LS
7/21/08
10:00 am

# Salience and Taxation: Theory and Evidence

Raj Chetty
UC-Berkeley and NBER

Adam Looney
Federal Reserve Board

Kory Kroft
UC-Berkeley

May 2008*

**Abstract**

This paper presents evidence that consumers under react to taxes that are not salient and characterizes the welfare consequences of tax policies when agents make such optimization errors. The empirical evidence is based on two complementary strategies. First, we conducted an experiment at a grocery store posting tax inclusive prices for 750 products subject to sales tax for a three week period. Scanner data show that this intervention reduced demand for the treated products by 8 percent. Second, we find that state-level increases in excise taxes (which are included in posted prices) reduce alcohol consumption significantly more than increases in sales taxes (which are added at the register and are hence less salient). We develop simple, empirically implementable formulas for the incidence and efficiency costs of taxation that account for salience effects as well as other optimization errors. Contrary to conventional wisdom, the formulas imply that the economic incidence of a tax depends on its statutory incidence and that a tax can create deadweight loss even if it induces no change in demand. Our approach to welfare analysis yields robust results because it does not require specification of a positive theory for why agents fail to optimize with respect to tax policies.

*We are very grateful to Sofia Berto Villas-Boas and Reed Johnson for help in implementing the experiment and to Christopher Carpenter, Jeffrey Miron, and Lina Tetelbaum for sharing data on alcohol regulations. Thanks to George Akerlof, David Ahn, Alan Auerbach, Douglas Bernheim, Kitt Carpenter, Judith Chevalier, Stefano DellaVigna, Amy Finkelstein, Michael Greenstone, Caroline Hoxby, Shachar Kariv, Peter Katuscak, Botond Koszegi, Erzo Luttmer, James Poterba, Matthew Rabin, Ricardo Reis, Emmanuel Saez, Jesse Shapiro, Andrei Shleifer, Dan Sichel, Uri Simonsohn, anonymous referees, and numerous seminar participants for helpful comments and discussions. Laurel Beck, Gregory Bruich, Matt Grandy, Matt Levy, Ankur Patel, Ity Shurtz, James Sly, and Philippe Wingender provided outstanding research assistance. Funding was provided by NSF grant SES 0452605. The analysis and conclusions set forth are those of the authors and do not indicate concurrence by other members of the research staff or the Board of Governors of the Federal Reserve. E-mail: chetty@econ.berkeley.edu, adam.looney@frb.gov, kroft@econ.berkeley.edu.

A central assumption in public economics is that agents optimize fully with respect to tax policies. For example, Frank P. Ramsey's (1927) seminal analysis of optimal commodity taxation assumes that agents respond to tax changes in the same way as price changes. Canonical results on tax incidence, efficiency costs, and optimal income taxation (e.g. Arnold C. Harberger 1964, James A. Mirrlees 1971, Anthony B. Atkinson and Joseph E. Stiglitz 1976) all rely on full optimization with respect to taxes.

Contrary to the full optimization assumption, there is accumulating evidence which suggests that individuals are inattentive to some types of incentives.[1] Inattention and imperfect optimization could be particularly important in the case of taxation because tax systems are complex and nontransparent in practice. Income tax schedules are typically highly nonlinear, benefit-tax linkages for social insurance programs are opaque (e.g. social security taxes and benefits), and taxes on commodities are often not displayed in posted prices (sales taxes, hotel city taxes, vehicle excise fees).

In the first half of this paper, we investigate empirically whether individuals optimize fully with respect to taxes by analyzing the effect of "salience" on behavioral responses to commodity taxation. Specifically, we show that commodity taxes that are included in the posted prices that consumers see when shopping (and are thus more salient) have larger effects on demand.[2] In Xavier Gabaix and David I. Laibson's (2006) terminology, our empirical analysis shows that some types of taxes are "shrouded attributes." In the second half of the paper, we develop a simple method of characterizing the welfare consequences of taxation when agents optimize imperfectly with respect to taxes.

We study the importance of salience empirically using two complementary strategies: (1) an experiment in a grocery store and (2) an observational study of the effect of alcohol

[1] Recent studies of inattentive behavior include prices vs. shipping fees (Tanjim Hossain and John Morgan 2005, 2008), financial markets (Stefano DellaVigna and Joshua Pollet 2008), and rebates for car purchases (Meghan Busse, Jorge Silva-Risso, and Florian Zettelmeyer 2006). See DellaVigna (2007) for additional examples and a review of this literature.

[2] We use "tax salience" to refer to the visibility of the tax-inclusive price. When taxes are included in the posted price, the total tax-inclusive price is more visible but the tax rate itself may be less clear. There is a longstanding theoretical literature on "fiscal illusion" which discusses how the lack of visibility of tax rates may affect voting behavior and the size of government (John S. Mill 1848). Unlike that literature, we define salience in terms of the visibility of the tax-inclusive price because we focus on behaviors that optimally depend on total tax-inclusive prices rather than behaviors which depend on the tax rate itself.

taxes on alcohol consumption. The experiment was implemented at a supermarket over a three-week period in early 2006. In this store, prices posted on the shelf exclude sales tax of 7.375 percent. If a product is subject to sales tax, it is added to the bill only at the register, as in most other retail stores in the United States.[3] To test if people under react to the sales tax because it is not included in the posted price, we posted tags showing the tax inclusive price below the original pretax price tags (shown in Exhibit 1). We posted these tags for all products (roughly 750 total) in three taxable groups: cosmetics, hair care accessories, and deodorants. A preliminary survey-based evaluation of the tags indicates that they succeed in reminding consumers of actual tax inclusive prices. Without the tags, nearly all survey respondents ignored taxes when calculating the total price of a basket of goods whereas with the tags, the vast majority computed the total tax inclusive price correctly.

We analyze the effect of posting tax inclusive prices on demand using a differences-in-differences research design. Scanner data show that quantity sold and total revenue in the treated group of products fell by about 8 percent during the intervention relative to two "control groups" – other products in the same aisle of the treatment store that were not tagged and products in two other stores in nearby cities. The null hypothesis that posting tax inclusive prices has no effect on demand is rejected using both t-tests and nonparametric permutation tests. To interpret the magnitude of the treatment effect, we compare it with the price elasticity of demand for these categories, which is in the range of 1 to 1.5. Since showing the tax inclusive price reduced demand by nearly the same amount as a 7.375 percent price increase, we infer that most consumers do not normally take the sales tax into account.

A concern with the experiment is that posting 750 new tags may have reduced demand because of a "Hawthorne effect" or a short-run violation of norms. This issue motivates our second empirical strategy, which compares the effect of price changes with tax changes using observational data over a longer horizon. To implement this test, we focus on alcohol consumption because alcohol is subject to two state-level taxes in the U.S.: an excise tax

[3]The sales tax affects relative prices because it does not apply to all goods. Approximately 40 percent of expenditure is subject to sales tax in the United States. Since food is typically exempt, the fraction of items subject to sales tax in grocery stores is much lower.

that is included in the posted price and a sales tax that is added at the register (and hence is less salient). Exploiting state-level changes in these two tax rates between 1970 and 2003 coupled with annual data on total beer consumption by state, we find that increases in the excise tax reduce beer consumption by an order of magnitude more than similar increases in the sales tax. A simple calibration shows that the magnitude of the difference in the elasticity estimates cannot be explained purely by the fact that the sales tax applies to a broader base, especially since food and nonalcoholic beverages are exempt from sales tax in most states. The difference in elasticities persists over time, indicating that behavioral responses to taxes and prices differ even in the long run.

Why do consumers under react to taxes that are not included in posted prices? One explanation is that customers are uninformed about the sales tax rate or which goods are subject to sales tax. An alternative hypothesis is that salience matters: the customers know what is taxed, but focus on the posted price when shopping. To distinguish between these hypotheses, we surveyed grocery shoppers about their knowledge of sales taxes. The median individual correctly reported the tax status of 7 out of the 8 products on the survey, indicating that our empirical findings are driven by salience effects. A key feature of salience is that it matters in steady state, and not just on the transition path after tax changes.

Motivated by this evidence, we analyze the implications of salience effects and other optimization errors for the welfare consequences of taxation. The challenge in this analysis, as in behavioral public economics more generally, is the calculation of welfare when behavior is inconsistent with full optimization. We characterize the effect of taxes on social surplus (excess burden) and distribution (incidence) using an approach that does not rely on a specific positive model of behavior, as in B. Douglas Bernheim and Antonio Rangel (2007). Our approach relies on two assumptions: (1) the tax affects welfare only through its effect on the consumption bundle chosen by the agent and (2) consumption choices when prices are fully salient are consistent with full optimization. Under these assumptions, we derive approximate formulas for the incidence and efficiency costs of taxation that depend only on the empirically observed demand function and not on the underlying model which generates that demand function. Intuitively, there are two demand curves that together are sufficient statistics for welfare calculations when individuals make optimization errors: the *tax-demand*

curve, which tells us how demand varies with taxes that are not included in posted prices, and the *price-demand* curve, which tells us how demand varies as (fully salient) posted prices change. We use the tax-demand curve to determine the effect of the tax on behavior and then use the price-demand curve to calculate the effect of that change in behavior on welfare. The price-demand curve can be used to recover the agent's underlying preferences and calculate welfare because it is generated by optimizing behavior.

The benefits of this approach to welfare analysis are its simplicity and adaptability. The formulas for excess burden and incidence can be derived using supply and demand diagrams and familiar notions of consumer and producer surplus. The formulas differ from the standard Harberger (1964) expressions by a single factor – the ratio of the compensated tax elasticity to the compensated price elasticity. Thus, one can calculate the (partial equilibrium) deadweight cost and incidence of any tax policy by estimating both the tax and price elasticities instead of just the tax elasticity as in the existing empirical literature. Although we motivate our welfare analysis by evidence of salience effects, the formulas account for *all* errors that consumers may make when optimizing with respect to taxes.[4] For example, confusion between average and marginal income tax rates (Charles de Bartolome (1995), Jeffrey B. Liebman and Richard J. Zeckhauser (2004), Naomi E. Feldman and Peter Katuščák (2006)) or over perception of estate tax rates (Robert J. Blendon et al. (2003), Joel B. Slemrod (2006)) can be handled using exactly the same formulas, without requiring knowledge of individuals' tax perceptions and information set.

Our welfare results challenge widely held intuitions based on the full optimization model. First, the agent who bears the statutory incidence of a tax bears more of the economic incidence, violating the classic tax neutrality result in competitive markets. Second, a tax increase can have a substantial efficiency cost even when demand for the taxed good does not change by distorting consumption allocations for inattentive individuals. Finally, holding fixed the tax elasticity of demand, we show that an increase in the price elasticity of demand *reduces* deadweight loss and *increases* incidence on consumers.

[4] Our formulas do not, however, permit errors in optimization relative to salient prices. Such errors can be accommodated by isolating a condition where the true price elasticity is revealed and applying the formulas here.

This paper builds on and relates to several recent papers in public economics. Our theoretical analysis can be viewed as an application of Bernheim and Rangel's (2007) choice-based approach to welfare, where the choices when taxes are salient reveal an agent's true rankings (see section IV for more details). Our analysis also relates to the recent work of Liebman and Zeckhauser (2004), who analyze optimal income taxation in a model where individuals misperceive tax schedules because of "ironing" or "spotlighting" behavior. Our approach does not require assumptions about whether individuals iron, spotlight, or respond in some other way to the tax schedule, as any of these behaviors are captured in the empirically observed tax and wage elasticities of labor supply.

Our empirical results are consistent with those of Amy N. Finkelstein (2007) and Tomer Blumkin, Bradley J. Ruffle, and Yosi Ganun (2008), who find evidence of salience effects in toll collection and a lab experiment on consumption vs. income taxes. One notable study that does not find that salience matters is Harvey S. Rosen (1976), who shows that the cross-sectional correlation between marginal tax rates and work hours is similar to the correlation between wage rates and work hours. The cross-sectional approach to estimation of wage elasticities has since been shown to suffer from serious identification problems, which could explain why our use of credibly exogenous variation to identify salience effects yields different results.

The remainder of the paper is organized as follows. Section I presents an organizing framework for our empirical analysis. Section II discusses the grocery experiment. Section III presents the evidence on alcohol sales. Section IV explores why consumers under react to taxes. Section V presents the theoretical welfare analysis. Section VI concludes.

# I Empirical Framework

To motivate the empirical analysis, consider consumer behavior in an economy with two goods, $x$ and $y$, that are supplied perfectly elastically.[5] Normalize the price of $y$ to 1 and let $p$ denote the pretax price of $x$. Assume that $y$ is untaxed and $x$ is subject to an ad

[5]The supply curve is effectively flat in both of our empirical strategies. In the grocery experiment, stocking patterns and prices are set at a regional level and are exogenous to our small intervention. In the alcohol analysis, we show that state-level changes in taxes on producers are shifted fully to consumers.

valorem sales tax $\tau^S$. The total price of $x$ is $q = (1 + \tau^S)p$. The price that consumers see when deciding what to purchase is $p$; the sales tax is not included in the posted price. Since consumers must calculate $q$ themselves but can see $p$ directly, we will say that the tax inclusive price $q$ is less "salient" than the pretax price $p$.

Let $x(p, \tau^S)$ denote demand as a function of the posted price and the ad-valorem sales tax. In the neoclassical full-optimization model, demand depends only on the total tax inclusive price: $x(p, \tau^S) = x((1 + \tau^S)p, 0)$. If consumers optimize fully, a 1 percent increase in $p$ and a 1 percent increase in the gross-of-tax price $(1 + \tau^S)$ reduce demand by the same amount: $\varepsilon_{x,p} \equiv -\frac{\partial \log x}{\partial \log p} = \varepsilon_{x,1+\tau^S} \equiv -\frac{\partial \log x}{\partial \log(1+\tau^S)}$. We hypothesize that in practice consumers under react to the tax $\tau^S$ because it is less salient: $\varepsilon_{x,p} > \varepsilon_{x,1+\tau^S}$. To test this hypothesis, we log-linearize the demand function $x(p, \tau^S)$ to obtain the following estimating equation:

$$\log x(p, \tau^S) = \alpha + \beta \log p + \theta_\tau \beta \log(1 + \tau^S) \tag{1}$$

In this equation, the parameter $\theta_\tau$ measures the degree to which agents under react to the tax.[6] In particular, $\theta_\tau$ is the ratio of the tax elasticity of demand ($\varepsilon_{x,1+\tau^S} = -\theta_\tau \beta$) to the price elasticity of demand ($\varepsilon_{x,p} = -\beta$):

$$\theta_\tau = \frac{\partial \log x}{\partial \log(1 + \tau^S)} \Big/ \frac{\partial \log x}{\partial \log p} = \frac{\varepsilon_{x,1+\tau^S}}{\varepsilon_{x,p}} \tag{2}$$

The null hypothesis of full optimization implies $\theta_\tau = 1$. We use two empirical strategies to estimate $\theta_\tau$.

*Strategy 1: Manipulate Tax Salience.* Our first empirical strategy is to make the sales tax as salient as the pretax price by posting the tax inclusive price $q$ on the shelf. When tax inclusive prices are posted, consumers presumably optimize relative to the tax inclusive price and set demand to $x((1 + \tau^S)p, 0)$. Hence, the effect of posting the tax inclusive price

[6] The parameter $\theta_\tau$ does not have a structural interpretation because we have not specified an economic model that generates (1). In a companion paper (Raj Chetty, Adam Looney, and Kory Kroft 2007), we develop a bounded rationality model in which agents face heterogeneous cognitive costs of computing tax-inclusive prices. In that model, $\theta$ is the fraction of individuals whose cognitive costs lie below the threshold where it is optimal to compute the tax inclusive price.

on demand is

$$\log x((1+\tau^S)p, 0) - \log x(p, \tau^S) = (1-\theta_\tau)\beta \log(1+\tau^S)$$

Recalling that $\varepsilon_{x,p} = -\beta$, we obtain the following estimator for $\theta_\tau$:

$$(1-\theta_\tau) = -\frac{\log x((1+\tau^S)p, 0) - \log x(p, \tau^S)}{\varepsilon_{x,p} \log(1+\tau^S)}. \tag{3}$$

The right hand side of this equation measures the effect of posting tax inclusive prices on demand divided by the effect of a price increase corresponding to the size of the tax. This ratio measures the degree of misperception of total prices when taxes are not included in posted prices. If all consumers normally take the sales tax into account, posting $q$ should have no effect on demand ($\theta_\tau = 1$), since it is redundant information. If all consumers ignore the sales tax, posting $q$ should reduce demand by $\varepsilon_{x,p}\log(1+\tau^S)$, implying $\theta_\tau = 0$.

*Strategy 2: Manipulate Tax Rate.* An alternative method of estimating $\theta_\tau$ is to exploit independent variation in $\tau^S$ and $p$ to estimate the sales tax elasticity $\varepsilon_{x,1+\tau^S}$ and the price elasticity $\varepsilon_{x,p}$, as in Rosen (1976). As shown in (2), the ratio of the two elasticities $\varepsilon_{x,1+\tau^S}/\varepsilon_{x,p}$ identifies $\theta_\tau$.

In the next section, we implement strategy 1 using a field experiment at a grocery store. In section III, we implement strategy 2 using observational data on alcohol consumption.

# II Evidence from an Experiment at a Grocery Store

## II.A Research Design

We conducted an experiment posting tax inclusive prices at one store of a national grocery chain. The store is a 37,000 square foot supermarket with annual revenue of approximately $25 million and is located in a middle-income suburb in Northern California. Approximately 30 percent of the products sold in the store are subject to the local sales tax of 7.375 percent, which is added at the register. Price tags on the shelves display only pretax prices, as in the upper half of the tag shown in Exhibit 1.

The grocery chain's managers expected that posting tax inclusive prices would reduce sales.[7] To limit revenue losses, they asked us to restrict our intervention to three product groups that were not "sales leaders" and to limit the duration of the intervention to three weeks. We looked for three product groups that met this requirement as well as two additional criteria: (1) products with relatively high prices so that the dollar amount of the sales tax is nontrivial; and (2) products belonging to "impulse purchase categories" – goods that exhibit high price elasticities – so that the demand response to the intervention would be detectable. We chose three groups of taxable toiletries: cosmetics, hair care accessories, and deodorants. These three product groups take up half an aisle of the store and together include about 750 distinct products.

We posted tax inclusive prices for all products in the three groups beginning on February 22, 2006 and ending on March 15, 2006.[8] Exhibit 1 shows how the price tags were altered. The original tags, which show pretax prices, were left untouched on the shelf. A tag showing the tax inclusive price was attached directly below this tag for each product. The added tag stated "Total Price: $\$p$ + Sales Tax = $\$q$," where $p$ denotes the pretax price (repeating the information in the original tag) and $q$ denotes the tax inclusive price. The original pretax price was repeated on the new tag to avoid giving the impression that the price of the product had increased. For the same reason, the fonts used for $p$, $q$, and the words "Sales Tax" exactly matched the font used by the store for the original price. Additional details on experiment implementation are given in Appendix A.

*Evaluation of Tags.* To determine whether the tags are effective in increasing tax salience and are understood by consumers, we conducted a preliminary survey-based evaluation in an undergraduate class. We showed the students a photograph of taxable products on the shelf at the grocery store similar to that in Exhibit 1. We distributed surveys (shown in Appendix Exhibit 1) asking each student to choose two goods and write down "the total bill

[7] We estimate that the loss in revenue due to our experiment was about $300 (8% of $3900). Extrapolating from this estimate, if taxes were included in posted prices for all taxable products, the revenue loss would be 2.4 percent, or $600,000 per year per store. Note that this calculation ignores general equilibrium effects that would arise if all retailers were required to post tax-inclusive prices.

