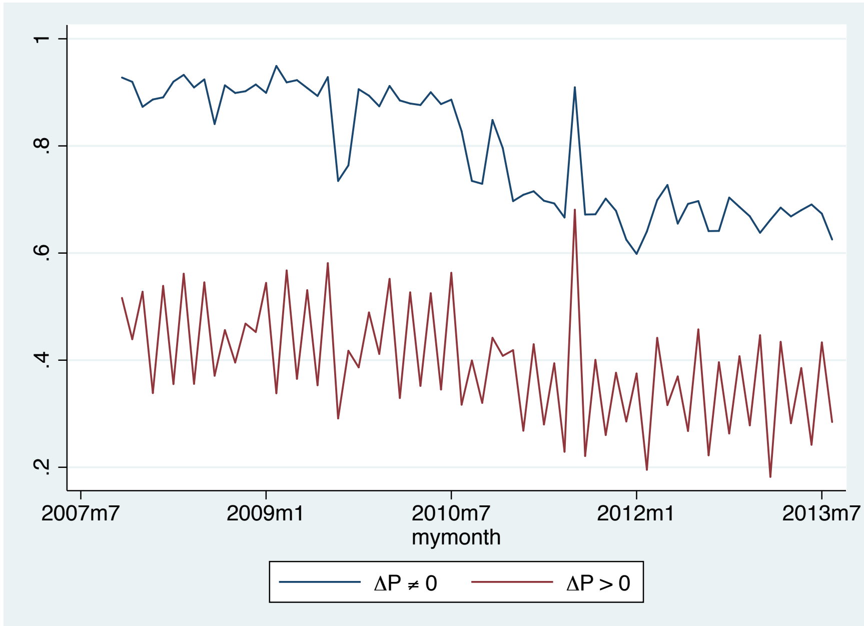
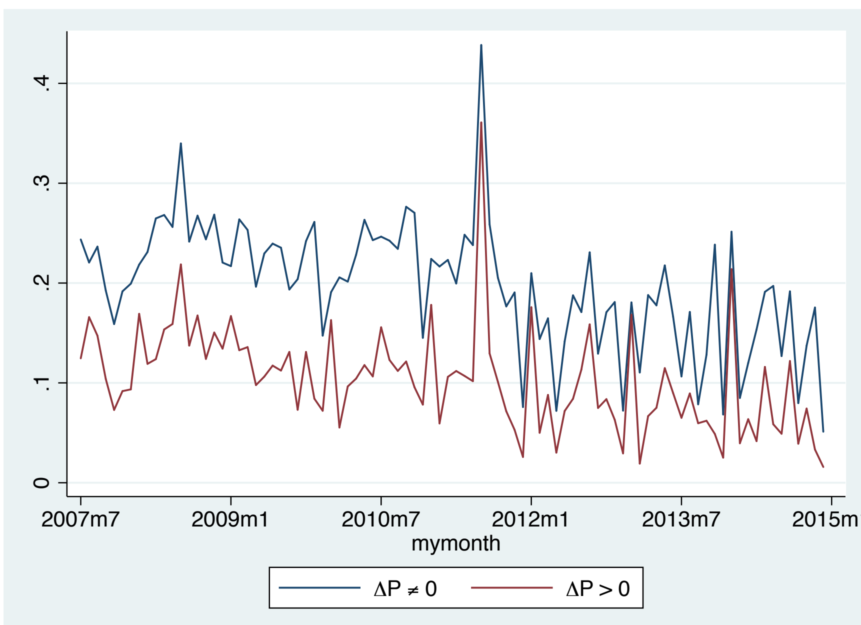


Figure 3: Frequency of Wholesale Price Adjustment by Month, CT



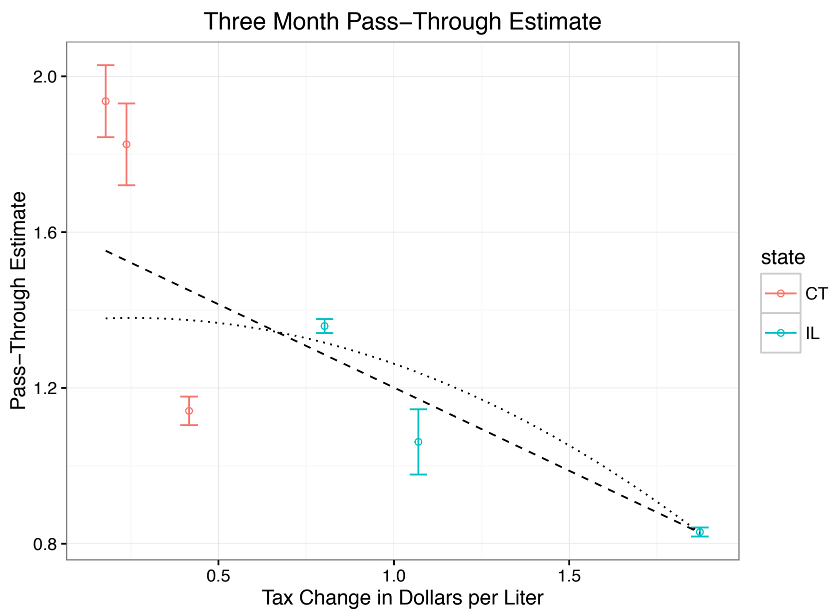
Note: The figure above reports the share of wholesale prices that change and that increase for each month between 2008 and 2012, weighted by retail units.