[8] The treatment of showing tax inclusive price tags could have been randomized at the individual product level. However, the concern that such an intervention could be deceptive (e.g. suggesting that one lipstick is taxed and another is not) dissuaded us from pursuing this strategy. We therefore tagged complete product groups, so that any direct substitute for a treated product would also be treated.

due at the register for these two items." We first showed the photograph with the regular tags displaying only the pretax prices. After collecting the survey responses, we showed a second photograph of products with our tax inclusive price tags and asked students to repeat the exercise. The results are summarized in the first panel of Table 1. When presented with the first photo, the modal response was the total pretax bill for the two products. Only 18 percent of students reported a total price within \$0.25 of the total tax inclusive amount. When presented with the second photo, the modal response included the sales tax, and 75 percent wrote down an amount within \$0.25 of the true tax inclusive total. This evidence shows that posting tax inclusive price tags does indeed have a strong "first stage" effect on tax salience. Moreover, the results allay concerns that the tags confused consumers into believing that these items were subject to an additional tax or that the pretax price of the product had been increased.

Although we are confident that the tags increased tax salience substantially, we cannot rule out the possibility that they also affected demand through other channels or "Hawthorne effects." For instance, the very fact that 750 new tags were posted on the shelves could have deterred customers from the aisle. We are only able to estimate the effect of posting the tags on demand, and have no means of decomposing the effect of the intervention into the various mechanisms through which the tags may have had an effect. The large first-stage effect of the tags on perceived prices leads us to believe that the primary mechanism is increased tax salience, but we ultimately rely on evidence from the second empirical approach (see section III) to address such concerns.

*Empirical Strategy.* To estimate the effect of our intervention on demand, we compare changes in quantity sold in the "treatment" group of products whose tags were modified with two "control" groups. We define the treatment group as products that belong to the cosmetics, hair care accessories, or deodorants product groups in the store where we conducted the experiment. The first control group is a set of products in the same aisles as the treatment products for which we did not change tags within the experimental store. These products include similar (taxable) toiletries such as toothpaste, skin care, and shaving products; see Table 2 in Appendix A for the full list. The second control group consists of all the toiletry products sold in a pair of stores in nearby cities. These control stores

were selected to match the treatment store prior to the experiment on the demographic and store characteristics shown in Table 2 in Appendix A. Using these two control groups, we implement a standard difference-in-difference methodology to test whether sales of the treated products fell during the intervention relative to the controls.

## II.B Data and Summary Statistics

We use scanner data from the treatment store and the two control stores, spanning week 1 of 2005 to week 15 of 2006. The dataset contains weekly information on price and quantity sold for all toiletry (treatment and control) products in each store. See Appendix A for details on the dataset.

Within the treatment group, there are 13 product "categories" (e.g. lipsticks, eye cosmetics, roll-on deodorants, body spray deodorants). The control product group contains 95 categories, which are listed in Appendix Table 1. We analyze the data at the category level (summing quantity sold and revenue over the individual products within categories) rather than the product level for two reasons. First, the intervention was done at the category level. Second, we cannot distinguish products that were on the shelf but did not sell (true zeros) from products that were not on the shelf. Analyzing the data at the category level circumvents this problem because there are relatively few category-weeks with missing data (4.7 percent of all observations). Since all the categories always existed in all stores throughout the sample period, we believe that these observations are true zeros, and code them as such.

Table 2 presents summary statistics for the treatment and control product groups in each store. The treatment store sold an average of 25 items per category and earned $98 of revenue per week per category over the sample period (column 1 of Table 2). The treatment products thus account for approximately $1,300 of revenue per week as a whole. Average weekly quantity sold per category is similar for the control products in the treatment store, but products in these categories are somewhat more expensive on average (column 2). Sales and revenue for the same categories in the control stores are very similar to those in the treatment store (columns 3-4).

## II.C Results

*Comparison of Means.* We begin the data analysis with a cross-tabulation of mean quantity sold in Table 3. The upper panel of the table shows data for the treatment store. The data is split into four cells. The rows split the data by time (pre experiment vs. the intervention period), while the columns split the data by product group (treated vs. control categories). Each cell shows the mean quantity sold for the group labelled on the axes, along with the standard error and the number of observations. All standard errors reported in this and subsequent tables in this section are clustered by week to adjust for correlation of errors across products.[9]

The mean quantity sold in the treatment categories fell by an average of 1.30 units per week during the experimental period relative to the pre period baseline. Meanwhile, quantity sold in the control categories within the treatment store went up by 0.84 units. Hence, sales fell in the treatment categories relative to the control categories by 2.14 units on average, with a standard error of 0.68. This change of $DD_{TS} = -2.14$ units is the "within treatment store" difference-in-difference estimate of the impact of posting tax inclusive prices. The identification assumption necessary for consistency of $DD_{TS}$ is the standard "common trends" condition (Bruce D. Meyer 1995), which in this case requires that sales of the treatment and control products would have evolved similarly absent our intervention.

One natural way of evaluating the validity of this identification assumption is to compare the change in sales of treatment and control products in the control stores, where no intervention took place. The lower panel of Table 3 presents such a comparison. In the control stores, sales of treatment products increased by a (statistically insignificant) $DD_{CS} = 0.06$ units relative to sales of control products. The fact that $DD_{CS}$ is not significantly different from zero suggests that sales of the treatment and control products would in fact have evolved similarly in the treatment store had the intervention not taken place.

Putting together the upper and lower panels of Table 3, one can construct a "triple difference" (DDD) estimate of the effect of the intervention, as in Jonathan Gruber (1994). This estimate is $DDD = DD_{TS} - DD_{CS} = -2.20$. This estimate is statistically significant

[9] Standard errors are similar when we cluster by category to adjust for serial correlation.

with $p < 0.01$, rejecting the null of full-optimization ($\theta_\tau = 1$). Note that both within-store and within-product time trends are differenced out in the DDD. The DDD estimate is therefore immune to both store-specific shocks – such as a transitory increase in customer traffic – and product-specific shocks – such as fluctuations in demand for certain goods. Hence, the identification assumption for consistency of the DDD estimate is that there was no shock during our experimental intervention that differentially affected sales of only the treatment products in the treatment store. In view of the planned, exogenous nature of the intervention, we believe that this condition is likely to be satisfied.

To gauge the magnitude of the estimated effect, we use the framework in section I. The mean quantity sold per category fell by 2.2 units per week, relative to a base of 29 units sold per week. Making the sales tax salient thus reduces demand by 7.6 percent. We show below that the estimated price elasticity of demand at the category level is $\varepsilon_{x,p} = 1.59$. Given the sales tax rate of 7.375 percent, plugging these values into (3) yields a point estimate of $\theta_\tau = 1 - \frac{7.6}{1.59 \times 7.375} = 0.35$. That is, a 10 percent tax increase reduces demand by the same amount as a 3.5 percent price increase.

*Regression Estimates.* We evaluate the robustness of the DDD estimate by estimating a series of regression models with various covariate sets and sample specifications in Table 4. Let the outcome of interest (e.g. quantity, log quantity, revenue) be denoted by $Y$. Let the variables $TS$ (treatment store), $TC$ (treatment categories), and $TT$ (treatment time) denote indicators for whether an observation is in the experimental store, categories, and time, respectively. Let $X$ denote a vector of additional covariates. We estimate variants of the following linear model, which generalizes the DDD method above (see Gruber 1994):

$$\begin{aligned} Y &= \alpha + \beta_1 TT + \beta_2 TS + \beta_3 TC + \gamma_1 TT \times TC + \gamma_2 TT \times TS + \gamma_3 TS \times TC \\ &\quad + \delta TT \times TC \times TS + \xi X + \varepsilon \end{aligned} \tag{4}$$

In this specification, the third-level interaction ($\delta$) captures the treatment effect of the experiment and equals the DDD estimate when no additional covariates are included.

Specification 1 of Table 4 estimates (4) for quantity sold, controlling for the mean price of the products in each category using a quadratic specification and including category, week,

and store fixed effects.[10] The estimated effect of the treatment is essentially the same as in the comparison of means, which is not surprising since there were no unusual price changes during our intervention period. Specification 2 shows that the intervention led to a significant reduction in revenue (price×quantity) from the treated products.[11]

In specification 3, we estimate an analogous model in logs instead of levels. In this specification, we weight each observation by the mean revenue over time by category by store, placing greater weight on the larger categories as in the levels regressions. The log specification is perhaps a better model for comparisons across categories with different baseline quantities, but it forces us to omit observations that have zero quantity sold. It yields a slightly larger estimate than the levels model for the reduction in quantity sold (10.1 percent). The estimated category-level price elasticity – the effect of a 1 percent increase in the prices of *all* goods within a category – is $\varepsilon_{x,p} = 1.59$. This elasticity is identified by the variation in average category-level prices across weeks within the stores. The estimate is consistent with those of Stephen J. Hoch et al. (1995), who estimate a full product-level demand system and obtain category-level price elasticities of 1 to 1.5 for similar products using scanner data from the same grocery chain.

*Placebo and Permutation Tests.* To further evaluate the "common trends" identification assumption, we check for unusual patterns in demand immediately before and after the experiment. We replicate specification 1 including indicator variables for the three week periods before the intervention began ($BT$) and after the intervention ended ($AT$). We also include second- and third-level interactions of $BT$ and $AT$ with the $TC$ and $TS$ variables, as for the $TT$ variable in (4). Column 4 of Table 4 reports estimates of the third-level interactions for the periods before, during, and after the experiment. Consistent with the

[10] The mean price is defined as the average price of the products in each category in the relevant week, weighted by quantity sold over the sample period. The fixed weights eliminate any mechanical relationship between fluctuations in quantity sold and the average price variable.

[11] Studies in the marketing literature (e.g., Eric T. Anderson and Duncan I. Simester 2003) find that demand drops discontinuously when prices cross integer thresholds (such as \$3.99 vs. \$4.01), and that retailers respond by setting prices that end in '9' to maximize profits. Indeed, the retailer we study sets most products' *pretax* prices just below the integer threshold – an observation that in itself supports our claim that individuals focus on the pretax price, since the tax inclusive price is often above the integer threshold. We find no evidence that demand fell more for the products whose price crossed the integer threshold once taxes were included (e.g. \$3.99 + Sales Tax = \$4.28), but the difference in the treatment effects is imprecisely estimated.

other results in Table 4, quantity sold in the treatment group is estimated to change by $\delta = -2.27$ units during the intervention. The corresponding "placebo" estimates for the periods before and after the treatment are close to zero, indicating that the fall in demand coincides precisely with the intervention period.

A concern in DD analysis is that serial correlation can bias standard errors, leading to overrejection of the null hypothesis of no effect (Marianne Bertrand, Esther Duflo, and Sendhil Mullainathan 2002). To address this concern, we implement a nonparametric permutation test for $\delta = 0$. We first choose a "placebo triplet" consisting of a store, three week time period, and a randomly selected set of 13 product categories. We then estimate (4), pretending that the placebo triplet is the treatment triplet. We repeat this procedure for all permutations of stores and contiguous three week periods and 25 different randomly selected groups of 13 categories, obtaining $63 \times 3 \times 25 = 4,725$ placebo estimates. Defining $G(\widehat{\delta}_P)$ to be the empirical cdf of these placebo effects, the statistic $G(\delta)$ gives a p-value for the hypothesis that $\delta = 0$. Intuitively, if the experiment had a significant effect on demand, we would expect the estimated coefficient to be in the lower tail of estimated placebo effects.[12] Since this test does not make parametric assumptions about the error structure, it does not suffer from the overrejection bias of the t-test.

Figure 1 illustrates the results of the permutation test by plotting the empirical distribution of placebo effects $G$ for log quantity (specification 3 of Table 4). The vertical line in the figure denotes the treatment effect reported in Table 4. For log quantity, $G(\delta) = 0.07$. An analogous test for log revenue yields $G(\delta) = 0.04$. Although these p-values are larger than those obtained using the t-tests, they confirm that the intervention led to an unusually low level of demand.

Finally, we consider subsets of the large set of counterfactuals across time, categories, and stores. In column 5 of Table 4, we restrict the sample to the treatment product categories and compare across time and stores. This DD estimate is quite similar to the DDD estimates. Results are also similar when we restrict the sample to the treatment store or limit the pre experiment sample to the three months immediately before the intervention.

[12] This test is an extension of Ronald A. Fisher's (1922) "exact test" for an association between two binary variables. See Paul R. Rosenbaum (1996) for more on permutation tests.

# III Evidence from Observational Data on Alcohol Sales

## III.A Research Design

We turn now to the second empirical test: comparing the effect of increases in posted prices and taxes on demand. We implement this strategy by focusing on alcohol consumption. Alcohol is subject to two taxes in most states: (1) an *excise* tax that is levied at the wholesale level and is included in the price posted on the shelf or restaurant menu and (2) a *sales* tax, which is added at the register (except in Hawaii, which we exclude). The total price of alcohol is therefore $q = p(1+\tau^E)(1+\tau^S)$ where $p$ is the pretax price, $\tau^E$ is the excise tax, and $\tau^S$ is the sales tax. Since the excise tax is included in the posted price, it is more salient than the sales tax.

We estimate the effect of $\tau^E$ and $\tau^S$ on alcohol consumption by exploiting the many state-level changes in the two taxes between 1970 and 2003. Our estimating equation is based on the demand specification in (1):

$$\log x = \alpha + \beta \log(1+\tau^E) + \theta_\tau \beta \log(1+\tau^S). \tag{5}$$

We estimate (5) in first-differences because both the tax rates and alcohol consumption are highly autocorrelated series. Letting $t$ index time (years) and $j$ index states, define the difference operator $\Delta z_{jt} = z_{jt} - z_{j,t-1}$. Introducing a set of other demand-shifters $X_{jt}$ and an error term $\varepsilon_{jt}$ to capture idiosyncratic state-specific demand shocks, we obtain the following estimating equation by first-differencing (5):

$$\Delta \log \widehat{x}_{jt} = \alpha' + \beta \Delta \log(1+\tau^E_{jt}) + \theta_\tau \beta \Delta \log(1+\tau^S_{jt}) + X_{jt}\rho + \varepsilon_{jt} \tag{6}$$

We estimate variants of (6) using OLS and test the hypothesis that $\theta_\tau = 1$.[13] The identification assumption is that the changes in sales and excise taxes are uncorrelated with state-specific shocks to alcohol consumption.

[13] The full-optimization model predicts $\theta_\tau = 1$ irrespective of the incidence of the taxes. If tax increases are passed through fully to the consumer – which appears to be the case in practice as we show below – $\beta$ equals the price elasticity of demand.

## III.B Data and Summary Statistics

Tax rates on alcohol vary across beer, wine, and spirits. In the interest of space, we present results for beer, which accounts for the largest share of alcohol consumption in the U.S. A parallel analysis using data on wine taxes and consumption for a smaller subset of years yields similar results. We use data on aggregate annual beer consumption by state from the National Institute of Alcohol Abuse and Alcoholism (NIAAA) (2006) from 1970-2003. These data are compiled from administrative state tax records, and are more precise than data from surveys because they reflect total consumption in each state.

We obtain data on beer excise tax and sales tax rates and revenues by state from the Brewer's Almanac (various years), World Tax Database (2006), and other sources.[14] The state sales tax is an ad valorem tax (proportional to price), while the excise tax is typically a specific tax (dollars per gallon of beer). We convert the excise tax rate into percentage units comparable to the sales tax by dividing the beer excise tax per case in year 2000 dollars by the average cost of a case of beer in the United States in the year 2000.[15] We normalize the excise tax by the average *national* price because each state's price is endogenous to its tax rate. Details on the data sources and construction of tax rates are given in Appendix A.

Table 5 shows summary statistics for the pooled dataset. Between 1970 and 2003, mean per capita consumption of beer is roughly 240 cans per year. The average state excise tax rate is 6.4 percent of the average price, while the mean state sales tax rate is 4.3 percent.[16] There is considerable independent variation within states in the two taxes over the sample period. There are 153 legislated changes to the sales tax and 131 legislated changes to excise taxes; the correlation between excise tax changes and sales tax changes is 0.06.

[14] We exclude West Virginia (WV) because of problems with the sales tax rate data described in Appendix A. Including WV magnifies the difference between the excise and sales tax elasticities.

[15] Real growth in the price of beer could lead to mismeasurement of beer prices and excise tax rates early in the sample. Using a subset of the data for which we have information on beer prices from the American Chamber of Commerce Researchers Association (ACRRA) cost-of-living survey, we find that beer price growth closely tracks changes in the CPI. Moreover, we show below that instrumenting for the actual ACCRA price in each state/year for which it is available using our construction of the excise tax rate yields similar results.

[16] Some cities also levy local sales taxes on top of the state sales tax. In Chetty, Looney, and Kroft (2007), we show that including local sales taxes by imputing them from data on local tax revenues does not affect the results.

## III.C Results

We begin with a graphical analysis to illustrate the relationship between alcohol consumption and taxes in Figures 2a and 2b. These figures plot annual state-level changes in log beer consumption per capita against log changes in the gross-of-excise-tax price $\Delta \log(1+\tau^E)$ and the gross-of-sales-tax price $\Delta \log(1+\tau^S)$. To construct Figure 2a, we first round each state excise tax change to the nearest tenth of a percent. We then compute the mean change in log beer consumption for observations with the same rounded excise tax change. Finally, we plot the mean consumption change against the rounded excise tax rates, superimposing a best-fit line on the points as a visual aid. Figure 2b is constructed analogously using sales tax changes. To make the range of changes in the excise tax comparable to the smaller range of changes in the sales tax, we restrict the range of both tax changes to $\pm.02$ log points. Figure 2a shows that increases in the beer excise tax sharply reduce beer consumption. Figure 2b shows that increases in the sales tax have a much smaller effect on beer consumption.

*Regression Estimates.* Table 6 presents estimates of the model for the state-level growth rate of alcohol consumption in (6). In these and all subsequent specifications, we adjust for potential serial correlation in errors by clustering the standard errors by state. Column 1 reports estimates of a baseline model that includes only year fixed effects (which remove aggregate trends) and log state population growth as covariates. In this specification, a 1 percent increase in the gross-of-excise-tax price is estimated to reduce beer consumption by 0.87 percent ($\varepsilon_{x,1+\tau^E} = 0.87$).[17] In contrast, a 1 percent increase in the gross-of-sales-tax price is estimated to reduce beer consumption by 0.20 percent ($\varepsilon_{x,1+\tau^S} = 0.20$). The null hypothesis that the excise and sales tax elasticities are equal is rejected with $p = 0.05$.

Columns 2-4 control for factors that may be correlated with the tax changes. One concern is that sales tax changes are correlated with the business cycle. In column 2, we control for the state-level business cycle by including changes in log state per capita income and the state unemployment rate as covariates. Introducing these controls reduces the estimated sales tax coefficient, and as a result the null hypothesis of equal elasticities is rejected with

[17] This elasticity estimate is consistent with estimates of the elasticity of beer consumption with respect to the excise tax rate ($\varepsilon_{x,\tau^E}$) reported in previous studies. For example, Philip J. Cook, Jan Ostermann, and Frank A. Sloan (2005) estimate that a \$0.01 increase in the beer tax per ounce of ethanol reduces beer consumption by 1.9 percent, which translates to $\varepsilon_{x,1+\tau^E} = 1.26$ at the sample mean.

$p = 0.01$. The sales tax effect is smaller because sales taxes are sometimes raised during budgetary shortfalls that occur in recessions. Since alcohol is a normal good (as indicated by the coefficients on per capita income and unemployment rate), failing to control for the business cycle biases the correlation between alcohol consumption and sales tax changes upward in magnitude. Hence, the endogeneity of sales tax rate appears to work *against* rejecting the null hypothesis that $\varepsilon_{x,1+\tau^E} = \varepsilon_{x,1+\tau^S}$.[18]

Another concern is that excise tax increases are sometimes associated with contemporaneous tightening of alcohol regulations. We evaluate this concern using data on four regulations: the legal drinking age, the blood alcohol content limit, implementation of stricter drunk driving regulations for youths, and introduction of administrative license revocation laws. We control for the change in the legal drinking age (in years) and separate indicator variables for a shift toward stricter regulations in each of the other three measures in column 3. The coefficient on the excise tax rate does not change significantly because regulation changes have modest effects on total beer consumption; on average, beer consumption falls by only 0.5 percent when one of the four regulations is tightened.

A third concern is that trends in excise tax rates may be correlated with changes in social norms, which directly influence alcohol consumption. For example, rising acceptance of alcohol consumption in historically conservative regions such as the South may have led to both a reduction in the excise tax as a percentage of price and an increase in alcohol consumption. To assess whether such trends lead to significant bias, we include region fixed effects in column 4 of Table 6, effectively identifying the model from changes in taxes in geographically adjacent states. The coefficient on the excise rate remains substantially larger than the coefficient on the sales tax, suggesting that our results are not spuriously generated by region-specific trends.

There are two sources of variation identifying the excise tax coefficient: (1) policy changes in the nominal tax rate, which produce sharp jumps in tax rates and (2) gradual erosion of the nominal value of the tax by inflation, which creates differential changes in excise tax

[18] Changes in excise taxes are not correlated with the business cycle. A more plausible source of endogeneity is that policymakers raise alcohol excise taxes when alcohol consumption is rising. This would also work against finding a difference in the elasticities, as the estimate of $\varepsilon_{x,1+\tau^E}$ will be biased downward.

rates across states because they have different initial tax rates.[19] To test whether the two sources of variation yield similar results, we isolate the effect of the policy changes using an instrumental variables strategy. We fix the price of beer at its sample average and compute the implied ad valorem excise tax as the nominal tax divided by this *time-invariant* price. The only variation in this simulated tax rate is due to policy changes. Using the simulated excise tax rate to instrument for the actual excise tax rate, we replicate the specification in column 3 of Table 6. The point estimates of both tax elasticities, reported in column 1 of Table 7, are similar to those in previous specifications. The standard errors rise as expected since part of the variation in excise tax rates has been excluded.

Thus far, we have focused on changes in tax rates and alcohol consumption at an annual frequency. One explanation of the difference between the sales and excise tax effects at the annual frequency is learning: people might immediately perceive excise taxes, but learn about changes in the sales tax over time. To test for such learning effects, we estimated specifications including lags and leads of the tax variables and differences over longer horizons. For example, Column 2 of Table 7 shows the effect of sales and excise tax changes on consumption over a three-year horizon (as in Jonathan Gruber and Emmanuel Saez (2002)). An increase in the excise tax rate continues to have a large negative effect on alcohol consumption after three years, whereas an equivalent increase in the sales tax still does not. This evidence suggests that consumers under react to taxes that are not salient even in the long run.

The premise underlying our analysis is that firms pass the excise tax through to consumers so that changes in the excise tax are reflected in the posted price of beer. We check this mechanism using data on posted prices of beer from the ACCRA cost of living survey from 1982 to 2000. Using these data, we estimate the price elasticity of demand for beer, instrumenting for changes in the posted price using changes in the excise tax rate. The estimated price elasticity of demand, reported in column 3 of Table 7, is $\varepsilon_{x,p} = 0.88$, almost identical to the estimates of $\varepsilon_{x,1+\tau^E}$ in the previous specifications. The standard error

[19] To clarify why inflation generates identifying variation, consider the following example. Suppose the pretax price of beer is \$1 and that state $A$ has a nominal alcohol tax of 50 cents, while state $B$ has no excise tax. If prices of all goods double, the gross-of-tax price of beer relative to other goods falls by $\frac{1.50-1.25}{1.50} = 17$ percent in state $A$ but is unchanged in state $B$.

rises because we have price data for only 55 percent of the observations. The reason that $\varepsilon_{x,p} = \varepsilon_{x,1+\tau^E}$ is that state-level excise tax increases are fully passed through to consumers – the coefficient on the excise tax variable in the first-stage regression is approximately 1. This finding supports the claim that the excise tax has a larger effect on demand than the sales tax because it is fully salient.

*Relative Price Changes and Excise vs. Sales Taxes.* An important concern in the comparison of sales and excise tax effects is that the sales tax applies to a broader set of goods than alcohol. Approximately 40 percent of consumption is subject to sales taxation.[20] A 1 percent increase in $\tau^S$ changes the relative price of alcohol and all other goods less than a 1 percent increase in $\tau^E$, which could potentially explain why the sales tax effect is smaller than the excise tax effect even absent salience effects.

We evaluate the magnitude of the bias due to this problem in two ways. First, we estimate the model using only the thirty states that fully exempted all food items from the sales tax in 2000.[21] In these states, changes in the sales tax always affect the relative price of alcohol and food (and nonalcoholic beverages), which is the most plausible substitute for alcohol. Column 4 of Table 7 shows that the sales tax elasticity remains quite small in this subsample.

As an alternative approach, we calibrate the effect of a 1 percent increase in a (hypothetical) tax $\tau^A$ that applies solely to alcohol ($x$) and is excluded from the posted price. Treating all goods other than alcohol as a composite commodity ($y$) of which 40 percent is subject to sales tax, observe that a 1 percent increase in the gross-of-sales-tax price $(1 + \tau^S)$ increases the price of $x$ relative to $y$ by $\frac{1.01}{1.004} - 1 \simeq 0.6$ percent. It follows that the effect of a 1 percent increase in the tax $\tau^A$ that applies solely to alcohol is given by $\varepsilon_{x,1+\tau^A} = \frac{1}{0.6}\varepsilon_{1,1+\tau^S}$. Scaling up the largest estimated response to the sales tax in Table 6 of -0.20 by $\frac{5}{3}$ yields an estimate of $\varepsilon_{x,1+\tau^A} = 0.33$, which remains substantially below the excise tax elasticity estimates.

A related concern is that increases in the beer excise tax may induce substitution to wine and spirits, thereby biasing the beer tax elasticity up relative to the sales tax elasticity. To assess the extent of substitution, we estimate the effect of the beer excise tax on the share

[20] In 2004, sales tax revenues were 2.1% of personal consumption expenditures (PCE). The average (state income-weighted) sales tax rate was 5.3 percent. Hence the tax base is approximately 40 percent of PCE.

[21] We do not have historical data on which goods are subject to the sales tax. However, case studies of some states suggest that the set of items subject to sales tax is fairly stable over time.

of beer in total alcohol (ethanol) consumption. The estimates in column 5 of Table 7 show that the beer share is insensitive to the beer tax rate. The reason is that excise tax rates on beer, wine, and spirits are highly correlated. For example, the correlation coefficient of changes in beer and wine tax rates is 0.94; in 86 percent of the instances in which a state changes its beer excise tax, it also changes its wine excise tax rate. We also find that the effect of beer excise taxes on *total* ethanol consumption is much larger than the effect of sales taxes. We conclude that differences in tax bases are unlikely to explain the substantial gap between the estimated sales and excise tax elasticities.

*Summary.* Averaging across the estimates in Tables 6 and 7, the mean estimate of the gross-of-excise-tax elasticity is 0.84. The mean estimate of the gross-of-sales-tax elasticity is 0.03. Scaling up the sales tax coefficient by $\frac{5}{3}$, we obtain an implied elasticity of 0.05 for a tax $\tau^A$ that is applied solely to alcohol at the register. Combining these estimates yields a point estimate of $\theta_\tau = \varepsilon_{x,1+\tau^A}/\varepsilon_{x,1+\tau^E} = 0.06$.

# IV Why Do Consumers Under React to Taxes?

There are two potential explanations for the finding that consumers under react substantially to taxes that are not included in posted prices. One is that customers are uninformed about sales tax rates. Showing the tax inclusive price tags may have provided new information about tax rates. An alternative explanation is that *salience* matters: individuals know about taxes when their attention is drawn to the subject, but do not pay attention to taxes that are not transparent while deciding what to buy.

A few pieces of auxiliary evidence from the empirical analysis cast doubt on the information hypothesis. First, sales of the (taxable) toiletries adjacent to those that were tagged in the grocery store experiment did not change significantly during the intervention. The fact that posting tax inclusive prices had no "spillover effects" suggests that individuals did not simply learn that toiletries are subject to sales tax. Second, demand returned to pre experiment levels after the intervention ended, suggesting that there were no persistent learning effects.[22] Third, the finding the sales and excise tax elasticities for alcohol demand do not

[22]We cannot rule out another equally plausible explanation of this finding: the set of individuals who

converge over time suggests that the under reaction to the sales tax is not caused by delays in acquiring information.

To distinguish between information and salience more directly, we surveyed 91 customers entering the store where we conducted the experiment about their knowledge of sales taxes. See Appendix A for details on survey implementation and Appendix Exhibit 2 for the survey instrument. We asked individuals the local sales tax rate and whether various products (e.g. milk, magazines, toothpaste) were subject to sales tax. Summary statistics for the survey data are displayed in Panel B of Table 1. 75 percent of those surveyed reported the sales tax rate within 0.5 percentage points of the true rate, and 97 percent reported a rate between 6.75 percent and 8.75 percent. The modal answer was exactly 7.375 percent. The median respondent answered 7 out of 8 questions about taxable status of the goods correctly. The respondents generally believe that food is not taxed, but inedible items and "sin" goods are taxed. Exceptions to this heuristic led to the most errors. In California, carbonated beverages are subject to sales tax, while cookies are not. Among respondents who got one question wrong, Coca Cola and cookies accounted for more than half the mistakes.

In summary, most consumers are well informed about commodity tax rates when their attention is drawn to the subject. However, they do not remember to include the tax when making consumption decisions, as shown by the survey of students discussed in section II. The two surveys and two strands of empirical evidence together indicate that salience and inattention are a central determinant of consumer responses to taxation in steady state.

*Positive Theories.* There are many positive theories that can explain under reaction to taxation. In a companion paper (Chetty, Looney, and Kroft 2007), we propose a bounded-rationality model in which agents pay cognitive costs to calculate tax inclusive prices. We show that small cognitive costs can generate substantial inattention to taxes because the utility loss from ignoring taxes is a second-order function of the tax rate. For example, an agent who spends $x_0 = \$1,000$, has $\varepsilon_{x,p} = 1$, and linear utility in $y$, loses only \$5 by ignoring a 10 percent sales tax. An economy populated by individuals who face small cognitive or time costs of paying attention to taxes can thus generate $\varepsilon_{x,p} >> \varepsilon_{x,1+\tau^S}$.

shop for these durable goods is likely to vary substantially across weeks, so customers in the weeks after the experiment may have been untreated.

More generally, agents with limited attention may use heuristics to achieve a consumption allocation that approximates the fully optimal bundle, but leads them to under react to taxes. For example, consumers may apply a tax rate of 5 percent or 10 percent instead of 7.375, percent or compute 7 percent of $5.00 instead of 7 percent of the exact price $4.95. A more sophisticated heuristic is to keep a separate shadow value of money in mind for taxed and untaxed goods. An entirely different theory of attention is a psychological model in which allocation of attention is triggered by cues (e.g. the visibility or color of pricing information) rather than economic optimization.

Our data does not allow us to distinguish between these models. We therefore proceed to analyze the welfare consequences of tax in a manner that does not depend on a specific positive theory of under reaction to taxes.

# V Welfare Analysis

The objective of this section is to develop partial-equilibrium formulas for incidence and efficiency costs that allow for salience effects. We first characterize tax incidence, which is essentially a mechanical calculation of price changes. We then characterize efficiency costs, which is a more complex problem because additional assumptions are required to calculate welfare changes when agents optimize imperfectly. We restrict attention to tax policies designed to raise revenue (e.g. to finance a public good).[23] The tools developed below can be adapted to analyze Pigouvian taxes intended to correct behavior, but we defer that analysis to future work.

## V.A Setup

Consider an economy with two goods, $x$ and $y$. The government levies two specific (unit) taxes on good $x$: a tax $t^E$ that is included in the posted price and a tax $t^S$ that is not

[23] Our welfare analysis focuses solely on the costs of raising tax revenue, taking the benefits of a given amount of revenue as invariant to the tax system used to generate it. For example, we ignore the possibility that more visible taxes may constrain inefficient spending by politicians (Finkelstein 2007).

included in the posted price.[24] In analogy with the empirical analysis, we will refer to $t^E$ as the excise tax and $t^S$ as the sales tax, but the formulas below can be applied to any tax, including labor and capital income taxes. Let $t = t^E + t^S$ denote the total tax on good $x$. Good $y$, the numeraire, is untaxed. Let $p$ denote the pretax price of $x$ and $q = p + t$ denote the tax inclusive price of $x$. As is standard in partial equilibrium analyses, assume that the tax revenue is not spent on the taxed good (i.e. it is used to buy $y$ or thrown away).

*Consumption.* The representative consumer has wealth $Z$ and has utility $u(x) + v(y)$. In the full-optimization model, the agent chooses a consumption bundle $(x^*(p+t^E, t^S, Z), y^*(p+t^E, t^S, Z))$ that satisfies

$$\begin{aligned} u'(x^*) &= (p+t)v'(y^*) \\ (p+t)x^* + y^* &= Z \end{aligned}$$

This model implies $\frac{\partial x^*}{\partial p} = \frac{\partial x^*}{\partial t}$, contradicting the data. To allow for different responses to prices and taxes, let $x(p + t^E, t^S, Z)$ denote the empirically observed demand for $x$ as a function of the posted price, sales tax, and wealth and $y(p + t^E, t^S, Z)$ the corresponding demand function for $y$.[25] We do not place structure on the positive model that generates $(x(p+t^E, t^S, Z), y(p+t^E, t^S, Z))$ other than to assume that the demand functions are smooth and that the choices are feasible:

$$(p+t)x(p+t^E, t^S, Z) + y(p+t^E, t^S, Z) = Z$$

Because the theory pertains to specific taxes rather than ad-valorem taxes, it is helpful to redefine the degree of under reaction to the tax as

$$\theta = \frac{\partial x(p+t^E, t^S, Z)}{\partial t^S} \Big/ \frac{\partial x(p+t^E, t^S, Z)}{\partial p} = \frac{\varepsilon_{x,q|t^S}}{\varepsilon_{x,q|p}}$$

[24] We analyze specific rather than ad valorem (percentage of price) taxes to simplify the algebra. The incidence and excess burden of introducing an ad valorem tax $\tau^S$ when there are no pre-existing taxes can be calculated by replacing $t^S$ by $\tau^S$ and $\frac{\partial x}{\partial t^S}$ by $\frac{\partial x}{\partial \tau^S}$ in the derivative-based formulas in Propositions 1-3.

[25] These demand functions differ from those estimated in the empirical analysis. With a slight abuse of notation, we switch the second argument of $x(\cdot)$ and $y(\cdot)$ to a specific tax rather than an ad-valorem tax from this point onward.

where $\varepsilon_{x,q|t^S} = -\frac{\partial x}{\partial t^S}\frac{q}{x(p+t^E,t^S,Z)}$ measures the percentage change in demand caused by a 1 percent increase in the total price of good $x$ through a tax change, while $\varepsilon_{x,q|p} = -\frac{\partial x}{\partial p}\frac{q}{x(p+t^E,t^S,Z)}$ represents the analogous measure for a 1 percent increase in $q$ through a change in $p$.[26]

When discussing the intuition for the results below, we will focus on the case where $\theta < 1$ and interpret $\theta$ as a measure of the degree of inattention to the tax. However, our analysis permits $\theta > 1$ and more generally permits $\frac{\partial x}{\partial t}$ to differ from $\frac{\partial x}{\partial p}$ for any reason, not just inattention.[27] The formulas derived below therefore account for any errors that consumers may make when optimizing with respect to taxes.

*Production.* Assume that the supply of the numeraire good $y$ is perfectly elastic. This assumption shuts down general equilibrium effects by ensuring that the price of $y$ is unaffected by the tax on $x$. Good $x$ is produced by price-taking firms, which use $c(S)$ units of $y$ to produce $S$ units of $x$. The marginal cost of production is weakly increasing: $c'(S) > 0$ and $c''(S) \geq 0$. Let $\pi(S) = pS - c(S)$ denote the representative firm's profits at a given pretax price $p$ and level of supply $S$. Assuming that firms optimize perfectly, the supply function for good $x$ is implicitly defined by the first-order condition for $S$ in the profit-maximization problem: $p = c'(S(p))$.[28] Let $\varepsilon_{S,p} = \frac{\partial S}{\partial p}\frac{p}{S(p)}$ denote the price elasticity of supply.

## V.B Incidence

How is the burden of a tax shared between consumers and producers in competitive equilibrium when consumers optimize imperfectly with respect to taxes? We derive formulas for the incidence of the sales tax on producers and consumers which parallel the derivations of Laurence J. Kotlikoff and Lawrence H. Summers (1987) for the full-optimization case. As is standard in the literature on tax incidence, we use $D(p, t^S, Z)$ instead of $x(p, t^S, Z)$ to refer

[26] The empirical estimates of $\theta_\tau$ can be directly mapped to values for $\theta$ using the equation $\theta_\tau = \theta\frac{1+t^S}{1+\theta t^S}$. The reason that $\theta_\tau < \theta$ is that agents under react to price increases when the tax is ad-valorem, because part of the price increase raises the amount of the tax $p\tau^S$. For small values of $t^S$, $\theta_\tau = \theta$ and hence the values of $\theta_\tau$ reported in sections II and III roughly correspond to estimates of $\theta$.

[27] Although our evidence shows that $\theta < 1$ for commodity taxes that are not salient, this need not be the case for all taxes. The opaque estate tax system, for example, appears to cause many individuals to over perceive tax rates on wealth (Slemrod 2006).

[28] The literature in psychology and economics has argued that firms are less prone to systematic errors than consumers (see e.g. section IV of DellaVigna 2007). It would be straightforward to extend our analysis to allow for salience effects on the firm side as well, in which case the formulas will depend on $\frac{\partial S}{\partial p}$ and $\frac{\partial S}{\partial t^S}$.

to the demand curve in this subsection. Let $p = p(t^E, t^S)$ denote the equilibrium pretax price that clears the market for good $x$ as a function of the tax rates. The market clearing price $p$ satisfies

$$D(p + t^E, t^S, Z) = S(p) \tag{7}$$

Implicit differentiation of (7) yields the following results.