Figure 4: Frequency of Retail Price Adjustment, CT



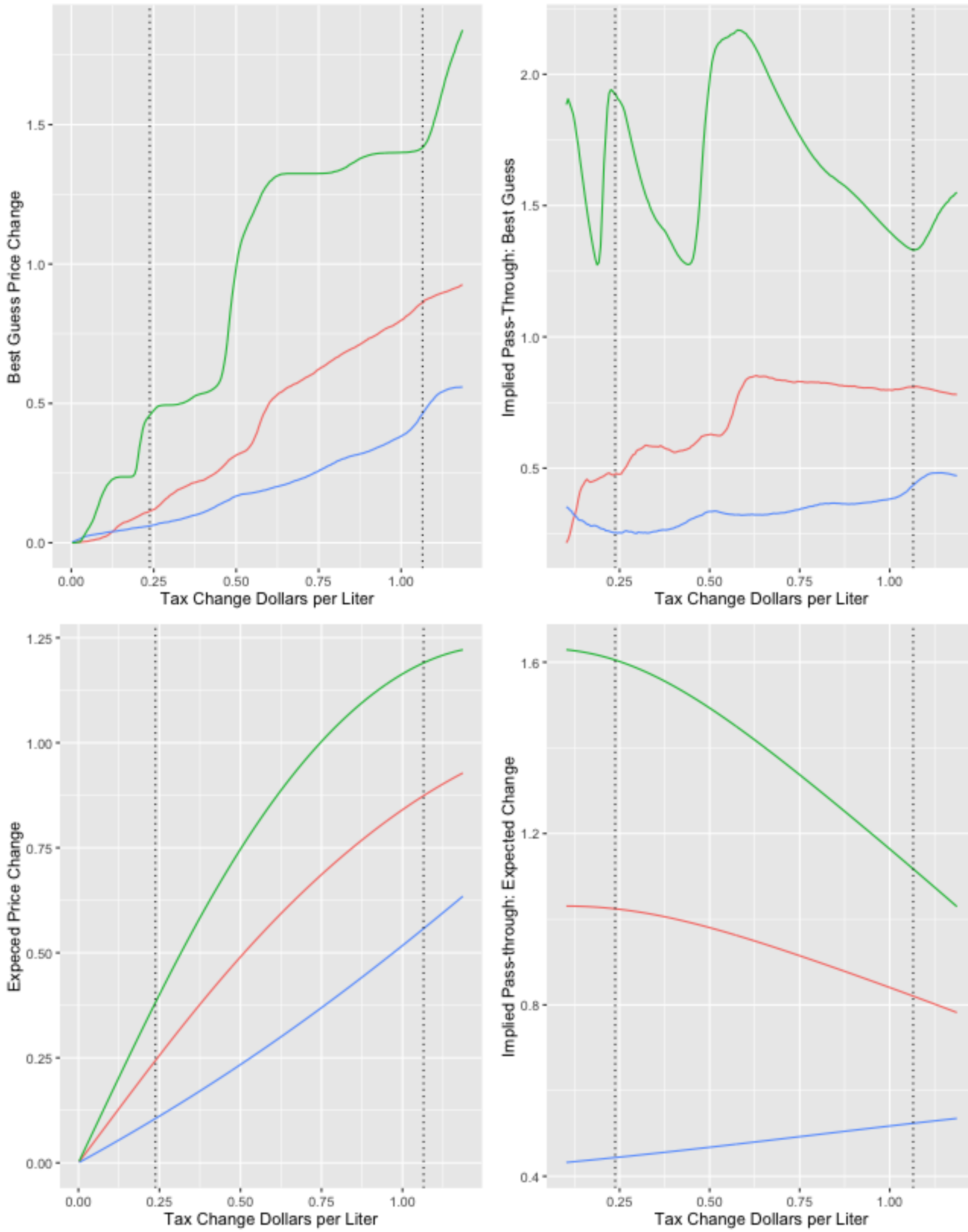
Note: The figure above plots the share of retail prices that change and that increase for each month between July 2007 and 2012, weighted by retail units.

Figure 5: Pass-Through Estimates



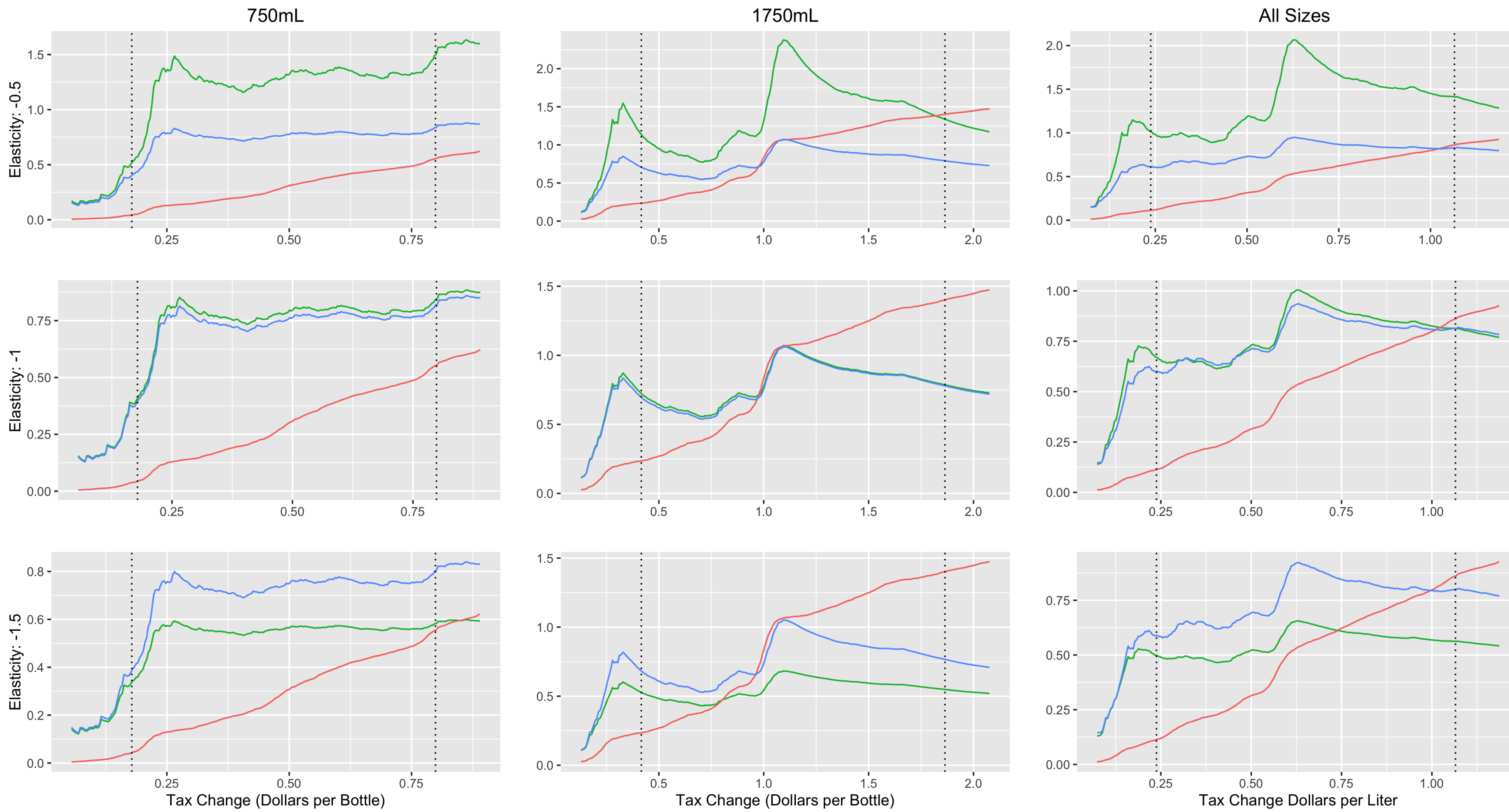
Note: The figure above plots the estimated pass-through rates for different size tax increases. The dashed line estimates the pass-through rate as a linear function of the tax, and the dotted line estimates the pass-through rate as a quadratic function of the tax.