**Proposition 1** *The incidence on producers of increasing $t^S$ is*

$$\frac{dp}{dt^S} = \frac{\partial D/\partial t^S}{\partial S/\partial p - \partial D/\partial p} = -\frac{\varepsilon_{D,q|t^S}}{\frac{q}{p}\varepsilon_{S,p} + \varepsilon_{D,q|p}} = -\frac{\theta\varepsilon_{D,q|p}}{\frac{q}{p}\varepsilon_{S,p} + \varepsilon_{D,q|p}} \tag{8}$$

*and the incidence on consumers is*

$$\frac{dq}{dt^S} = 1 + \frac{dp}{dt^S} = \frac{\frac{q}{p}\varepsilon_{S,p} + \varepsilon_{D,q|p} - \varepsilon_{D,q|t^S}}{\frac{q}{p}\varepsilon_{S,p} + \varepsilon_{D,q|p}} = \frac{\frac{q}{p}\varepsilon_{S,p} + (1-\theta)\varepsilon_{D,q|p}}{\frac{q}{p}\varepsilon_{S,p} + \varepsilon_{D,q|p}}$$

*where $\partial D/\partial t^S$ and $\partial D/\partial p$ are both evaluated at $(p + t^E, t^S, Z)$ and $\partial S/\partial p$ is evaluated at $p$.*

Figure 3 illustrates the incidence of introducing a sales tax $t^S$ in a market that is initially untaxed. The figure plots supply and demand as a function of the pretax price $p$. The market initially clears at a price $p_0 = p(0, 0)$. When the tax is levied, the demand curve shifts inward by $t^S \partial D/\partial t^S$ units, creating an excess supply of $E = t^S \partial D/\partial t^S$ units of the good at the initial price $p_0$. To re-equilibriate the market, producers cut the pretax price by $E/(\partial S/\partial p - \partial D/\partial p)$ units. The only difference in the incidence diagram in Figure 3 relative to the traditional model without salience effects is that the demand curve shifts inward by $t^S \partial D/\partial t^S$ instead of $t^S \partial D/\partial p$. With salience effects, the shift in the demand curve is determined by the tax elasticity, while the price adjustment needed to clear the market is determined by the price elasticity. This is why one must estimate *both* the tax and price elasticities to calculate incidence.

Three general lessons about tax incidence emerge from the formulas in Proposition 1.

1. [Attenuated Incidence on Producers] Incidence on producers is attenuated by $\theta = \frac{\partial D/\partial t^S}{\partial D/\partial p}$ relative to the traditional model. Intuitively, producers face less pressure to reduce

the pretax price when consumers under react to the sales tax. In the extreme case where $\partial D/\partial t^S = 0$, consumers bear all of the tax, because there is no need to change the pretax price to clear the market. More generally, the incidence of a tax on consumers is inversely related to the degree of attention to the tax ($\theta$).

One interpretation of this result is that the demand curve becomes more inelastic when individuals are inattentive. Though changes in inattention and the price elasticity both affect the gross-of-tax-elasticity $\varepsilon_{D,q|t^S} = \theta\varepsilon_{D,q|p}$ in the same way, their effects on incidence are not equivalent. To see this, consider two markets, $A$ and $B$, where $\varepsilon^A_{S,p} = \varepsilon^B_{S,p} = 0.1$. In market $A$, demand is inelastic and consumers are fully attentive to taxes: $\varepsilon^A_{D,q|p} = 0.3$ and $\theta^A = 1$. In market $B$, demand is elastic but consumers are inattentive: $\varepsilon^B_{D,q|p} = 1$ and $\theta^B = 0.3$. An econometrician would estimate the same tax elasticity in both markets: $\varepsilon^A_{D,q|t^S} = \varepsilon^B_{D,q|t^S} = 0.3$. However, $[\frac{dp}{dt^S}]^A = -0.75$ whereas $[\frac{dp}{dt^S}]^B = -0.27$. In market $A$, suppliers bear most of the incidence since demand is 3 times more elastic to price than supply. In market $B$, even though demand is 10 times as price elastic as supply, producers are able to shift most of the incidence of the tax to consumers because of inattention.

Intuitively, a low price elasticity of demand has two effects on incidence: it reduces the shift in the demand curve but increases the size of the price cut needed to re-equilibriate the market for a given level of excess supply. Inattention to the tax also reduces the shift in the demand curve, but does not have the second offsetting effect. This difference is apparent in the formula for $\frac{dp}{dt}$ in (8), where $\varepsilon_{D,q|p}$ appears in both the numerator and denominator whereas $\theta$ appears only in the numerator. As a result, a 1 percent reduction in attention leads to greater incidence on consumers than a 1 percent reduction in the price elasticity. As $\varepsilon_{S,p}$ approaches 0, $\frac{dq}{dt^S}$ approaches $1-\theta$ irrespective of $\varepsilon_{D,q|p}$. If consumers are sufficiently inattentive, they bear most of the incidence of a tax even if supply is inelastic.

2. [No Tax Neutrality] Taxes that are included in posted prices have greater incidence on producers because they are fully salient: $\frac{dp}{dt^E} = \frac{\partial D/\partial t^S}{\partial S/\partial p - \partial D/\partial p} < \frac{dp}{dt^S}$. Taxes levied on producers are more likely to be included in posted prices than taxes levied on consumers because producers must actively "shroud" a tax levied on them in order to reduce its salience. Together, these observations suggest that producers will generally bear more of the incidence when a tax is levied on them than when it is levied on the consumers. Statutory incidence

affects economic incidence, contrary to intuition based on the full-optimization model.[29]

3. [Effect of Price Elasticity] Holding fixed the size of the tax elasticity $\varepsilon_{D,q|t^S}$, an increase in the price elasticity of demand *raises* incidence on consumers ($\partial[\frac{dp}{dt^S}]/\partial\varepsilon_{D,q|p} > 0$). This is because holding fixed the shift in the demand curve created by the introduction of the tax, a smaller price reduction is needed to clear the market if demand is very price elastic. In contrast, if the degree of inattention $\theta$ is held fixed as $\varepsilon_{D,q|p}$ varies, we obtain the conventional result $\partial[\frac{dp}{dt^S}]/\partial\varepsilon_{D,q|p} < 0$ because $\varepsilon_{D,q|t^S}$ and $\varepsilon_{D,q|p}$ vary at the same rate. Thus, taxing markets with more elastic demand could lead to greater or lesser incidence on consumers, depending on the extent to which the tax elasticity $\varepsilon_{D,q|t^S}$ covaries with the price elasticity $\varepsilon_{D,q|p}$.

## V.C Efficiency Cost

We begin by characterizing the excess burden of introducing a sales tax $t^S$ in an initially untaxed market with fixed producer prices. We then extend the analysis to allow for endogenous producer prices and pre existing excise and sales taxes.

### V.C.1 Definitions

We first define generalized indirect utility and expenditure functions that permit prices and taxes to have different effects. Let $V(p+t^E, t^S, Z) = u(x(p+t^E, t^S, Z)) + v(y(p+t^E, t^S, Z))$ denote the agent's indirect utility as a function of the posted price of good $x$, the sales tax, and wealth. Let $e(p+t^E, t^S, V)$ denote the agent's expenditure function, which represents the minimum wealth necessary to attain utility $V$ at a given posted price and sales tax. Let $R(t^E, t^S, Z) = tx(p+t^E, t^S, Z)$ denote tax revenue.

Following Herbert Mohring (1971) and Alan J. Auerbach (1985), we measure excess burden using the concept of equivalent variation. When $p$ is fixed, the excess burden of

[29] Consistent with this prediction, Busse, Silva-Risso, and Zettelmeyer (2006) find that 35 percent of manufacturer rebates given to car dealers are passed through to the buyer, while 85 percent of rebates given to buyers stay with the buyer. The reason is that most consumers did not find out about the dealer rebates. Rudolf Kerschbamer and Georg Kirchsteiger (2000) find that statutory evidence affects economic incidence in a lab experiment.

introducing a sales tax $t^S$ in a previously untaxed market is

$$EB(t^S) = Z - e(p, 0, V(p, t^S, Z)) - R(0, t^S, Z) \tag{9}$$

The value $EB(t^S)$ is the amount of additional tax revenue that could be collected from the consumer while keeping his utility constant if the distortionary tax were replaced with a lump-sum tax. Roughly speaking, $EB(t^S)$ can be interpreted as the total value of the purchases that fail to occur because of the tax. Our objective is to derive approximate expressions for (9) in terms of empirically estimable elasticities.

### V.C.2 Preference Recovery

The efficiency cost of a tax policy depends on two elements: (1) the change in behavior induced by the tax and (2) the effect of that change in behavior on the consumer's utility. The first element is observed empirically – we can estimate the demand function $x(p + t^E, t^S, Z)$. The second element is the key challenge for behavioral welfare economics. How do we compute indirect utility $V(p + t^E, t^S, Z)$ when the agent's behavior is not consistent with optimization? The following two assumptions allow us to recover $V$ without specifying a positive model for the demand function $x(p + t^E, t^S, Z)$.

**A1** Taxes affect utility only through their effects on the chosen consumption bundle. The agent's indirect utility given taxes of $(t^E, t^S)$ is

$$V(p + t^E, t^S, Z) = u(x(p + t^E, t^S, Z)) + v(y(p + t^E, t^S, Z))$$

**A2** When tax inclusive prices are fully salient, the agent chooses the same allocation as a fully-optimizing agent:

$$x(p, 0, Z) = x^*(p, 0, Z) = \arg\max u(x(p, 0, Z)) + v(Z - px(p, 0, Z))$$

Assumption A1 requires that consumption is a sufficient statistic for utility – that is, holding fixed the consumption bundle $(x, y)$, the tax rate or its salience has no effect on $V$. To understand the content of this assumption, consider the following situation in which it is

violated. In a bounded rationality model, the cognitive cost that the agent pays to calculate the total price when $t^S > 0$ makes his utility lower than pure consumption utility. Taxes that are not included in posted prices therefore generate deadweight burden beyond that due to the distortion in the consumption bundle (Chetty, Looney, and Kroft 2007). In such models, the excess burden computations in this paper correspond to the deadweight cost net of any increase in cognitive costs.[30]

Assumption A2 requires that the agent behaves like a fully-optimizing agent when all taxes are fully salient. That is, the agent's choices when total prices are fully salient reveal his true rankings. This assumption is violated when the agent's choices are suboptimal even without taxes. For example, if there are other "shrouded attributes" or if agents suffer from biases when optimizing relative to prices (Nina Mazar, Botond Koszegi, and Dan Ariely 2008), we would not directly recover true preferences from $x(p, 0, Z)$. The excess burden formulas derived below ignore errors in optimization relative to prices.

Using assumptions A1 and A2, we calculate consumer welfare and excess burden in two steps. We first use the demand function without taxes $x(p, 0, Z)$ to recover the agent's underlying preferences $(u(x), v(y))$ as in the full-optimization model. We then use the demand function with taxes $x(p+t^E, t^S, Z)$ to calculate the agent's indirect utility $V(p+t^E, t^S, Z)$ as a function of the tax rate. Conceptually, this method pairs the libertarian criterion of calculating welfare from individual choice with the assumption that the agent optimizes relative to true incentives only when tax inclusive prices are perfectly salient.

Our calculation of excess burden can be viewed as an application of Bernheim and Rangel's (2007) choice-based approach to welfare analysis. Bernheim and Rangel show that one can obtain bounds on welfare without specifying a positive theory of behavior by separating the inputs that matter for utility from "ancillary conditions" that do not. By applying a "refinement" to identify ancillary conditions under which an agent's choices reveal his true rankings, one can sharpen the bounds. In Bernheim and Rangel's terminology, our assumption A1 is that tax salience is an "ancillary condition" that affects choices but not true utility. Assumption A2 is a "refinement" which posits that the choices made when the tax is not

[30] Chetty, Looney, and Kroft (2007) show that the additional deadweight burden due to cognitive costs is likely to be negligible since relatively small cognitive costs generate substantial amounts of inattention.

perfectly salient are "suspect," and should be discarded when inferring the utility relevant for welfare analysis. This refinement allows us to obtain exact measures of equivalent variation and efficiency costs without placing structure on the model that generates $x(p + t^E, t^S, Z)$.

### V.C.3 Fixed Producer Prices

We derive simple formulas for excess burden using approximations analogous to those used by Harberger (1964) and Edgar K. Browning (1987). Like the widely applied Harberger-Browning formula, the formulas below ignore the third- and higher-order terms in the Taylor expansion for excess-burden. Hence, the formulas provide accurate measures of excess burden for small tax changes.

In this section, we characterize excess burden of introducing a sales tax in a market where production is constant-returns-to-scale ($c'' = 0$). In this case, the pretax price of $x$ is fixed at $p = c'(0)$ because the supply curve is flat. Moreover, since firms earns zero profits ($\pi = 0$), only consumer welfare matters for excess burden. To state the formula compactly, we introduce notation for income-compensated elasticities. Let $\partial x^c/\partial p = \partial x/\partial p + x\partial x/\partial Z$ denote the income-compensated (Hicksian) price effect. Define $\partial x^c/\partial t^S = \partial x/\partial t^S + x\partial x/\partial Z$ as the analogous income-compensated tax effect. Note that this "compensated tax effect" does not necessarily satisfy the Slutsky condition $\partial x^c/\partial t^S < 0$. It is possible to have an upward-sloping compensated tax-demand curve because $x(p, t^S, Z)$ is not generated by utility maximization. In contrast, assumption A2 guarantees $\frac{\partial x^c}{\partial p} < 0$ through the Slutsky condition. Let $\varepsilon^c_{x,q|p} = -\frac{\partial x^c}{\partial p}\frac{q}{x}$ and $\varepsilon^c_{x,q|t^S} = -\frac{\partial x^c}{\partial t^S}\frac{q}{x}$ denote the compensated price and tax elasticities.

**Proposition 2** *Suppose producer prices are fixed ($\varepsilon_{s,p} = \infty$). Under assumptions A1-A2, the excess burden of introducing a small tax $t^S$ in an untaxed market is approximately*

$$\begin{aligned} EB(t^S) &\simeq -\frac{1}{2}(t^S)^2\theta^c\partial x^c/\partial t^S \\ &= \frac{1}{2}(t^S)^2\theta^c x(p, t^S, Z)\frac{\varepsilon^c_{x,q|t^S}}{p + t^S} \end{aligned} \tag{10}$$

*where $\partial x^c/\partial t^S$ and $\partial x^c/\partial p$ are evaluated at $(p, 0, Z)$ and $\theta^c = \frac{\partial x^c/\partial t^S}{\partial x^c/\partial p} = \frac{\varepsilon^c_{x,q|t^S}}{\varepsilon^c_{x,q|p}}$ is the ratio of the compensated tax and price effects.*

**Proof.** We provide an instructive proof for the case of quasilinear utility ($v(y) = y$). The derivation for the general case is given in Appendix B. To reduce notation, we suppress wealth and write $x(p, t^S)$ since $\frac{\partial x}{\partial Z} = 0$ when utility is quasilinear.

We first use assumption A1 to obtain an expression for indirect utility:

$$\begin{aligned} V(p, t^S, Z) &= u(x(p, t^S) + y(p, t^S) = u(x(p, t^S)) + Z - (p + t^S)x(p, t^S) \\ &= Z + u(0) + \int_0^{x(p,t^S)} [u'(x) - (p + t^S)]dx \end{aligned} \tag{11}$$

Recognizing that $e$ is the inverse of $V$ and using (9), it follows that excess burden is

$$EB(t^S) = \int_{x(p,0)}^{x(p,t^S)} [u'(x) - p]dx.$$

To recover $u'(x)$ empirically, we use A2, which implies that

$$\begin{aligned} u'(x(p, 0)) &= u'(x^*(p, 0)) = p \\ \Rightarrow u'(x) &= P(x) \end{aligned}$$

where $P(x) = x^{-1}(p, 0)$ is the inverse price-demand function. It follows that

$$EB(t^S) = \int_{x(p,0)}^{x(p,t^S)} [P(x) - p]dx \tag{12}$$

which measures the area under the inverse-demand curve between $x(p, 0)$ and $x(p, t^S)$. This is an exact formula for excess burden that could be implemented with a non-parametric estimate of the demand curve. To obtain a simple analytical formula, we write $EB(t^S)$ using a Taylor expansion. We ignore the $(t^S)^3$ and higher-order terms, which is equivalent to assuming that $x(\cdot)$ is linear when utility is quasilinear. Evaluating the integral in (12) with this approximation yields

$$EB(t^S) \simeq -\frac{1}{2}(t^S)^2 \frac{\partial x/\partial t^S}{\partial x/\partial p} \partial x/\partial t^S \tag{13}$$

which corresponds to (10) since $\frac{\partial x}{\partial Z} = 0$ in the quasilinear case.

*Graphical Derivation.* Figure 4 illustrates the calculation of deadweight loss for the quasilinear case. The initial price of the good is $p_0$ and the price after the imposition of the sales tax is $p_0+t^S$. The figure plots two demand curves. The first is the standard Marshallian demand curve as a function of the total price of the good, $x(p,0)$. This *price-demand* curve coincides with the marginal utility $u'(x)$ as shown in the proof above. The second, $x(p_0,t^S)$ represents how demand varies with the tax on $x$. This *tax-demand* curve is drawn assuming $\partial x/\partial p \leq \partial x/\partial t^S$, consistent with the empirical evidence.

The agent's initial consumption choice prior to the introduction of the tax is depicted by $x_0 = x(p_0,0)$. Initial consumer surplus is given by triangle $ABC$, which equals total utility (up to a constant) as shown by (11). When the tax $t^S$ is introduced, the agent cuts consumption of $x$ by $\Delta x = -t^S \partial x/\partial t^S$. Notice that at the new consumption choice $x_1$, the agent's marginal willingness-to-pay for $x$ is below the total price $p_0 + t^S$ because he underreacts to the tax. This optimization error leads to a loss of surplus corresponding to triangle $DEF$. The consumer's surplus after the implementation of the tax is therefore given by triangle $DGC$ minus triangle $DEF$. The revenue raised from the tax corresponds to the rectangle $GBEH$. It follows that the change in total surplus – government revenue plus consumer surplus – equals the shaded triangle $AFH$. This is precisely the measure in (12) – the area under the price-demand curve between $x_0$ and $x_1$. The base of the triangle $(AH)$ has length $-t^S \frac{\partial x}{\partial t^S}$ while the height of the triangle $(AF)$ is $t^S \frac{\partial x/\partial t^S}{\partial x/\partial p}$, yielding (13).[31]

When utility is not quasilinear ($\frac{\partial x}{\partial Z} > 0$), the form of the formula remains exactly the same, but all the inputs are replaced by income-compensated effects, exactly as in the Harberger formula. The intuition for this difference is analogous to that in the full-optimization model: behavioral responses due to pure income effects are nondistortionary, since they would occur under lump sum taxation as well. Deadweight loss is determined by difference between the actual behavioral response ($\frac{\partial x}{\partial t^S}$) and the socially optimal response given the reduction in net-of-tax income ($-x\frac{\partial x}{\partial Z}$), which is $\frac{\partial x}{\partial t^S} - (-x\frac{\partial x}{\partial Z}) = \frac{\partial x^c}{\partial t^S}$.

[31] Another instructive derivation starts from the excess burden of taxation for a fully-optimizing agent, $EB^*$ (triangle $AID$). Starting from $EB^*$, we obtain excess burden for the agent who does not optimize fully (triangle $AFH$) by making two adjustments: (1) subtracting the additional revenue earned by the government because the agent under reacts to the tax (rectangle $HIDE$) and (2) adding the private welfare loss due to the optimization error (triangle $FED$).

Note that in Proposition 2 and all subsequent excess burden calculations, $\partial x^c/\partial p$ is evaluated at a point with zero sales tax $(p, 0)$. The reason is that we recover true preferences only when the posted price equals the total price: $x(p, t^S, Z) = x^*(p, t^S, Z)$ if and only if $t^S = 0$. If an environment without sales tax is not observed, one could implement the formula by assuming that the price elasticity does not depend on the tax rate ($\frac{d^2x^c}{dpdt^S} = 0$), a plausible assumption for small tax rates. Under this assumption, $\frac{dx^c}{dp}(p, 0, Z) = \frac{dx^c}{dp}(p, t^S, Z)$, which can be estimated empirically as in section III.

*Discussion.* The only difference between (10) and the canonical Harberger formula ($EB^*(t^S) = -\frac{1}{2}(t^S)^2\frac{\partial x^c}{\partial t^S}$) is the ratio of the tax and price effects $\frac{\partial x^c/\partial t^S}{\partial x^c/\partial p}$. Three general lessons about excess burden emerge from this ratio.

1. [Inattention Reduces Excess Burden if $\frac{\partial x}{\partial Z} = 0$] When utility is quasilinear, the tax $t^S$ generates deadweight cost equivalent to that created by a perfectly salient tax of $\theta t^S$. If agents ignore taxes completely and $\theta = 0$, then $EB = 0$. Taxation creates no inefficiency when $\theta = 0$ because the agent's consumption allocation coincides with the first-best bundle that he would have chosen under lump sum taxation.[32] As the degree of attention to taxes rises, excess burden rises at a quadratic rate: $EB \propto \theta^2$. Excess burden rises with the square of $\theta$ for the same reason that it rises with the square of the $t^S$ – the increasing marginal social cost of deviating from the first-best. Because $EB$ is a quadratic function of $\theta$ but a linear function of $\varepsilon_{x,q|p}$, inattention (reductions in $\theta$) and inelasticity (reductions in $\varepsilon_{x,q|p}$) have different effects on excess burden, as in the incidence analysis. Like incidence, excess burden depends on which side of the market is taxed. Since a tax on producers is likely to be included in posted prices, it leads to a larger reduction in demand and more deadweight loss than an equivalent tax levied on consumers when utility is quasilinear.

2. [Inattention Can Raise Excess Burden if $\frac{\partial x}{\partial Z} > 0$] Unlike in the quasilinear case, making a tax less salient to reduce $\frac{\partial x}{\partial t^S}$ can *increase* deadweight loss when there are income effects. In fact, a tax can create deadweight cost even if the agent completely ignores it and demand for the taxed good does not change, i.e. $\frac{\partial x}{\partial t^S} = 0$. This result contrasts with the canonical

[32] The consumer's *private* welfare always rises with $\theta$ – increased salience of tax-inclusive prices is always desirable from the consumer's perspective. However, the gain in the consumer's private welfare from full attention (triangle $FED$ in Figure 4) is more than offset by the resulting loss in government revenue (rectangle $HIDE$), which is why total surplus falls with $\theta$ when utility is quasilinear.

intuition that taxes generate deadweight costs only if they induce changes in demand. In the full-optimization model, taxation of a normal good creates a deadweight cost only if $\frac{\partial x}{\partial p} < 0$ since $\frac{\partial x}{\partial p} = 0 \Rightarrow \frac{\partial x^c}{\partial p} = 0$ given $\frac{\partial x}{\partial Z} > 0$. This reasoning fails when the tax-demand is not the outcome of perfect optimization, because there is no Slutsky condition for $\frac{\partial x^c}{\partial t^S}$. A zero uncompensated tax elasticity does not imply that the compensated tax elasticity is zero. Instead, when $\frac{\partial x}{\partial t^S} = 0$, $\frac{\partial x^c}{\partial t^S} = \frac{\partial x}{\partial Z}$ and (10) becomes

$$EB(t^S) = -\frac{1}{2}(t^S x)^2 \frac{\partial x/\partial Z}{\partial x^c/\partial p} \partial x/\partial Z$$

This equation shows that $EB > 0$ even when $\partial x/\partial t^S = 0$ in the presence of income effects. To understand this result, recall that the excess burden of a distortionary tax is determined by the extent to which the agent deviates from the allocation he would optimally choose if subject to a lump sum tax of an equivalent amount. In the quasilinear case, the agent's consumption bundle when ignoring the tax coincides with the bundle he would optimally choose under lump sum taxation, because the socially optimal choice of $x$ does not depend on total income. When utility is not quasilinear, an optimizing agent would reduce consumption of both $x$ and $y$ when faced with a lump sum tax. An agent who does not change his demand for $x$ at all when the tax is introduced ends up overconsuming $x$ relative to the social optimum. The income-compensated tax elasticity $\frac{\partial x^c}{\partial t^S} = \frac{\partial x}{\partial Z}$ is positive because the tax effectively distorts demand for $x$ *upward* once the income effect is taken into account, leading to inefficiency.

As a concrete example, consider an individual who consumes cars ($x$) and food ($y$). Suppose he chooses the same car he would have bought at a total price of $p_0$ because he does not perceive the tax ($\frac{\partial x}{\partial t^S} = 0$) and therefore has to cut back on food to meet his budget. This inefficient allocation of net-of-tax income leads to a loss in surplus. The lost surplus is proportional to the income effect on cars $\frac{\partial x}{\partial Z}$ because this elasticity determines how much the agent should have cut spending on the car to reach the social optimum given the tax. This example illustrates that policies which "hide" taxes can potentially create substantial deadweight loss despite attenuating behavioral responses, particularly when the income elasticity and expenditure on the taxed good are large.

Note that inattention to a tax on $x$ need not necessarily lead to $\frac{\partial x}{\partial t^S} = 0$. The effect of inattention on $\frac{\partial x}{\partial t^S}$ depends on how the agent meets his budget given the tax. The agent must reduce consumption of at least one of the goods to meet his budget when the tax on $x$ is introduced: $\frac{\partial x}{\partial t^S} + \frac{\partial y}{\partial t^S} = -x$. The way in which agents meet their budget may vary across individuals (Chetty, Looney, and Kroft 2007). For example, credit-constrained agents may be forced to cut back on consumption of $y$ if they ignore the tax when buying $x$, as in the car purchase example above, leading to $\frac{\partial x}{\partial t^S} = \theta = 0$ and $EB > 0$. Agents who smooth intertemporally, in contrast, may cut both $y$ as well as future purchases of $x$ (buying a cheaper car next time). Such intertemporal smoothing could lead to a long-run allocation closer to the socially optimal $\frac{\partial x}{\partial t^S} = -x\frac{\partial x}{\partial Z}$, in which case hidden taxes would lead to $\theta^c = 0$ and $EB = 0$. Importantly, Proposition 2 holds irrespective of how the agent meets his budget. Variations in the budget adjustment process are captured in the value of $\frac{\partial x^c}{\partial t^S}$.

3. [Role of Price Elasticity] Holding fixed $\varepsilon_{x,q|t^S}$, excess burden is inversely related to $\varepsilon_{x,q|p}$. As demand becomes *less* price-elastic, $EB$ *increases*. This can be seen in Figure 4, where the shaded triangle becomes larger as $x(p, 0)$ becomes steeper, holding $x(p_0, t^S)$ fixed. Intuitively, an agent with price-inelastic consumption has rapidly increasing marginal utility as his consumption level deviates from the first-best level. A given reduction in demand thus leads to a larger loss of surplus for an agent with more price-inelastic demand. As in the incidence analysis, taxing markets with more elastic demand could lead to greater or lesser excess burden, depending on the covariance between $\varepsilon_{x,q|t^S}$ and $\varepsilon_{x,q|p}$.

### V.C.4 Endogenous Producer Prices

We now drop the constant-returns to scale assumption and consider a market where the supply curve is upward sloping ($\varepsilon_{S,p} < \infty$). In this case, pretax prices are endogenous to the tax rate and firms earn positive profits, which must be accounted for in the welfare calculation. Following Auerbach (1985), assume that profits $\pi(S(p))$ are paid to the consumer using the numeraire $y$. In this subsection, we assume that utility is quasilinear ($v(y) = y$). We do not treat the case with income effects and $\varepsilon_{S,p} < \infty$ in this paper.[33] When $p(t_0^E, t_0^S)$

[33] Even in the full-optimization model, explicit formulas for excess burden cannot be obtained when there are income effects and non zero producer surplus (Auerbach 1985).

is endogenous and utility is quasilinear, excess burden is

$$EB(t^S) = Z - e(p_0, 0, V(p_1, t^S, Z)) + \pi_0 - \pi_1 - R$$

where $p_0 = p(0,0)$ and $p_1 = p(0,t^S)$ denote the equilibrium price before and after the introduction of the tax, $\pi_i = \pi(S(p_i))$, and $R = t^S x(p_1, t^S)$ denotes tax revenue (Auerbach 1985). Intuitively, excess burden equals the sum of the change in consumer surplus and producer surplus minus government revenue ($R$). Let $\frac{dx}{dt^S}$ denote the total reduction in equilibrium quantity caused by the tax, taking into account the effect of the price response: $\frac{dx}{dt^S} = \frac{\partial x}{\partial p}\frac{\partial p}{\partial t^S} + \frac{\partial x}{\partial t^S}$. Correspondingly, let $\varepsilon^{TOT}_{x,q|t^S} = -\frac{dx}{dt^S}\frac{q}{x(p,t^S)}$ denote the total change in demand caused by a 1 percent increase in the price $q = p_1 + t^S$ through an increase in $t^S$, taking into account the effect of the endogenous price response.

**Proposition 3** *Suppose utility is quasilinear ($v(y) = y$). Under assumptions A1-A2, the excess burden of introducing a small tax $t^S$ in a previously untaxed market is approximately*

$$\begin{aligned} EB(t^S) &\simeq -\frac{1}{2}(t^S)^2\theta\frac{dx}{dt^S} \\ &= \frac{1}{2}(t^S)^2\theta x(p_1, t^S)\frac{\varepsilon^{TOT}_{x,q|t^S}}{p_1 + t^S}. \end{aligned} \tag{14}$$

*where $\frac{dx}{dt^S}$ is evaluated at $(p_0, 0, Z)$ and $\theta = \frac{\partial x/\partial t^S}{\partial x/\partial p} = \frac{\varepsilon_{x,q|t^S}}{\varepsilon_{x,q|p}}$.*

**Proof.** The equation can be derived heuristically by calculating the area of the triangle that lies between the supply and the (no tax) price-demand curve $x(p,0)$ between the initial and final equilibrium quantities in Figure 3. The width of the triangle is $t^S\frac{dx}{dt^S}$ and the height is $\theta t^S$. See Appendix B for a formal derivation.

The lessons discussed above with fixed producer prices carry over to the case with endogenous prices. Indeed, the formula for excess burden with upward-sloping supply has exactly the same form as in (10). The only difference is that the size of the deviation in demand from the social optimum is given by the total derivative $\frac{dx}{dt^S}$ instead of the partial derivative $\frac{\partial x}{\partial t^S}$. When $p$ is fixed, these two derivatives coincide, so (14) collapses to the formula in Proposition 2 without income effects. When $p$ is endogenous, part of the distortion

in behavior is offset by the reduction in prices by producers to clear the market, leading to $|\frac{dx}{dt^S}| < |\frac{\partial x}{\partial t^S}|$ and smaller deadweight loss.

### V.C.5 Pre Existing Taxes

Finally, we calculate the marginal deadweight cost of increasing the sales tax by $\Delta t$ when there are pre existing taxes on good $x$. The initial excise tax rate is $t_0^E$ and sales tax rate is $t_0^S$. Let $p_0 = p(t_0^E, t_0^S)$ denote the initial pre-tax equilibrium price, $q_0 = p_0 + t_0^E + t_0^S$ denote the initial tax-inclusive price, and $x_0 = x(p_0 + t_0^E, t_0^S)$ denote initial quantity sold. Let $p_0^E = p(t_0^E, 0)$ denote the price when there is only an excise tax and $p_1 = p(t_0^E, t_0^S + \Delta t)$ denote the price after the tax increase. The following proposition provides approximate formulas for excess burden that are accurate for small initial tax rates $(t_0^E, t_0^S)$ and a small tax increase $\Delta t$. We explain why the approximation requires small initial tax rates after stating the result, and provide a formal statement of this requirement in Appendix B.

**Proposition 4** *Under assumptions A1-A2, the excess burden of a small sales tax increase $\Delta t$ starting from small initial tax rates $(t_0^E, t_0^S)$ is approximately given by the following formulas, with all derivatives evaluated at the no-sales-tax equilibrium $(p_0^E + t_0^E, 0)$.*
*(i) If producer prices are fixed ($\varepsilon_{s,p} = \infty$):*

$$\begin{aligned} EB(\Delta t|t_0^E, t_0^S) &\simeq -\frac{1}{2}(\Delta t)^2\theta^c\frac{\partial x^c}{\partial t^S} - \Delta t\frac{\partial x^c}{\partial t^S}[t_0^E + \theta^c t_0^S] \\ &= \frac{1}{2}(\Delta t)^2\theta^c x_0\frac{\varepsilon^c_{x,q|t^S}}{q_0} + \Delta t x_0\frac{\varepsilon^c_{x,q|t^S}}{q_0}[t_0^E + \theta^c t_0^S] \end{aligned}$$

*(ii) If utility is quasilinear ($v(y) = y$):*

$$\begin{aligned} EB(\Delta t|t_0^E, t_0^S) &\simeq -\frac{1}{2}(\Delta t)^2\theta\frac{dx}{dt^S} - \Delta t\frac{dx}{dt^S}[t_0^E + \theta t_0^S] \\ &= \frac{1}{2}(\Delta t)^2\theta x_0\frac{\varepsilon^{TOT}_{x,q|t^S}}{q_0} + \Delta t x_0\frac{\varepsilon^{TOT}_{x,q|t^S}}{q_0}[t_0^E + \theta t_0^S]. \end{aligned}$$

**Proof.** See Appendix B and Appendix Figure 1.

The first term in these formulas, proportional to $(\Delta t)^2$, is analogous to the triangle in the classic "Harberger trapezoid." This term comes from the loss in consumer and producer

surplus due to the tax increase, and is exactly the same as in the case without pre existing taxes. The second term, proportional to $\Delta t$, is analogous to the rectangle in the Harberger trapezoid. This term reflects the fiscal externality that the agents impose on the government by changing their behavior. Government revenue falls by $\Delta t \frac{dx^c}{dt^S}[t_0^E + t_0^S]$ because of the behavioral response to the tax increase. However, part of this fiscal externality is offset by a gain of $\Delta t \frac{dx^c}{dt^S}[1 - \theta^c]t_0^S$ in private utility to the consumer, since he was initially over consuming $x$ relative to his private optimum.

With pre existing taxes, tax increases can have a first-order (large) deadweight cost. The first-order deadweight cost due to $t_0^S$ is multiplied by $\theta^c$ because the deviation from the socially optimal level of $x$ caused by $t_0^S$ is proportional to $\theta^c$. If utility is quasilinear, levying a tax on top of a pre existing tax $t_0^S$ that is completely hidden ($\theta = \theta^c = 0$) generates only second-order (small) excess burden. If utility is not quasilinear, the same tax increase generates a first-order deadweight cost. Intuitively, the agent's consumption bundle is distorted relative to the social optimum to begin with if $\frac{dx}{dt^S} = 0$ when there are income effects. An increase in the tax rate exacerbates this pre existing distortion, creating a first-order deadweight cost even though there is no change in uncompensated demand.

We close with a technical remark about the approximations used in Proposition 4. The classic "Harberger trapezoid" formula requires that $\Delta t$ is small and that either (1) initial tax rates are small or (2) demand is linear ($\frac{d^2x}{dp^2} = 0$) over the $\Delta t$ interval (Auerbach 1985). In our case, for small $\Delta t$, condition (1) suffices to obtain simple formulas for $EB$ but (2) does not. The reason is that we can only recover the utility of $x(p, t^S)$ when $t^S = 0$ under A2. To calculate $V(p, t_0^S)$, we assume that $t_0^S$ is small and take a Taylor expansion around $V^*(p, t_0^S)$, ignoring the third- and higher-order terms. Linearity over the $\Delta t$ interval itself does not allow us to calculate $V(p, t_0^S)$ when $t_0^S > 0$.[34] Note that all of the approximations in Propositions 2-4 are needed only to obtain simple analytical formulas for $EB$. Exact measures of excess burden can be calculated using a non-parametric estimate of $x(p, t)$, as in the full-optimization model (Jerry A. Hausman and Whitney K. Newey 1995).

[34] Linearity of demand over the $\Delta t$ interval permits large $t_0^E$ but not $t_0^S$. To allow for large $t_0^S$, one must make additional parametric assumptions about the demand and utility functions between $x(p_0^E + t_0^E, 0)$ and $x(p_0 + t_0^E, t_0^S)$ to calculate $V(p_0 + t_0^E, t_0^S)$.

# VI Conclusion

The finding that individuals optimize imperfectly with respect to relatively simple commodity taxes suggests that they optimize imperfectly with respect to many tax policies, including income and capital taxation. The formulas developed here can be applied to characterize the incidence and efficiency costs of such policies. Much as Harberger identified the compensated price elasticity as the key parameter to be estimated in subsequent work, this paper has identified the compensated tax and price elasticities ($\varepsilon^c_{x,q|t^S}$ and $\varepsilon^c_{x,q|p}$) as "sufficient statistics" for empirical studies in behavioral public economics. Existing elasticity estimates based on variation in tax rates cannot be interpreted as wage or price elasticities unless the tax is perfectly salient.

A natural next step would be to extend the welfare analysis in this paper to characterize optimal taxation when agents optimize imperfectly, generalizing the results of Ramsey (1927) and Mirrlees (1971). Combining the formulas here with a positive theory would be useful for this analysis. For example, Chetty, Looney, and Kroft's (2007) bounded-rationality model predicts that attention and behavioral responses to taxation are larger when (1) tax rates are high, (2) the price-elasticity of demand is large, and (3) the amount spent on the good is large. Combined with the welfare analysis here, these predictions imply that in markets with these three characteristics, tax incidence should fall more heavily on producers and excess burden should be closer to the Harberger measure.

Finally, the approach to welfare analysis proposed here – using a domain where incentives are fully salient to characterize the welfare consequences of policies that are not salient – can be applied in other contexts. Many social insurance and transfer programs (e.g. Medicare, Social Security, the Earned Income Credit) have complex features and may induce suboptimal behaviors. By estimating behavioral responses to analogous programs whose incentives are more salient, one can characterize the welfare consequences of the existing programs more accurately. Another potential application is to optimal regulation (e.g. consumer protection laws, financial market regulations). By identifying "suboptimal" transactions using data on consumer's choices in domains where incentives are more salient, one could develop rules to maximize consumer welfare that do not rely on paternalistic judgements.

# References


[1] American Chamber of Commerce Researchers Association (ACCRA). ACCRA Cost of Living Index Quarterly Reports, 1982-2000. Louisville, KY.

[2] Anderson, Eric T. and Duncan I. Simester. 2003. "Effects of $9 Price Endings on Retail Sales: Evidence from Field Experiments." *Quantitative Marketing and Economics*, 1(1): 93-110.

[3] Atkinson, Anthony B. and Joseph E. Stiglitz. 1976. "The design of tax structure: direct versus indirect taxation." *Journal of Public Economic*s, 6(1): 55-75.

[4] Auerbach, Alan J. 1985. "The Theory of Excess Burden and Optimal Taxation," in *Handbook of Public Economics* vol. 1, ed. Alan Auerbach and Martin Feldstein, 61-128. Amsterdam: Elsevier Science Publishers B. V.

[5] Bernheim, B. Douglas and Antonio Rangel. 2007. "Beyond Revealed Preference: Choice-Theoretic Foundations for Behavioral Welfare Economics." National Bureau of Economic Research, Inc., NBER Working Papers: No. 13737.

[6] Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2004. "How Much Should We Trust Differences-in-Differences Estimates?" *Quarterly Journal of Economic*s, 119(1), 249-275.

[7] Beer Institute. Various years. *Brewer's Alamanac.* http://www.beerinstitute.org/statistics.asp?bid=200

[8] Blendon, Robert J., Stephen R. Pelletier, Marcus D. Rosenbaum, and Mollyann Brodie. 2003. "Tax Uncertainty." *Brookings Review* (Summer): 28-31.