Figure 6: Predicted Price Changes and Incidence; Ordered Logit



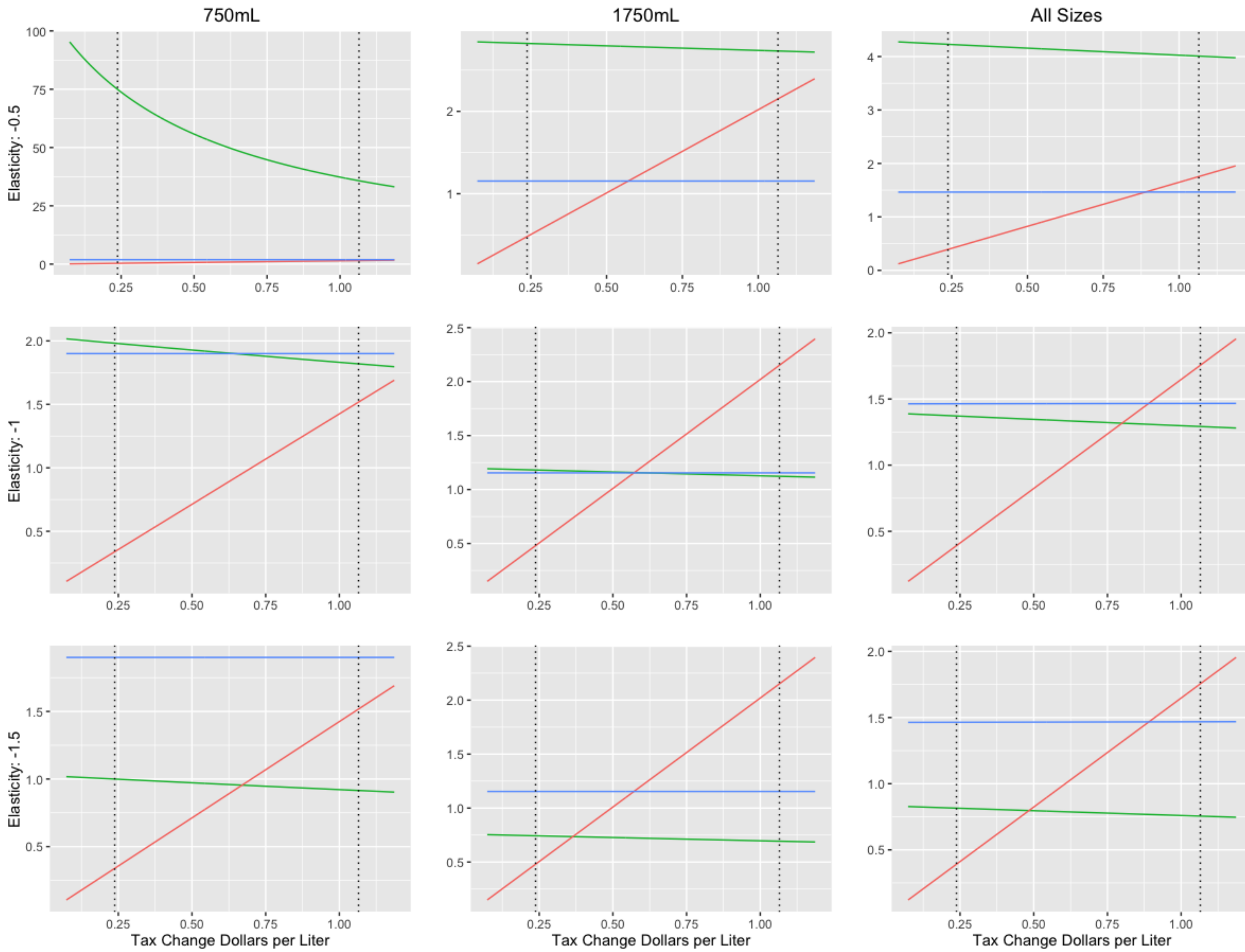
Red Line: Averaged Model; Blue Line: CT Only Linear in $\Delta\tau$; Green Line: Both States Cubic in $\Delta\tau$

Figure 7: Counterfactual Welfare: Ordered Logit



Red Line: Average Price Increase; Green Line: Incidence $\Delta CS/\Delta PS$; Blue Line: Implied Pass-Through $E[\Delta p_j/\Delta \tau_j]$.

Figure 8: Counterfactual Welfare: Constant Pass-Through/ Linear Model



Red Line: Average Price Increase; Green Line: Incidence $\Delta CS/\Delta PS$; Blue Line: Implied Pass-Through $E[\Delta p_j/\Delta\tau_j]$.