[9] Blumkin, Tomer, Bradley J. Ruffle, and Yosi Ganun. 2008. "Are Income and Consumption Taxes Ever Really Equivalent? Evidence from a Real-Effort Experiment." University Library of Munich, Germany, MPRA Paper 6479.

[10] Browning, Edgar K. 1987. "On The Marginal Welfare Cost of Taxation." *American Economic Review*, 77: 11-23.

[11] Busse, Meghan, Jorge Silva-Risso and Florian Zettelmeyer. 2006. "$1000 Cash Back: The Pass-Through of Auto Manufacturer Promotions." *American Economic Review*, 96(4): 1253-1270.

[12] Chetty, Raj. 2006. "A New Method of Estimating Risk Aversion." *American Economic Review*, 96(5): 1821-1834.

[13] Chetty, Raj, Adam Looney, and Kory Kroft. 2007. "Salience and Taxation: Theory and Evidence." National Bureau of Economic Research, Inc., NBER Working Papers: No. 13330.

[14] Commerce Clearing House. Various years. *State Tax Handbook*, Chicago: Commerce Clearing House Inc.

[15] Cook, Philip J., Jan Ostermann and Frank A. Sloan. 2005. "Are Alcohol Excise Taxes Good for us? Short and Long Term Effects on Mortality Rates." National Bureau of Economic Research, Inc., NBER Working Papers: No. 11138.

[16] de Bartolome, Charles. 1995. "Which Tax Rate Do People Use: Average or Marginal?" *Journal of Public Economics*, 56: 79-96.

[17] DellaVigna, Stefano. 2007. "Psychology and Economics: Evidence from the Field." National Bureau of Economic Research, Inc., NBER Working Papers: No. 13420.

[18] DellaVigna, Stefano and Joshua Pollet. Forthcoming, 2008. "Investor Inattention and Friday Earnings Announcements." *Journal of Finance*.

[19] Feldman, Naomi E. and Peter Katuščák. 2006. "Should the Average Tax Rate be Marginalized?" CERGE Working Paper No. 304.

[20] Finkelstein, Amy N. 2007. "E-Z Tax: Tax Salience and Tax Rates." National Bureau of Economic Research, Inc., NBER Working Papers: No. 12924.

[21] Fisher, Ronald A. 1922. "On the interpretation of $\chi^2$ from contingency tables, and the calculation of P." *Journal of the Royal Statistical Society*, 85(1):87-94.

[22] Gabaix, Xavier and David Laibson. 2006. "Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets" *Quarterly Journal of Economics*, 121(2): 505-540.

[23] Gruber, Jonathan. 1994. "The Incidence of Mandated Maternity Benefits." *American Economic Review*, 84(3): 622-641.

[24] Gruber, Jonathan and Emmanuel Saez. 2002. "The Elasticity of Taxable Income: Evidence and Implications." *Journal of Public Economics*, 84: 1-32.

[25] Harberger, Arnold C. 1964. "The Measurement of Waste." *American Economic Review*, 54(3): 58-76.

[26] Hausman, Jerry A. and Whitney K. Newey. 1995. "Nonparametric Estimation of Exact Consumers Surplus and Deadweight Loss." *Econometrica*, 63(6): 1445-1476.

[27] Hoch, Stephen J., Byung-Do Kim, Alan L. Montgomery, and Peter E. Rossi. 1995. "Determinants of Store-Level Price Elasticity." *Journal of Marketing Research*, 32(1): 17-29.

[28] Hossain, Tanjim and John Morgan. 2006. "...Plus Shipping and Handling: Revenue (Non) Equivalence in Field Experiments on eBay." *Advances in Economic Analysis & Policy*, 6(3): Article 3.

[29] Hossain, Tanjim and John Morgan. 2008. "Shrouded Attributes and Information Suppression: Evidence from Field Experiments." UC-Berkeley Working Paper.

[30] Kerschbamer, Rudolf and Georg Kirchsteiger. 2000. "Theoretically robust but empirically invalid? An experimental investigation into tax equivalence." *Economic Theory*, 16: 719-734.

[31] Kotlikoff, Laurence J. and Lawrence H. Summers. 1987. "Tax Incidence," in *Handbook of Public Economics* Vol. 2, ed. Alan J. Auerbach and Martin Feldstein, 1043-1092. Amsterdam: Elsevier Science Publishers B. V.

[32] Lakins, Nekisha E., Gerald D. Williams, Hsiao–ye Yi, and Barbara A. Smothers. 2004. "Surveillance Report #66: Apparent Per Capita Alcohol Consumption: National, State, and Regional Trends, 1977-2002." Bethesda, MD: NIAAA, Alcohol Epidemiologic Data System.

[33] Liebman, Jeffrey B. and Richard J. Zeckhauser. 2004. "Schmeduling." Harvard KSG Working Paper.

[34] Mazar, Nina, Botond Koszegi, and Dan Ariely. 2008. "Price-Sensitive Preferences." UC-Berkeley Working Paper.

[35] Meyer, Bruce D. 1995. "Natural and Quasi-Experiments in Economics." *Journal of Business and Economic Statistic*s. 13(2): 151–61.

[36] Mill, John S. 1848. *Principles of Political Economy.* Oxford: Oxford University Press.

[37] Mirrlees, James A. 1971. "An Exploration in the Theory of Optimum Income Taxation." *The Review of Economic Studies*, 38(2): 175-208.

[38] Mohring, Herbert. 1971. "Alternative welfare gain and loss measures." *Western Economic Journal*, 9: 349-368.

[39] National Institute of Alcohol Abuse and Alcoholism (NIAAA). 2006. "Per Capita Alcohol Consumption." Data available at http://www.niaaa.nih.gov/Resources/DatabaseResources/QuickFacts/AlcoholSales.

[40] Nephew, Thomas M., Hsiao–ye Yi, Gerald D. Williams, Frederick S. Stinson, and Mary C. Dufour. 2004. "U.S. Alcohol Epidemiologic Data Reference Manual." NIH Publication No. 04-5563.

[41] Ramsey, Frank P. 1927. "A Contribution to the Theory of Taxation." *Economic Journal*, 37(145): 47-61.

[42] Rosen, Harvey S. 1976. "Taxes in a Labor Supply Model with Joint Wage-Hours Determination." *Econometrica*, 44(3): 485-507.

[43] Rosenbaum, Paul R. 1996. "Observational Studies and Nonrandomized Experiments," in *Handbook of Statistics* Vol. 13, ed S. Ghosh and C.R. Rao, 181-197. New York: Elsevier.

[44] Slemrod, Joel B. 2006. "The Role of Misconceptions in Support for Regressive Tax Reform." *National Tax Journal*, 59(1): 57-75.

[45] Tax Foundation. Various years. *Special Report: State Tax Rates and Collections*, Washington D.C. http://www.taxfoundation.org

[46] World Tax Database. 2006. University of Michigan Business School. http://www.bus.umich.edu/OTPR/otpr/introduction.htm

## Appendix A: Data and Empirical Methods

**Grocery Experiment.** The store changes product prices on Wednesday nights and leaves the prices fixed (with rare exceptions) for the following week, termed a "promotional week." To synchronize our intervention with this pricing cycle, a team of researchers and research assistants printed tags every Wednesday night and attached them to each of the 750 products. The tags were changed between 11 pm and 2 am, which are low-traffic times at the store. The tags were printed using a template and card stock supplied by the store (often used for sales or other additional information on a product) in order to match the color scheme and layout familiar to customers. The two control stores were chosen by a minimum-distance criterion based on the characteristics listed in Appendix Table 1.

The raw scanner data provided by the grocery chain contains information on weekly revenue and quantity sold for each product (UPC id) that was sold among the 108 categories listed in Appendix Table 2 in the three stores from 2005 week 1 to 2006 week 15. The original dataset contains 331,508 product-week-store observations. The quantity and revenue variables are measured net of returns (i.e., returns count as negative sales). We exclude 1,756 observations where the weekly quantity or revenue was zero or negative, which are cases where as many or more items were returned than purchased in that week. Including these observations does not affect the results. Finally, we aggregate to the category-week-store level by summing quantity and revenue across products, setting the sum to zero if no products were sold in a given category-week-store.

The average price for each category of goods is defined as $P_{ct} = \Sigma_{i \in c}(p_{it}\bar{q}_i)/\Sigma_{i \in c}\bar{q}_i$ where $c$ indexes the category, $t$ time, and $i$ products, $p_{it}$ is the price of good $i$ at time $t$, and $\bar{q}_i$ is the average quantity sold of good $i$. This "category price" is effectively a price index for a fixed basket of products where each product's weight in the basket is determined by its average weekly sales over the period before and during the experiment. Since the scanner data reports only items that have sold each week, we impute prices for unsold items when constructing $P_{ct}$. In particular, we use the price in the last observed transaction for unsold products; if no previous price is available, we use the next available price. Alternative imputation methods – such as using the closest observed price, or an average of previous and subsequent prices – give similar results. Varying the imputation technique has little impact on the estimates in Tables 4 and 5 because items requiring imputation have low sales volume, and therefore receive little weight in the category-level price variable.

**Grocery Store Survey.** We surveyed 91 customers entering the treatment store in August 2006 about their knowledge of sales taxes. Survey respondents were offered candy bars and sodas to spend a few minutes filling out the survey displayed in Appendix Exhibit 2. After collecting basic demographic information, the survey asked individuals to report whether each of eight goods were subject to sales tax or not. Many individuals remarked while filling out the survey that they did not think about taxes while shopping, and therefore were hesitant to report which goods were taxed. These individuals were asked to mark their best guess to avoid nonresponse bias. To assess whether knowledge of taxes is correlated with experience, we also asked whether individuals had purchased each of these goods recently. Finally, we asked questions about tax rates and bases – the sales tax rate in the city where the store is located, the state income tax rate, and the tax base for the federal estate tax.

**Alcohol Analysis.** Data on aggregate annual beer, wine, spirits, and ethanol consumption by state are available from the NIAAA (2006) from 1970-2003. These data contain information on total gallons of beer sold by wholesalers because this measure determines tax liabilities. See Thomas M. Nephew et al. (2004) and Nekisha E. Lakins et al. (2004) for details on data construction.

State excise tax rates on beer are primarily obtained from the Brewer's Almanac (various years), published annually by the Beer Institute. These rates were verified and corrected using the Tax Foundation's State Tax Collections and Rates (various years) and the State Tax Handbook. Our measure of the excise rate includes taxes that are statutorily 'local' excise taxes – which are sometime excluded from state statistics available in the Brewer's Almanac – that are applied state-wide. Specifically, in Alabama, Georgia, and Louisiana all counties or localities levy an excise tax in addition to the state excise tax.

Excise taxes on alcohol frequently differ by product, packaging, and whether sold for on- or off-premise consumption. In states where rates differ, our measure corresponds to the excise tax on packaged 12oz. beer, sold for off-premise consumption, with an alcohol content of 3.2 percent or more. Excise rates on other beer products are highly correlated with this measure across states, and the timing of tax changes for different categories of alcoholic beverages within a state are virtually identical. Per-gallon taxes are converted to per-case rates by multiplying by 2.25, the number of gallons in 24 12oz. cans or bottles. The excise tax rate is converted into an ad valorem rate by dividing the real CPI-adjusted beer excise tax per case in year 2000 dollars by the average cost of a case of beer in the United States in 2000, as measured by the Beer Institute. Since Alaska has a higher price level than the continental United States, we follow Census Bureau practice and adjust its price level up by 25 percent when calculating the percentage excise tax rate. None of our results are affected by this adjustment, or by excluding Alaska entirely. For a subset of years (1982-2000) and states, we have actual beer price data from the ACCRA cost of living index survey, which samples the price of a six pack of beer (Budweiser, Schlitz, or Miller Lite) in large cities. We define the ACCRA price variable as the annual average of all prices in each state.

State sales taxes are obtained primarily from the World Tax Database (2006) at the University of Michigan. These data were verified and corrected using state Department of Revenue websites and the State Tax Handbook. Four states (KS, VT, DC, MN) apply a higher sales tax rate to alcohol than other products. In those states we include the alcohol rate rather than the general sales rate when they differ. We supplement the data on state-level sales taxes with data on average local sales tax rates, which are imputed from data on local revenues from the Census Bureau's Survey of State and Local Government Finances and a tax base defined as state revenues divided by the state rate.

Since our estimation strategy relies on the timing and magnitude of the tax changes, we evaluate the precision of the data by regressing the change in the log of state tax revenues on the change in the log of the sales tax rate, controlling for state income. In the full sample, the coefficient estimate on the sales tax rate is 0.76 (s.e. 0.03). A state-by-state analysis of changes in rates and changes in revenues also yields similarly high correlations, with the exception of West Virginia. In WV, the correlation between sales tax rates and revenues is near zero and statistically insignificant, perhaps because the tax base is often changed at the same time as the rate. Since this problem could artificially attenuate the sales tax elasticity, we exclude West Virginia from our analysis.

## Appendix B: Proofs

We prove Proposition 4, since it nests the results in Propositions 2 and 3 as the case where $t_0^E = 0$ and $t_0^S = 0$. The proofs below provide approximations for $EB$ that are accurate for small initial tax rates and tax changes. To be precise, let $t_0^E = \lambda \bar{t}^E$, $t_0^S = \lambda \bar{t}^S$, and $\Delta t = \lambda \overline{\Delta t}$. We derive expressions for $EB(\Delta t | t_0^S, t_0^E)$ using Taylor expansions that ignore terms proportional to $\lambda^n$ where $n \geq 3$. That is, the formulas ignore terms proportional to $(\Delta t)^3$, $(t_0^S)^3$, $(\Delta t)^2 t_0^S$, $(\Delta t)^2 t_0^E$, $(\Delta t)(t_0^E)^2$, etc.

**Proof of Proposition 4i**

Definitions: Let $x_0^E = x(p+t_0^E, 0)$, $x_0 = x(p+t_0^E, t_0^S)$, $x_1 = x(p+t_0^E, t_0^S + \Delta t^S)$, $x_0^* = x(p+t_0^E + t_0^S, 0)$ and $x_1^* = x(p+t_0^E + t_0^S + \Delta t^S, 0)$. Let $V^*(p+t_0^E, t_0^S, Z)$ denote the utility attained by a fully optimizing agent who consumes the optimal bundle $(x^*(p+t^E, t^S, Z), y^*(p+t^E, t^S, Z))$. Let $R^*(p+t^E, t^S, Z) = (t^E + t^S)x^*(p+t^E, t^S, Z)$ denote tax revenue obtained from a fully optimizing agent.

Let the agent's loss from failing to optimize relative to the tax be denoted by

$$G(t_0^E, t_0^S) = e(p, 0, V^*(p+t_0^E, t_0^S)) - e(p, 0, V(p+t_0^E, t_0^S))$$

The gain in revenue due to the agent's underreaction to the tax is

$$\Delta R(t_0^E, t_0^S) = R(p+t_0^E, t_0^S, Z) - R^*(p+t_0^E, t_0^S, Z)$$

Recall that excess burden in the full optimization case is

$$EB^*(t_0^E, t_0^S) = Z - e(p, 0, V(p+t_0^E, t_0^S, Z)) - R^*(p+t_0^E, t_0^S, Z).$$

Combining these three equations, we can rewrite the formula for excess burden in (9) as

$$EB(t_0^E, t_0^S) = EB^* - \Delta R + G.$$

Using this formulation for $EB$, the excess burden of a sales tax increase $\Delta t$ is

$$EB(\Delta t | t_0^E, t_0^S) = \Delta EB^* - \Delta\Delta R + \Delta G \tag{15}$$

where the difference operator $\Delta X(t_0^E, t_0^S) = X(t_0^E, t_0^S + \Delta t) - X(t_0^E, t_0^S)$. We will use Taylor expansions to obtain simple expressions for each of these three terms below.

i) Auerbach (1985) shows that ignoring third-order terms, excess burden for an optimizing agent is approximately

$$\Delta EB^* = -\frac{\partial x^c}{\partial p} \Delta t [t_0^E + t_0^S + \frac{1}{2}\Delta t]$$

ii) The $\Delta\Delta R$ term can be written as:

$$\Delta\Delta R = -(t_0^E + t_0^S + \Delta t)(x_1^* - x_1) + (t_0^E + t_0^S)(x_0^* - x_0)$$

Ignoring the third- and higher-order terms (proportional to $\lambda^n, n \geq 3$) in $EB$, we can write this equation as

$$\Delta\Delta R = (\Delta t)^2 \left(\frac{\partial x}{\partial t} - \frac{\partial x}{\partial p}\right) + \Delta t\left(\frac{\partial x}{\partial t} - \frac{\partial x}{\partial p}\right)(t_0^E + 2t_0^S)$$

iii) Simplifying the expression for $G$ requires more work. First recall that the expenditure function is

$$e(p, t^S, V) = (p + t^S)x^c(p, t^S, V) + y^c(p, t^S, V)$$

and hence

$$\frac{\partial e}{\partial V} = (p + t^S)\frac{\partial x^c}{\partial V} + \frac{\partial y^c}{\partial V}.$$

The expenditure minimization problem is

$$\min(p + t^S)x^c + y^c \text{ s.t. } u(x) + v(y) = V$$

Differentiating the utility constraint for the expenditure minimization problem (EMP) yields

$$u'(x^c)\frac{dx^c}{dV} + v'(y^c)\frac{dy^c}{dV} = 1$$

The first-order-condition for the EMP implies

$$u'(x^{*c}) = (p + t^S)v'(y^{*c})$$

and hence we obtain the equation

$$(p + t_0^E + t_0^S)\frac{\partial x^{*c}}{\partial V} + \frac{\partial y^{*c}}{\partial V} = \frac{1}{v'(y^{*c})} = \frac{\partial e(p + t_0^E, t_0^S, V^*)}{\partial V}$$

where all the derivatives are evaluated at $(p + t_0^E, t_0^S, V^*)$. Using a Taylor expansion, we write

$$G = \frac{\partial e(p + t_0^E, t_0^S, V^*)}{\partial V}[V^*(p+t_0^E, t_0^S, Z) - V(p+t_0^E, t_0^S, Z)] - \frac{1}{2}\frac{\partial^2 e(p + t_0^E, t_0^S, V^*)}{\partial V^2}[V^* - V]^2 + ...$$

We show below that $V^* - V$ is proportional to $\lambda^2$; hence, the $[V^* - V]^2$ and higher-order terms in this expansion can be ignored under the second-order approximation. Hence, we can write

$$G = \frac{[V^*(p + t_0^E, t_0^S, Z) - V(p + t_0^E, t_0^S, Z)]}{v'(y^{*c}(p + t_0^E, t_0^S, V^*))}$$

Define the utility gain from choosing the optimal level $x^*$ vs. another point $x$ as

$$\begin{aligned}\widetilde{G}(x) &= u(x^*) - u(x) + v(y^*) - v(y) \\ &= u'(x^*)(x^* - x) - \frac{1}{2}u''(x^*)(x^* - x)^2 + O_u^3 + v'(y^*)(y^* - y) - \frac{1}{2}v''(y^*)(y^* - y)^2 + O_v^3\end{aligned}$$

where $O_u^3$ and $+O_v^3$ represent the third- and higher order terms of the Taylor expansions for $u$ and $v$. All of the terms in $O_u^3$ and $+O_v^3$ are ultimately proportional to $\lambda^n$ where $n \geq 3$, so we ignore these terms from this point onward.

Using the first-order-condition that characterizes the choice of the fully-optimizing agent,

$$u'(x^*) = (p + t_0^E + t_0^S)v'(y^*)$$

and the identity

$$(p + t_0^E + t_0^S)(x^* - x) = (y - y^*)$$

we obtain

$$\begin{aligned} \widetilde{G} &= -\frac{1}{2}u''(x^*)(x^* - x)^2 - \frac{1}{2}v''(y^*)(y^* - y)^2 \\ &= -\frac{1}{2}(x^* - x)^2[u''(x^*) + v''(y^*)(p + t_0^E + t_0^S)^2] \end{aligned} \tag{16}$$

Totally differentiating the fully-optimizing agent's first-order-condition with respect to $p$ yields

$$\begin{aligned} u''(x^*)\frac{\partial x^*}{\partial p} &= v'(y^*) + (p + t_0^E + t_0^S)v''(y^*)\frac{\partial y^*}{\partial p} \\ &= v'(y^*) + (p + t_0^E + t_0^S)[-(p + t_0^E + t_0^S)\frac{\partial x^*}{\partial p} - x^*]v''(y^*). \end{aligned}$$

It follows that

$$[u''(x^*) + (p + t_0^E + t_0^S)^2 v''(y^*)]\frac{\partial x^*}{\partial p} = v'(y^*) - (p + t_0^E + t_0^S)x^* v''(y^*)$$

and hence

$$\widetilde{G} = -\frac{1}{2}(x^* - x)^2 \frac{[v'(y^*) - (p + t_0^E + t_0^S)x^* v''(y^*)]}{\partial x^*/\partial p}. \tag{17}$$

Defining $\gamma_y = -y^* v''(y^*)/v'(y^*)$ it follows that

$$G \simeq \frac{\widetilde{G}}{v'(y^*)} = -\frac{1}{2}(x^* - x)^2 \frac{1}{\partial x^*/\partial p}[1 + (p + t_0^E + t_0^S)\frac{x^*}{y^*}\gamma_y]. \tag{18}$$

Finally, we use a result from Chetty (2006) which relates the coefficient of relative risk aversion to the ratio of the income effect to the substitution effect:

$$\gamma_y = \frac{-y^*}{p + t_0^E + t_0^S} \frac{\frac{\partial x^*}{\partial z}}{\frac{\partial x^{*c}}{\partial p}}. \tag{19}$$

Inserting this expression into (18) yields

$$G \simeq -\frac{1}{2}(x^* - x)^2 \frac{1}{\partial x^*/\partial p}[1 - x^* \frac{\frac{\partial x^*}{\partial z}}{\frac{\partial x^{*c}}{\partial p}}] = -\frac{1}{2}(x^* - x)^2 \frac{1}{\partial x^{*c}/\partial p}$$

Note that $\partial x^c/\partial p = \partial x^{*c}/\partial p$ and $\partial x/\partial p = \partial x^*/\partial p$ at the no-sales-tax point $p + t_0^E$ under A2. Thus, for small $\Delta t$, $t_0^S$, and $t_0^E$,

$$\begin{aligned}
\Delta G &\simeq -\frac{1}{2}\frac{1}{\partial x^c/\partial p}\{(x_1^* - x_1)^2 - (x_0^* - x_0)^2\} \\
&\simeq -\frac{1}{2}\frac{1}{\partial x^c/\partial p}(\Delta t)^2 \{(\frac{\partial x}{\partial p})^2 + (\frac{\partial x}{\partial t})^2 - 2\frac{\partial x}{\partial t}\frac{\partial x}{\partial p}\} - \Delta t \frac{1}{\partial x^c/\partial p}(\frac{\partial x}{\partial p} - \frac{\partial x}{\partial t})^2 t_0^S
\end{aligned}$$

where the second approximation ignores third- and higher-order terms and all derivatives are evaluated at $(p + t_0^E, 0)$.

Combining the expressions for $\Delta G$, $\Delta\Delta R$, and $\Delta EB^*$ above using (15) and collecting terms yields the formula in Proposition 4ii.

**Proof of Proposition 4ii**

Definitions: Let $p_0^E = p(t^E, 0)$, $p_0 = p(t^E, t^S)$, and $p_1 = p(t^E, t^S + \Delta t)$. Let $\pi_0$ and $\pi_1$ denote profits before and after implementation of the tax. To reduce notation, we suppress wealth in the demand function and write $x(p, t)$ since $Z$ does not affect $x$ when utility is quasilinear. Let $x_0^E = x(p_0^E + t_0^E, 0)$, $x_0 = x(p_0 + t_0^E, t_0^S)$, $x_1 = x(p_1 + t_0^E, t_0^S + \Delta t^S)$. Let $R_0 = (t_0^E + t_0^S)x_0$ and $R_1 = (t_0^E + t_0^S + \Delta t^S)x_1$.

Excess burden with pre-existing taxes and quasi-linear utility is (Auerbach 1985):

$$EB(\Delta t | t_0^S, t_0^E) = Z - e(p_0 + t_0^E, t_0^S, V(p_1 + t_0^E, t_0^S + \Delta t, Z + \pi_1)) + (\pi_0 - \pi_1) - (R_1 - R_0) \quad (20)$$

Using the definition of the expenditure function for the quasilinear case and the definition of the profit function, we can write (20) as

$$EB = u(x_0) - u(x_1) + c(x_1) - c(x_0) \quad (21)$$

This expression measures the area of the trapezoid that lies between the price-demand and supply curves between $x_0$ and $x_1$, shown in Appendix Figure 1. The derivation below is essentially an algebraic calculation of the area of that trapezoid using a series of Taylor expansions. To begin, we write (21) as

$$EB = u'(x_0)(x_0 - x_1) - \frac{1}{2}u''(x_0)(x_0 - x_1)^2 - O_u^3 + c'(x_0)(x_1 - x_0) + \frac{1}{2}c''(x_0)(x_1 - x_0)^2 + O_c^3$$

where $O_f^3 = \sum_3^\infty \frac{f^n(x_0)}{n!}(x_1 - x_0)^n$, $f = u, c$ consists of the third- and higher-order elements of the Taylor series. Observe that the first-order-condition for the optimal choice of $x$ for

the consumer is

$$u'(x^*(p)) = p$$

Total differentiation of this condition yields

$$u''(x^*(p)) = \frac{1}{\partial x^*(p)/\partial p}$$

Recognizing that $x^*(p_0^E + t_0^E, 0) = x(p_0^E + t_0^E, 0) = x_0^E$, we take a Taylor approximation around $x_0^E$ to write

$$\begin{aligned} u'(x_0) &= p_0^E + t_0^E + u''(x_0^E)(x_0 - x_0^E) + ... \\ u''(x_0) &= \frac{1}{\partial x/\partial p} + u'''(x_0^E)(x_0 - x_0^E) + ... \\ x_0 - x_0^E &= \frac{dx}{dt^S} t_0^S + \frac{1}{2}\frac{dx}{d(t^S)^2}(t_0^S)^2 + ... \end{aligned}$$

Note that the derivatives in these equations are evaluated at the no-sales-tax equilibrium $(p_0^E + t_0^E, 0)$ because this is the only point at which the first order conditions hold.

Similarly, the first-order-conditions from firm optimization and a Taylor approximation around $x_0^E$ can be used to write

$$\begin{aligned} c'(x_0) &= p_0^E + c''(x_0^E)(x_0 - x_0^E) \\ c''(x_0) &= \frac{1}{\partial S/\partial p} + c'''(x_0^E)(x_0 - x_0^E) + ... \end{aligned}$$

Finally, a Taylor expansion around $x_0$ yields:

$$(x_0 - x_1) = \frac{dx}{dt^S}\Delta t - \frac{1}{2}\frac{d^2x}{d(t^S)^2}(\Delta t)^2 + ...$$

Ignoring the third- and higher-order terms (proportional to $\lambda^n, n \geq 3$) in $EB$, we can combine the Taylor expansions above to write

$$\begin{aligned} EB &= -\frac{1}{2}\left(\frac{1}{\partial x/\partial p} - \frac{1}{\partial S/\partial p}\right)\left(\frac{dx}{dt^S}\Delta t\right)^2 \\ &\quad -\left(\frac{dx}{dt^S}\Delta t - \frac{1}{2}\frac{d^2x}{d(t^S)^2}(\Delta t)^2\right)\left(t_0^E + \frac{dx}{dt^S}\left(\frac{1}{\partial x/\partial p} - \frac{1}{\partial S/\partial p}\right)t_0^S\right) \end{aligned} \tag{22}$$

To simplify this expression, we use the expression for $\frac{\partial p}{\partial t^S}$ in Proposition 1 to obtain the following result:

$$\frac{dx}{dt^S}\left(\frac{1}{\partial x/\partial p} - \frac{1}{\partial S/\partial p}\right) = \left(\frac{\partial x}{\partial p}\frac{\partial p}{\partial t^S} + \frac{\partial x}{\partial t^S}\right)\left(\frac{1}{\partial x/\partial p} - \frac{1}{\partial S/\partial p}\right) = \theta \tag{23}$$

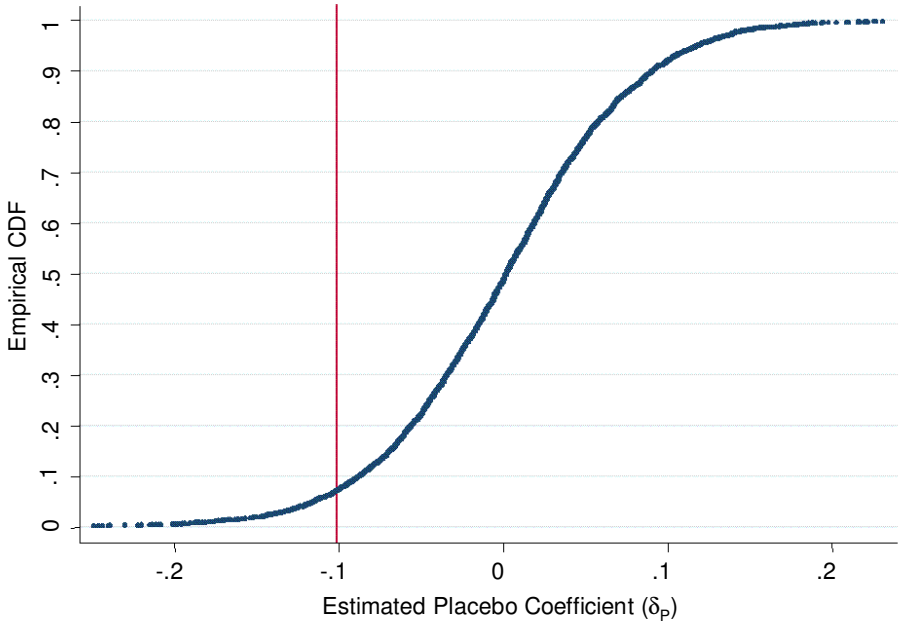
Combining (23) with (22), we obtain

$$EB = -\frac{1}{2}(\Delta t)^2\theta\frac{dx}{dt^S} - \frac{dx}{dt^S}\Delta t(t_0^E + \theta t_0^S).$$

**EXHIBIT 1**
Tax-Inclusive Price Tags

**Figure 1**

Distribution of Placebo Estimates: Log Quantity




NOTE– This figure plots the empirical distribution of placebo effects ($G$) for log quantity. The CDF is constructed from 4,725 estimates of $\delta_P$ using the specification in column 3 of Table 4. No parametric smoothing is applied: the CDF appears smooth because of the large number of points used to construct it. The vertical line shows the treatment effect estimate reported in Table 4.

**Figure 2a**

Per Capita Beer Consumption and State Beer Excise Taxes

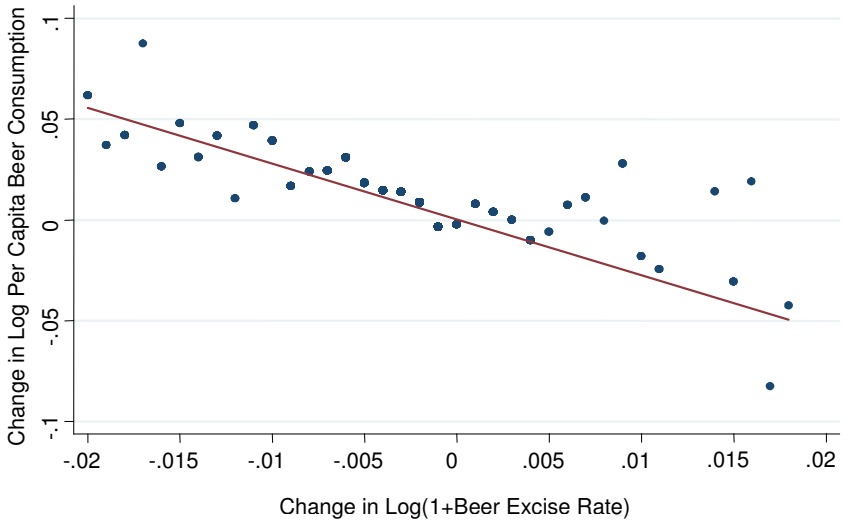
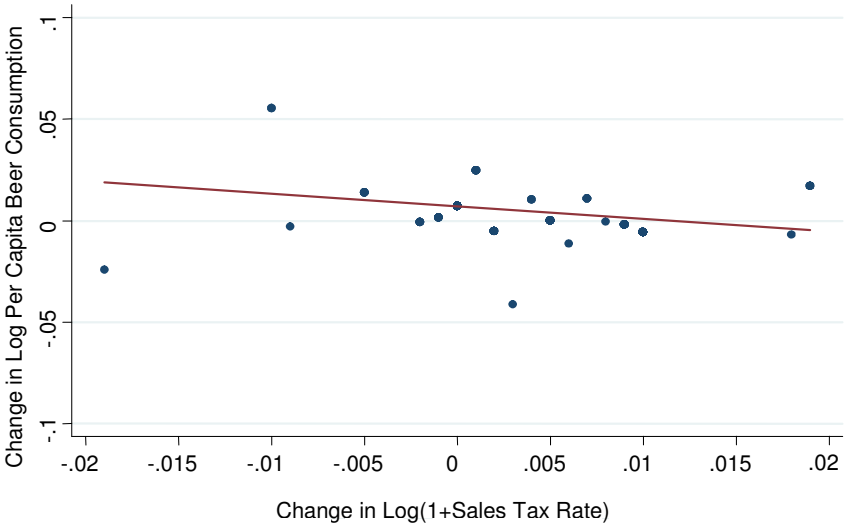


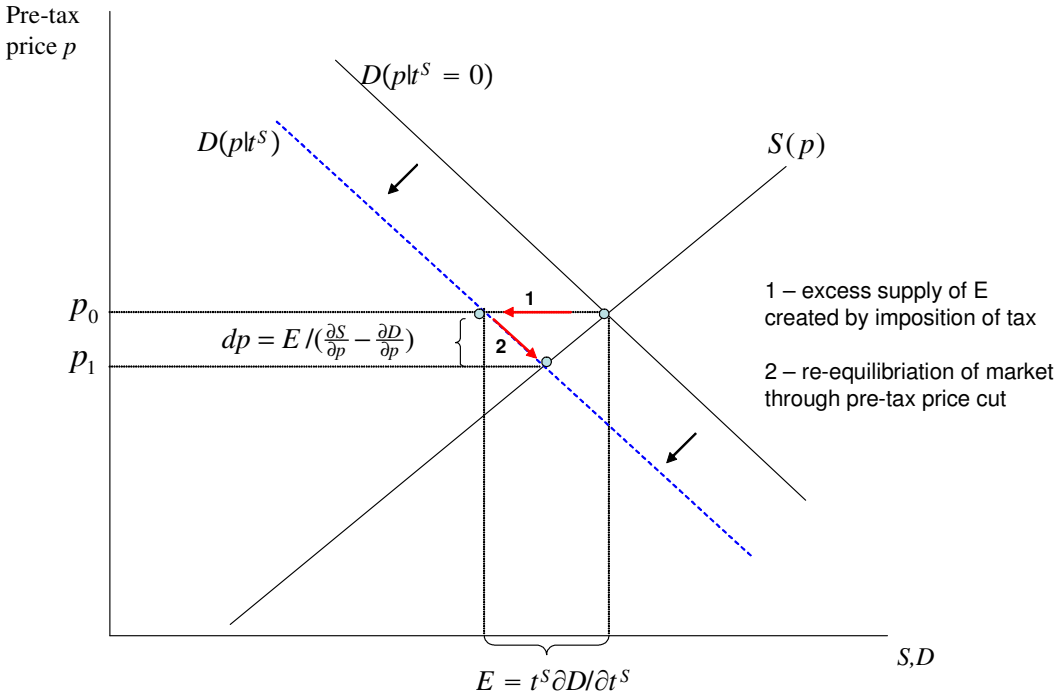

**Figure 2b**

Per Capita Beer Consumption and State Sales Taxes



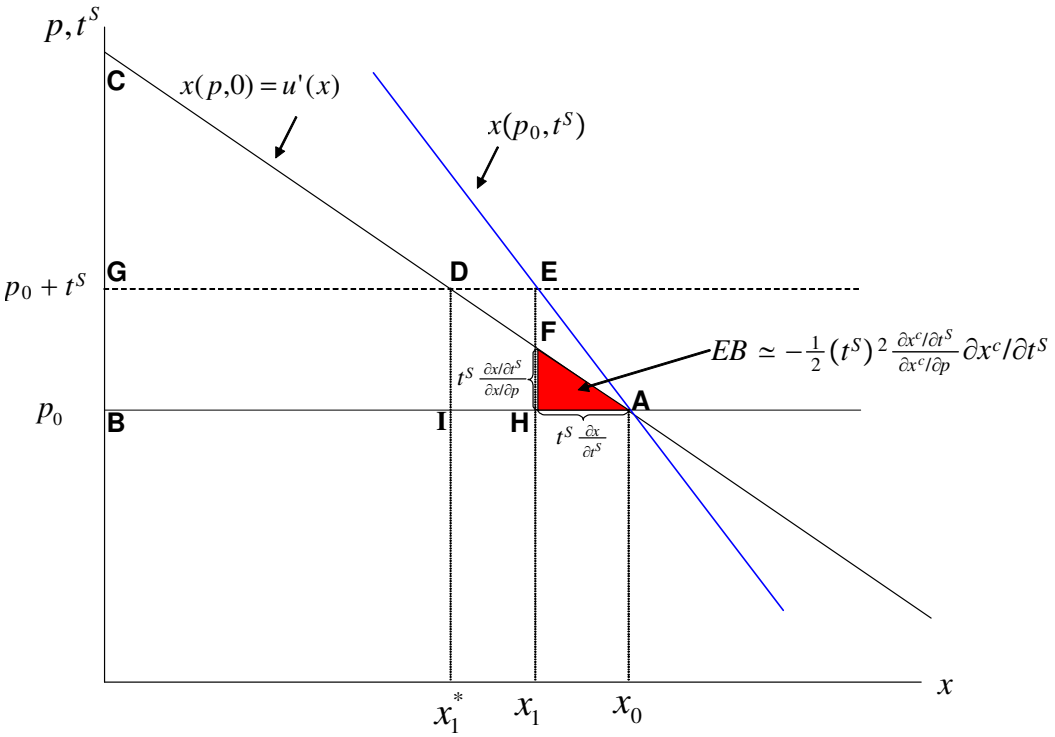

NOTE–These figures plot within-state annual changes in beer consumption against within-state changes in gross-of-tax-prices ($1 + t^E$ and $1 + t^S$). To construct Figure 2a, we round each state excise tax change to the nearest tenth of a percent (0.1%), and compute the mean change in log beer consumption for observations with the same rounded excise tax change. Figure 2a plots the mean consumption change against the rounded excise tax rates. Figure 2b is constructed analogously, rounding sales tax changes to the nearest 0.1%. See Appendix A for data sources and sample definition.