Figure 9: Deadweight Loss Per Dollar of Tax Revenue: Ordered Logit

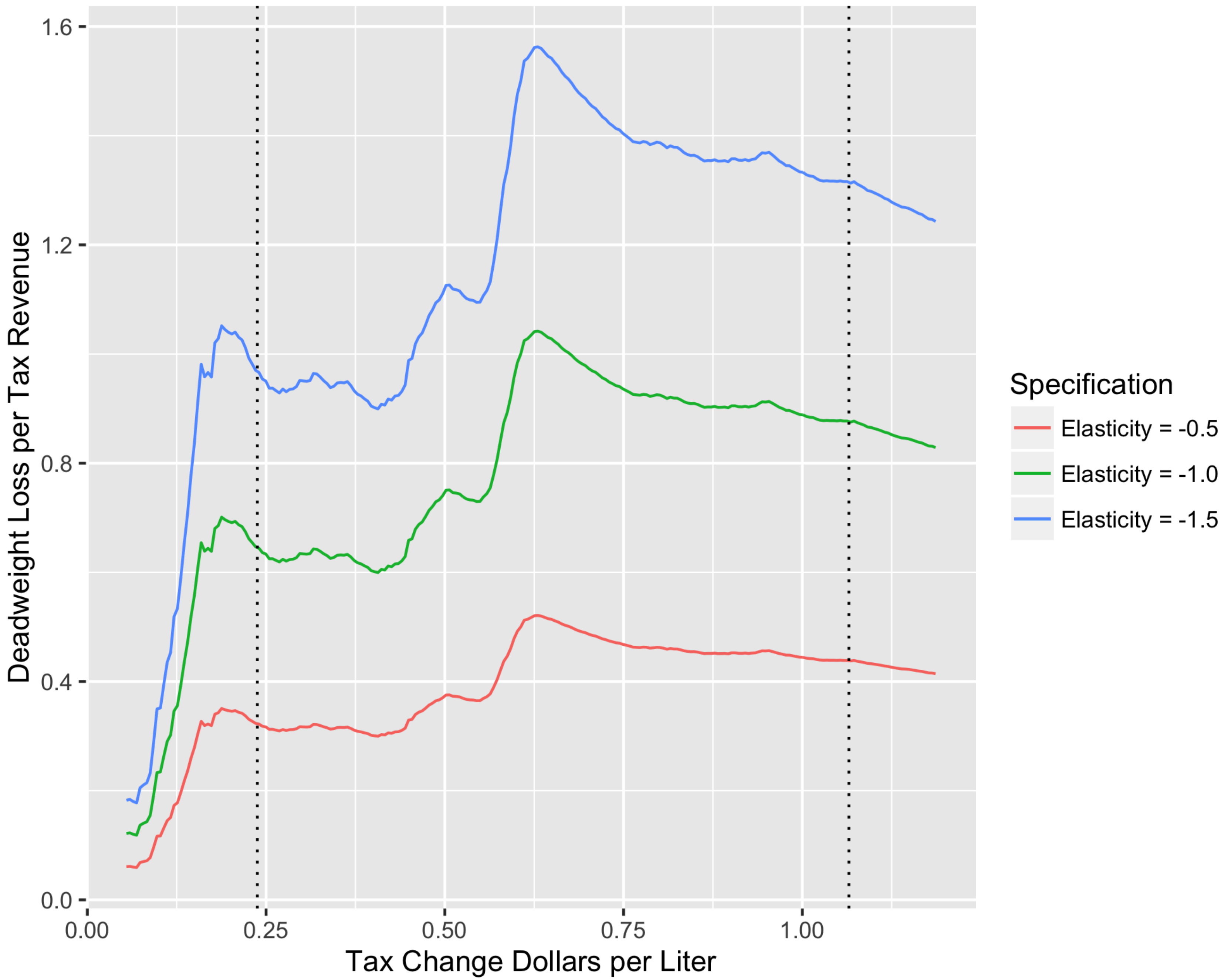
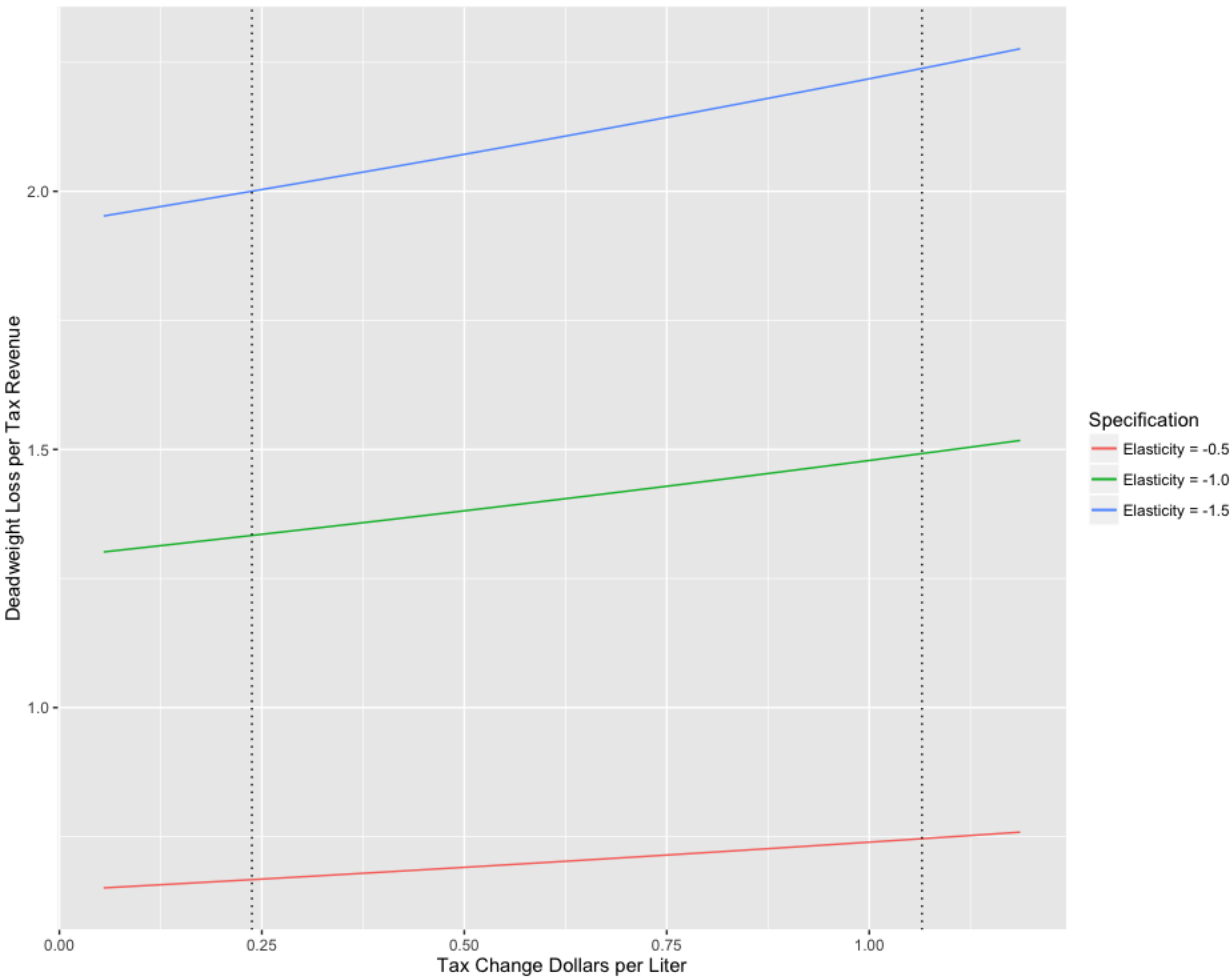


Figure 10: Deadweight Loss Per Dollar of Tax Revenue: Constant Pass-Through/Linear



Online Appendix A: Binary Choice Specification

In this Appendix, we fit a more restricted version of our main specification where we restrict price changes only to $\Delta P \in \{\$0, \$1\}$. Instead of an ordered logit regression, we employ a binary logit regression where the two choices are: keep the same price, increase price by exactly \$1. We relabel all price decreases as $\Delta p_j = 0$ and all price increases of $\geqslant \$1.00$ as $\Delta p_j = 1$. We report the results in Table 11. The parameters all have similar to signs for those in the ordered logit model. However, the fit is substantially improved as we no longer have to predict which products increase prices by \$2.00 from those that increase by \$1.00. We include an additional specification that includes potentially endogenous regressors, which we label as such. The additional endogenous regressor is (a fifth order polynomial in) the fraction of competitors selling the same product that also increased their prices during the same period. This substantially improves the classification accuracy from approximately 81% to 85% on the CT data and 69% to 80% for the CT and IL combined dataset. We can interpret this extra endogenous variable as the observed action of the competitors of the dynamic game. It also presents a challenge for prediction, because as taxes increase we might also expect the fraction of competitors changing prices to increase, thus we would need some sort of fixed point or notion of equilibrium in order to make realistic predictions.

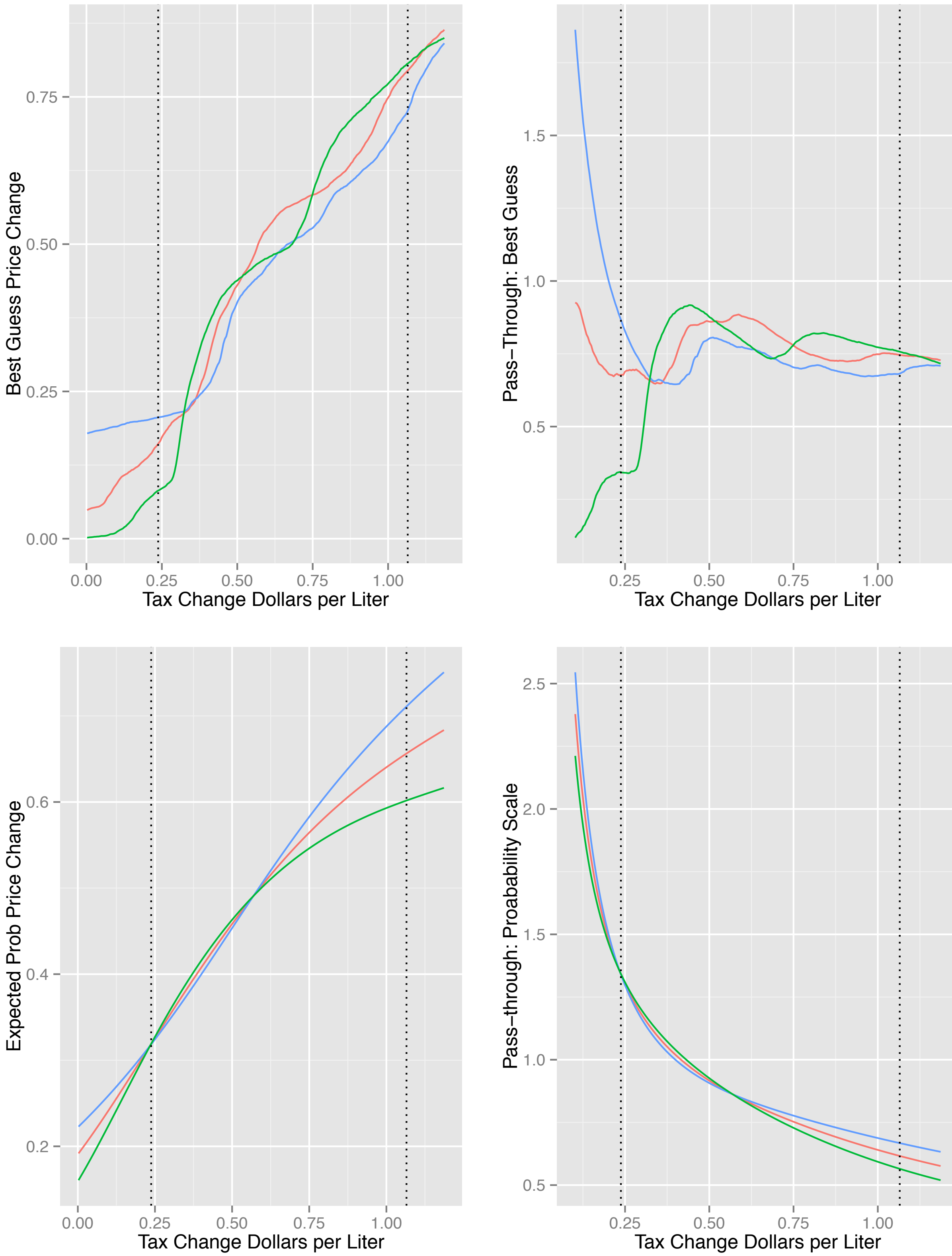
We also present fit comparisons in Figure 11. Unlike the ordered logit, we see that the model which includes data from both states (CT and IL) is quite similar to the model which includes data from CT only but allows for heterogeneity in initial conditions (changes in wholesale prices since the last retail price change). Thus whether we use one specification or the other (or any convex combination) we get qualitatively similar results. The main difference between the two models lies in predicted changes outside the range of the two observed tax increases (specifically for very small tax increases).

We also provide the same welfare results that we did for the ordered logit in Figures 12 and 13. We see the same non-monotonicities and U-shaped pattern in both the incidence and the deadweight loss per dollar of tax revenue as we do in the main specification. The downside of the binary logit specification is that should be skeptical of predicted effects for tax increases larger than 60 cents per liter, because the implied tax change for 1.75L bottles would exceed \$1.00 and we are ruling out \$2.00 or larger price changes.

Table 11: Quarterly \$1 Price Change, Logit Estimates

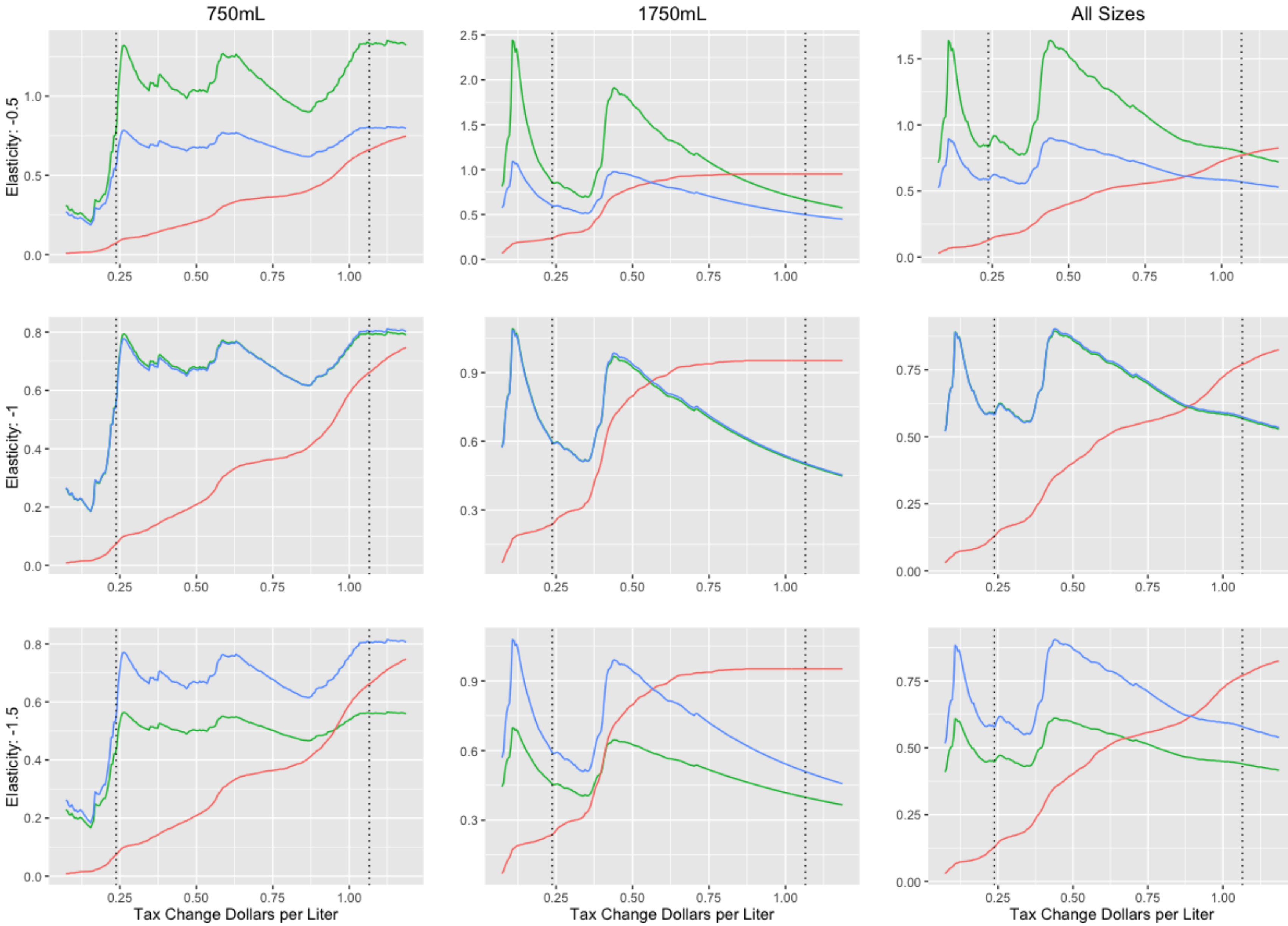
	$Pr(\Delta Price \geqslant \$1 \mid X)$			
	CT Only		CT and IL	
	Exog.	Endog.	Exog.	Endog.
	(1)	(2)	(3)	(4)
$\Delta\tau$	2.695***	0.924***	4.285***	2.128***
	(0.131)	(0.154)	(0.254)	(0.305)
$\Delta\tau^2$			−2.505***	−3.060***
			(0.455)	(0.546)
$\Delta\tau^3$			0.483***	1.140***
			(0.171)	(0.205)
$\log(p_{j,t-1})$	0.401***	−0.057***	0.054***	−0.254***
	(0.015)	(0.017)	(0.004)	(0.004)
log(2010 Sales)	−0.031***	0.065***	−0.016***	0.010***
	(0.004)	(0.005)	(0.001)	(0.001)
log(Store Size)	0.016***	−0.012***	−0.040***	−0.044***
	(0.003)	(0.004)	(0.001)	(0.001)
log(# Competitors)	0.163***	0.256***	0.325***	0.337***
	(0.011)	(0.016)	(0.002)	(0.003)
$\Delta p_w = 0$	−3.131***	−2.804***		
	(0.070)	(0.083)		
Δp_w	−27.051***	−34.271***		
	(6.256)	(7.483)		
Δp_w^2	43.770***	57.804***		
	(10.782)	(12.500)		
Δp_w^3	−40.008***	−63.272***		
	(10.847)	(13.108)		
Δp_w^4	42.240**	81.986***		
	(18.678)	(22.186)		
Highest Price	−0.463***	−0.377***	−0.103***	−0.179***
	(0.028)	(0.031)	(0.005)	(0.007)
Lowest Price	0.286***	0.564***	0.459***	0.726***
	(0.023)	(0.026)	(0.005)	(0.006)
Relative Price	−149.273***	−227.526***	−735.393***	−1,023.625***
	(6.087)	(6.801)	(5.063)	(6.355)
Relative Price^2	42.626***	27.415***	79.554***	108.615***
	(4.815)	(5.182)	(4.150)	(19.317)
Relative Price^3	−12.762**	10.365**	10.535**	223.008***
	(5.455)	(4.901)	(4.094)	(20.737)
Constant	0.153	−0.146		
	(0.093)	(0.124)		
Year FE	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes
Polynomial: Competitor $Pr(\Delta P \geqslant 0)$	No	5th	No	5th
Classification Accuracy	0.808	0.846	0.692	0.797
Null Accuracy	0.675	0.675	0.675	0.675
N	224,276	224,276	2,693,076	2,693,076
Log Likelihood	−79,632.270	−59,412.640	−1,369,850.000	−968,791.700
Akaike Inf. Crit.	159,318.500	118,893.300	2,739,772.000	1,937,669.000

Figure 11: Logit Fit Comparison



Note: The figure above plots the fit for the binary logit using both predicted probability, and best-guess criteria. Vertical lines denote the observed tax change in CT and IL respectively. Blue: CT data, linear specification. Green: CT and IL, Cubic Specification. Red: 50-50 averaged model.

Figure 12: Counterfactual Welfare: Binary Logit



Red Line: Probability of a Price Increase; Green Line: Incidence $\Delta CS/\Delta PS$; Blue Line: $|\Delta CS|/\Delta GovRev$.

Figure 13: Deadweight Loss Per Dollar of Tax Revenue: Binary Logit

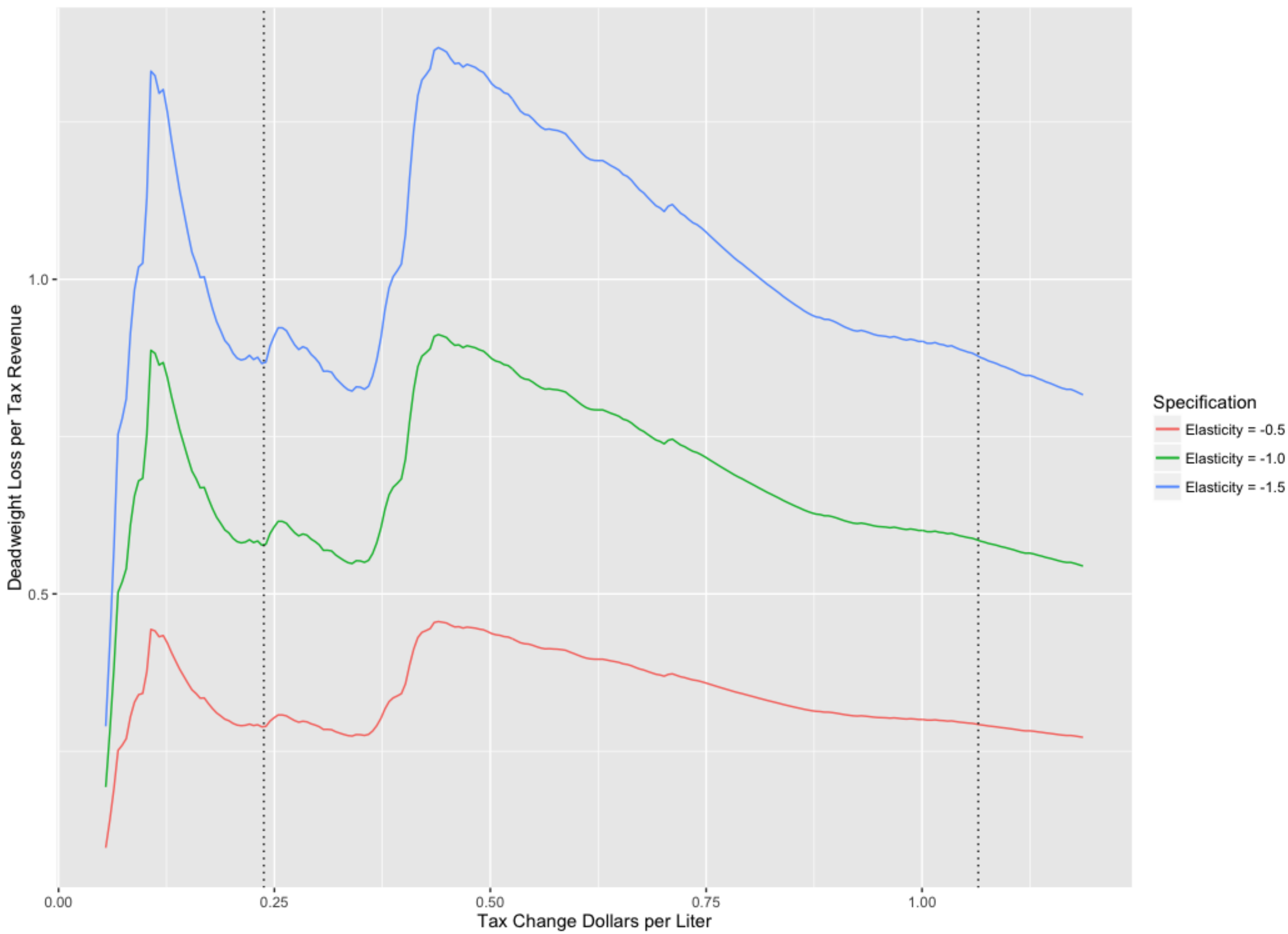


table notes here

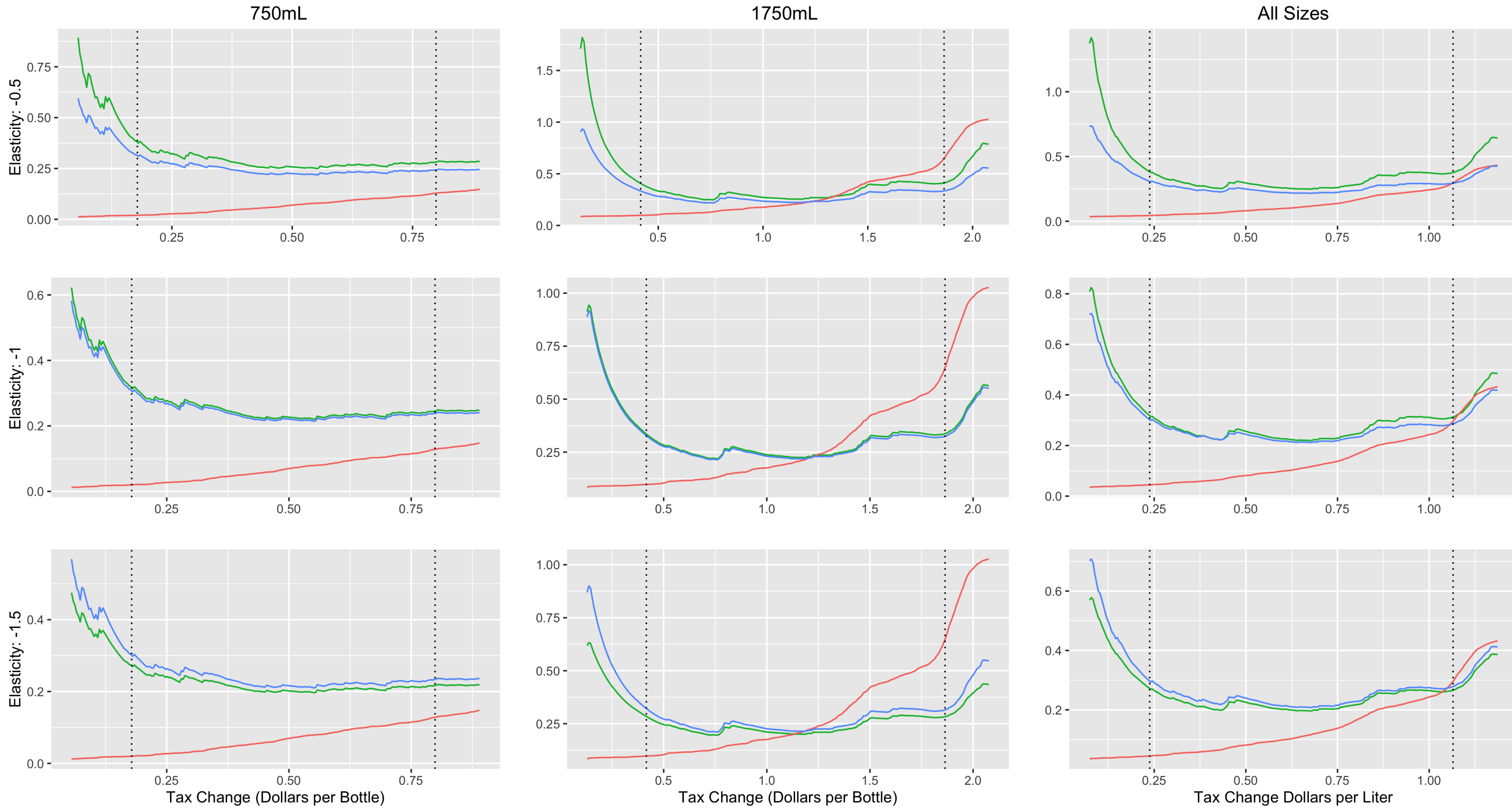
Online Appendix B: Individual Ordered Logit Predictions

In the text we provide incidence and welfare calculation for an averaged ordered logit model. Here we present the same predictions for each of the two models separately. The first model uses only the data from Connecticut and includes changes in wholesale prices, while the second also includes data from Illinois, which expands the domain of the tax increase but does not include wholesale price information. We report counterfactual welfare predictions under each model individually rather than averaged across models.

We present both the incidence plots in Figures 14 and 15 and the deadweight loss per dollar of tax revenue plots in Figures 16 and 17. The major differences from the averaged model are as follows. For the data from CT and IL which do not include changes in wholesale prices as a state variable, we see that price changes tend to be sharply clustered around particular price points: around 55 cents per bottle for both 750mL and 1.75L bottles. The Connecticut only data, which allows for heterogeneity in "initial conditions" but restricts the relationship between the ordered logit index and the tax change to be linear shows less clustering around particular price changes. Both models demonstrate non-monotonicities in both incidence and deadweight loss per dollar of tax revenue. There is some disagreement for the effects of very small tax increases (smaller than those we observe in Connecticut). The CT only model shows high incidence and large deadweight loss for small tax increases, while the CT and IL model shows virtually no price changes, incidence or deadweight loss for very small tax increases. Both of these are outside the domain of the observed data, so we should take both with a grain of salt. Ignoring the smallest tax increases (to the left of the veritcal line for the observed CT tax increase) we see that both models produce similar u-shaped patterns for both incidence and deadweight loss. (The U-shaped pattern for the CT only data is somewhat obscured by the very high incidence and deadweight loss on the leftmost portion of the graph).

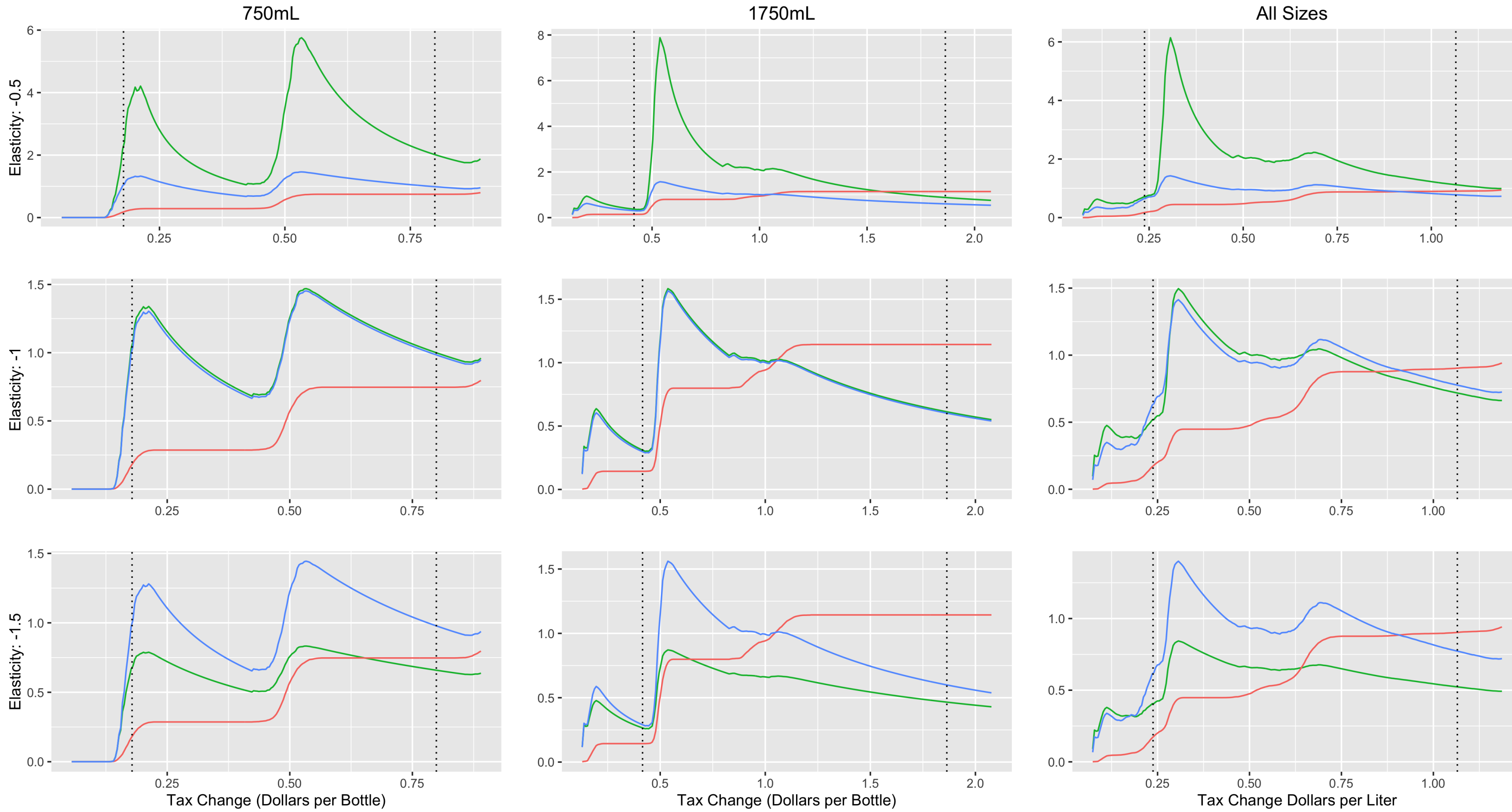
In general, we think this provides both good motivation for our use of an averaged model in the main specification, as well as an indication that the main features are present in both of the disaggregated models and not being driven by one model or the other.

Figure 14: Counterfactual Welfare: Ordered Logit (CT Only)



Red Line: Probability of a Price Increase; Green Line: Incidence $\Delta CS/\Delta PS$; Blue Line: $|\Delta CS|/\Delta GovRev$.

Figure 15: Counterfactual Welfare: Ordered Logit (CT and IL)



Red Line: Probability of a Price Increase; Green Line: Incidence $\Delta CS/\Delta PS$; Blue Line: $|\Delta CS|/\Delta GovRev$.

Figure 16: Deadweight Loss Per Dollar of Tax Revenue: Ordered Logit (CT Only)

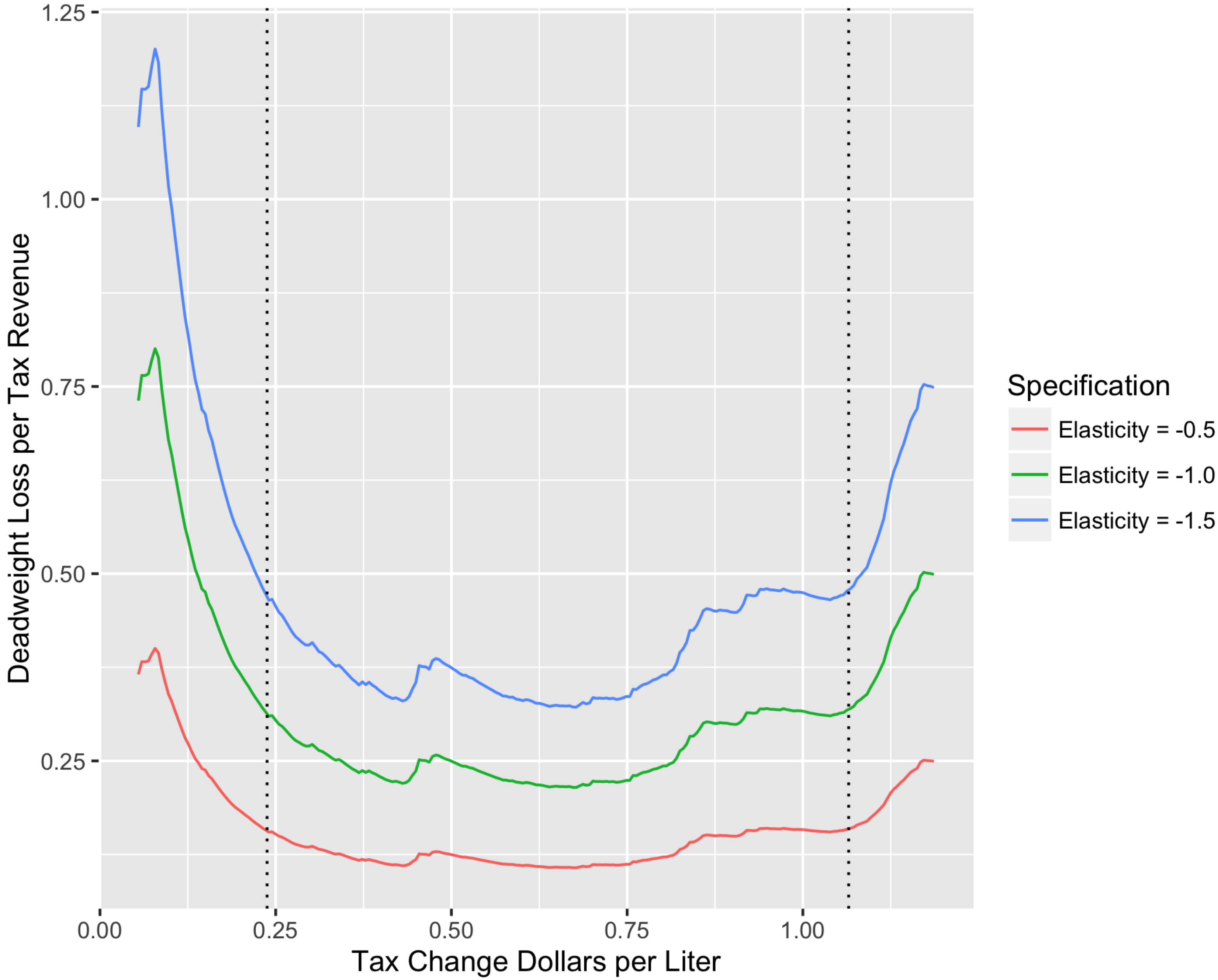


Figure 17: Deadweight Loss Per Dollar of Tax Revenue: Ordered Logit (CT and IL)