**Figure 3**
Incidence of Taxation




NOTE–This figure illustrates the incidence of introducing a tax $t^s$ levied on consumers in a market that is initially untaxed. The figure plots supply and demand as a function of the pre-tax price $p$. The initial price-demand curve is $D(p|t^S = 0)$; the price-demand curve after the tax is introduced is $D(p|t^S)$. When the tax is levied, the demand curve shifts inward by $t^S \times \partial D/\partial t^S$ units, creating an excess supply of $E = t^S \times \partial D/\partial t^S$. To re-equilibriate the market, producers cut the pre-tax price by $E/(\partial S/\partial p - \partial D/\partial p)$ units, implying $\frac{dp}{dt^S} = \frac{\partial D/\partial t^S}{\partial S/\partial p - \partial D/\partial p}$.

**Figure 4**
Excess Burden with Quasilinear Utility and Fixed Producer Prices




NOTE–This figure illustrates the efficiency cost of introducing a tax $t^s$ levied on consumers when utility is quasilinear and producer prices are fixed. The figure plots two demand curves: (1) the price-demand curve $x(p,0)$, which shows how demand varies with the pre-tax price of the good and (2) the tax-demand curve $x(p_0, t^S)$, which shows how demand varies with the tax. The figure is drawn assuming $|\partial x/\partial t^S| \leq |\partial x/\partial p|$ consistent with the empirical evidence. The tax reduces demand from $x_0$ to $x_1$. The consumer's surplus after the implementation of the tax is given by triangle *DGC* minus triangle *DEF*. The revenue raised from the tax corresponds to the rectangle *GBEH*. The change in total surplus – government revenue plus consumer surplus – equals the shaded triangle *AFH*.

TABLE 1
Survey Evidence: Summary Statistics

| Panel A: Classroom Survey: | Mean | Median | Standard Deviation |
|---|---|---|---|
| Original Price Tags: | | | |
| Correct tax-inclusive price w/in $0.25 | 0.18 | 0.00 | 0.39 |
| Experimental Price Tags: | | | |
| Correct tax-inclusive price w/in $0.25 | 0.75 | 1.00 | 0.43 |
| T-test for equality of means: $p < 0.001$ | | | |
| N=49 | | | |

| Panel B: Grocery Store Survey | Mean | Median | Standard Deviation |
|---|---|---|---|
| Local Sales Tax Rate (Actual rate is 7.375 percent) | 7.48 | 7.39 | 0.80 |
| Fraction correctly reporting tax status: | | | |
| All items | 0.82 | 1.00 | 0.38 |
| Beer | 0.90 | 1.00 | 0.30 |
| Cigarettes | 0.98 | 1.00 | 0.15 |
| Cookies | 0.65 | 1.00 | 0.48 |
| Magazines | 0.87 | 1.00 | 0.34 |
| Milk | 0.82 | 1.00 | 0.38 |
| Potatoes | 0.81 | 1.00 | 0.39 |
| Soda | 0.76 | 1.00 | 0.43 |
| Toothpaste | 0.80 | 1.00 | 0.40 |
| N=91 | | | |

Notes: *Panel A*: Reports summary statistics for a survey of 49 students who were shown regular (non-tax inclusive) price tax and the experimental (tax-inclusive) price tags. Statistics shown are for an indicator for whether individual reported total bill within 25 cents of total tax-inclusive price. See Appendix Exhibit 1 for survey instrument. *Panel B*: Reports summary statistics for a survey of 91 customers at the treatment grocery store. See Appendix Exhibit 2 for survey instrument.

TABLE 2
Grocery Experiment: Summary Statistics

| | Treatment Store | | Control Stores | | Total |
|---|---|---|---|---|---|
| | Treatment Products | Control Products | Treatment Products | Control Products | All Stores and Products |
| *A. Category Level Statistics:* | (1) | (2) | (3) | (4) | (5) |
| Weekly quantity sold per category | 25.08 (24.1) | 26.63 (38.1) | 27.84 (27.4) | 30.64 (47.0) | 29.01 (42.5) |
| Weekly revenue per category | $97.85 (81.9) | $136.05 (169.9) | $107.04 (92.3) | $154.66 (207.7) | $143.10 (187.1) |
| Number of categories | 13 | 95 | 13 | 95 | 108 |
| *B. Product Level Statistics* | | | | | |
| Pre-tax product price | 4.46 (1.8) | 6.26 (4.3) | 4.52 (1.7) | 6.31 (4.2) | 6.05 (4.1) |
| Pre-tax product price (weighted by quantity sold) | 4.27 (4.7) | 5.61 (3.9) | 4.29 (1.6) | 5.59 (3.8) | 5.45 (3.7) |
| Weekly quantity sold per product (conditional >0) | 1.47 (0.9) | 1.82 (1.6) | 1.61 (1.1) | 1.98 (191.0) | 1.88 (1.7) |

Notes: Statistics reported are means with standard deviations in parentheses. Statistics are based on sales between 2005 week 1 and 2006 week 15. Data source is scanner data obtained from a grocery chain. The "Treatment Store" is the store where the intervention took place; the "Control Stores" are two nearby stores in the same chain. "Treatment products" are cosmetics, hair care accessories, and deodorants. "Control Products" are other toiletries located in the same aisles; see Appendix Table 2 for complete list. Product price reflects actual price paid, including any discount if product is on sale. See Appendix A for data sources and sample definition.

**TABLE 3**
Effect of Posting Tax-Inclusive Prices: DDD Analysis of Mean Quantity Sold

TREATMENT STORE

| Period | Control Categories | Treated Categories | Difference |
|---|---|---|---|
| Baseline (2005:1-2006:6) | 26.48<br>(0.22)<br>[5510] | 25.17<br>(0.37)<br>[754] | -1.31<br>(0.43)<br>[6264] |
| Experiment (2006: 8-2006:10) | 27.32<br>(0.87)<br>[285] | 23.87<br>(1.02)<br>[39] | -3.45<br>(0.64)<br>[324] |
| Difference over time | 0.84<br>(0.75)<br>[5795] | -1.30<br>(0.92)<br>[793] | **$DD_{TS}$ = -2.14**<br>(0.64)<br>[6588] |

CONTROL STORES

| Period | Control Categories | Treated Categories | Difference |
|---|---|---|---|
| Baseline (2005:1-2006:6) | 30.57<br>(0.24)<br>[11020] | 27.94<br>(0.30)<br>[1508] | -2.63<br>(0.32)<br>[12528] |
| Experiment (2006: 8-2006:10) | 30.76<br>(0.72)<br>[570] | 28.19<br>(1.06)<br>[78] | -2.57<br>(1.09)<br>[648] |
| Difference over time | 0.19<br>(0.64)<br>[11590] | 0.25<br>(0.92)<br>[1586] | **$DD_{CS}$ = 0.06**<br>(0.90)<br>[13176] |
| | | **DDD Estimate** | ***-2.20***<br>(0.58)<br>[19764] |

Notes: Each cell shows mean quantity sold per category per week, for various subsets of the sample. Standard errors (clustered by week) in parentheses, number of observations in square brackets. Experimental period spans week 8 in 2006 to week 10 in 2006. Baseline period spans week 1 in 2005 to week 6 in 2006. Lower panel reflects averages across the two control stores.

**TABLE 4**

Effect of Posting Tax-Inclusive Prices: Regression Estimates

| Dependent Variable: | Quantity per category | Revenue per category ($) | Log quantity per category | Quantity per category | Quantity (treat. categories only) |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| **Treatment** | **-2.20** | **-13.12** | **-0.101** | **-2.27** | **-1.55** |
| | (0.59) | (4.88) | (0.03) | (0.60) | (0.35) |
| Average Price | -3.15 | -3.24 | | -3.04 | -15.06 |
| | (0.26) | (1.74) | | (0.25) | (3.55) |
| Average Price Squared | 0.05 | 0.06 | | 0.05 | 1.24 |
| | (0.00) | (0.03) | | (0.00) | (0.34) |
| Log Average Price | | | -1.59 | | |
| | | | (0.11) | | |
| Before Treatment | | | | -0.21 | |
| | | | | (1.07) | |
| After Treatment | | | | 0.20 | |
| | | | | (0.78) | |
| Category, Store, Week FEs | x | x | x | x | x |
| Sample size | 19,764 | 19,764 | 18,827 | 21,060 | 2,379 |

Notes: Standard errors, clustered by week, reported in parentheses. All columns report estimates of the linear regression model specified in equation (4). Quantity and revenue reflect total sales of products within a given category per week in each store. Average price is a weighted average of the prices of the products for sale in each category using a fixed basket of products (weighted by total quantity sold) over time. In column 3, observations are weighted by total revenue by category-store. Specification 4 includes "placebo" treatment variables (and their interactions) for the 3 week period before the experiment and the 3 week period after the experiment. Specification 5 reports DD estimates restricting the sample to treatment product categories only (at both treatment and control stores). In this specification, the "Treatment" variable is defined as the interaction between the treatment store dummy and treatment time dummy.

**TABLE 5**
Summary Statistics for State Beer Consumption, Taxes, and Regulation

| | |
|---|---|
| Per-Capita Beer Consumption (cans) | 243.2<br>(46.1) |
| State Beer Excise Tax ($/case) | 0.51<br>(0.50) |
| State Beer Excise Tax (percent) | 6.4<br>(8.2) |
| Sales Tax (percent) | 4.3<br>(1.9) |
| Drinking Age is 21 | 0.73<br>(0.44) |
| Drunk Driving Standard | 0.66<br>(0.47) |
| Any Alcohol Regulation Change | 0.19<br>(0.39) |
| N (number of state-year pairs) | 1,690 |

Notes: Statistics reported are means with standard deviations in parentheses. Observations are by state for each year from 1970 to 2003. "Drinking Age is 21" is an indicator for whether the state-year has a legal drinking age of 21; "Drunk Driving Standard" indicates state-year has a threshold blood alcohol content level above which one is automatically guilty of drunk driving. "Any Alcohol Regulation Change" is a dummy variable equal to one in any year where a state has raised the drinking age, implemented a stricter drunk driving standard, implemented an administrative license revocation law, or a zero tolerance youth drunk driving law. See Appendix A for data sources and sample definition.

**TABLE 6**
Effect of Excise and Sales Taxes on Beer Consumption

Dependent Variable: Change in Log(per capita beer consumption)

| | Baseline | Bus. Cycle | Alc. Regulations | Region Trends |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| **ΔLog(1+Excise Tax Rate)** | **-0.87** | **-0.91** | **-0.89** | **-0.71** |
| | (0.17) | (0.17) | (0.17) | (0.18) |
| **ΔLog(1+Sales Tax Rate)** | **-0.20** | **-0.00** | **-0.02** | **-0.05** |
| | (0.30) | (0.30) | (0.30) | (0.30) |
| ΔLog(Population) | 0.03 | -0.07 | -0.07 | -0.09 |
| | (0.06) | (0.07) | (0.07) | (0.08) |
| ΔLog(Income per Capita) | | 0.22 | 0.22 | 0.22 |
| | | (0.05) | (0.05) | (0.05) |
| ΔLog(Unemployment Rate) | | -0.01 | -0.01 | -0.01 |
| | | (0.01) | (0.01) | (0.01) |
| Alcohol Regulation Controls | | | x | x |
| Year Fixed Effects | x | x | x | x |
| Region Fixed Effects | | | | x |
| F-Test for Equality of Tax Elasticities (Prob>F) | 0.05 | 0.01 | 0.01 | 0.06 |
| Sample Size | 1,607 | 1,487 | 1,487 | 1,487 |

Notes: Standard errors, clustered by state, in parentheses. All specifications are estimated on full sample for which data are available (state unemployment rate data is unavailable in early years). Column 3 includes three indicators for whether the state implemented per se drunk driving standards, administrative license revocation laws, or zero tolerance youth drunk driving laws, and the change in the minimum drinking age (measured in years). Column 4 includes fixed effects for each of nine census regions. F-test tests null hypothesis that coefficients on excise and sales tax rate variables are equal.

**TABLE 7**

Effect of Excise and Sales Taxes on Beer Consumption: Robustness Checks

| | Dependent Variable: Change in Log(per capita beer consumption) | | | | Dep Var: Share Ethanol from Beer |
|---|---|---|---|---|---|
| | IV for Excise w/ Policy | 3-Year Differences | IV for ACCRA Beer Price | Food Exempt | |
| | (1) | (2) | (3) | (4) | (5) |
| **ΔLog(1+Excise Tax Rate)** | **-0.62** | **-1.11** | | **-0.91** | **0.16** |
| | (0.28) | (0.46) | | (0.22) | (0.13) |
| **ΔLog(Beer Price)** | | | **-0.88** | | |
| | | | (0.42) | | |
| **ΔLog(1+Sales Tax Rate)** | **-0.03** | **-0.00** | **0.09** | **-0.14** | **0.25** |
| | (0.30) | (0.32) | (0.59) | (0.30) | (0.22) |
| ΔLog(Population) | -0.06 | -1.30 | -0.05 | 0.03 | 0.09 |
| | (0.07) | (0.37) | (0.15) | (0.07) | (0.05) |
| ΔLog(Income per Capita) | 0.22 | 0.09 | 0.23 | 0.22 | 0.01 |
| | (0.05) | (0.07) | (0.10) | (0.05) | (0.03) |
| ΔLog(Unemployment Rate) | -0.01 | -0.00 | 0.00 | -0.01 | 0.00 |
| | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) |
| Alcohol Regulation Controls | x | x | x | x | x |
| Year Fixed Effects | x | x | x | x | x |
| F-Test for Equality of Tax Elasticities (Prob>F) | 0.16 | 0.05 | 0.12 | 0.04 | 0.74 |
| Sample Size | 1487 | 1389 | 825 | 937 | 1487 |

Notes: Standard errors, clustered by state, in parentheses. Column 1 replicates column 3 of Table 6, instrumenting for excise tax rate changes with the nominal excise tax rate divided by the average price of a case of beer from 1970 to 2003 to eliminate variation in the tax rate due to inflation erosion. In Column 2, all variables are defined using three-year differences instead of first-differences. Column 3 instruments for the the log change in the ACCRA survey price of beer using the log change in the gross-of-excise tax rate. Column 4 restricts the sample to states where all food was exempt from taxation in 2000. In column 5, the dependent variable is the fraction of total ethanol consumption in each state-year accounted for by beer. F-test tests null hypothesis that coefficients on excise and sales tax rate variables are equal.

**APPENDIX TABLE 1**
Descriptive Statistics: Grocery Stores

| | Treatment Store | Control Store #1 | Control Store #2 |
|---|---|---|---|
| *A. Store Characteristics* | | | |
| Mean Weekly Revenue ($) | 307,297 | 268,193 | 375,114 |
| Total Floor Space (sq ft) | 41,609 | 34,187 | 37,251 |
| Store Opening Year | 1992 | 1992 | 1990 |
| *B. City Characteristics (in 1999)* | | | |
| Population | 88,625 | 96,178 | 90,532 |
| Median Age (years) | 33.9 | 31.1 | 32.3 |
| Median Household Income ($) | 57,667 | 51,151 | 60,359 |
| Mean Household Size | 2.8 | 2.9 | 3.1 |
| Percent bachelor's degree or higher | 19.4 | 20.4 | 18.2 |
| Percent Married | 60.2 | 56.9 | 58.1 |
| Percent White | 72.1 | 56.2 | 65.3 |
| Distance to Treatment Store (miles) | | 7.7 | 27.4 |

Notes: Data on store characteristics obtained from grocery chain. Weekly revenue statistics based on sales in calendar year 2005. Data for city characteristics are obtained from the U.S. Census Bureau, Census 2000. Control stores were chosen using a least-squares minimum-distance criterion based on this set of variables.

**APPENDIX TABLE 2**

Category Classification in Grocery Store Data

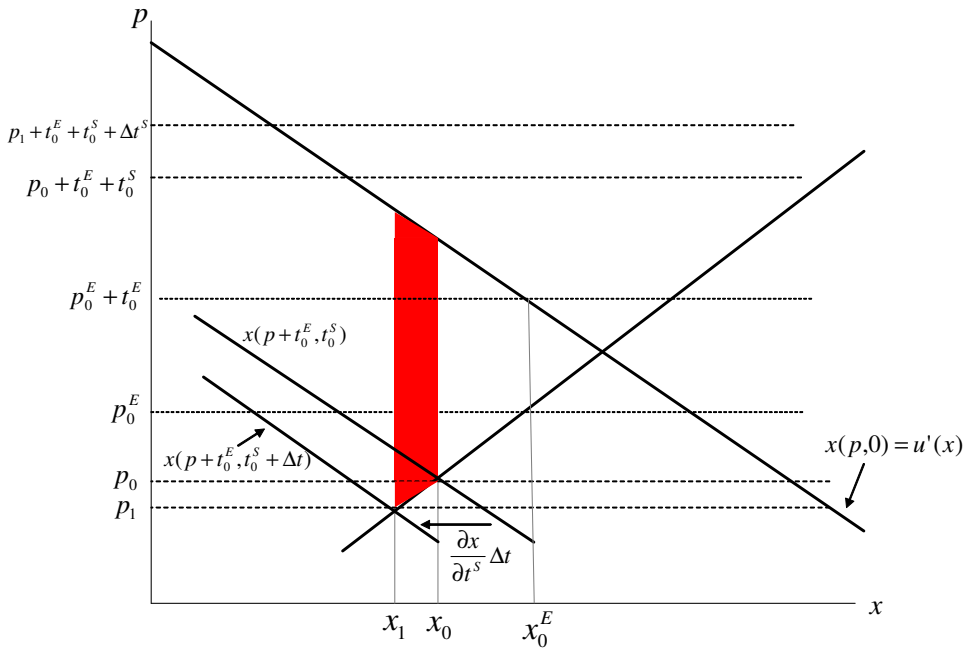
| Categories | Group Description | Category Description | Mean Weekly Revenue |
|---|---|---|---|
| Treatment | | | |
| 5101 | Deodorant | Aerosols | 82.40 |
| 5103 | Deodorant | Body Sprays | 55.22 |
| 5105 | Deodorant | Roll-ons | 44.12 |
| 5110 | Deodorant | Clear Solids | 323.38 |
| 5115 | Deodorant | Clear Soft | 35.13 |
| 5120 | Deodorant | Clear | 123.48 |
| 5125 | Deodorant | Visible Sticks | 75.57 |
| 5245 | Hair Care | Accessories | 189.47 |
| 5501 | Cosmetics | Facial | 84.20 |
| 5505 | Cosmetics | Eye | 195.00 |
| 5510 | Cosmetics | Nail | 73.38 |
| 5515 | Cosmetics | Lipstick | 48.39 |
| 5520 | Cosmetics | Accessories | 19.37 |
| Control | | | |
| 5005 | Oral Hygiene | At Home Whitening | 107.24 |
| 5010 | Oral Hygiene | Manual Toothbrush | 340.57 |
| 5012 | Oral Hygiene | Power Toothbrush | 120.89 |
| 5015 | Oral Hygiene | Oral Rinse/Mouthwash | 314.75 |
| 5020 | Oral Hygiene | Denture Care | 96.82 |
| 5025 | Oral Hygiene | Dental Floss Products | 116.75 |
| 5030 | Oral Hygiene | Interdental Implements | 26.76 |
| 5035 | Oral Hygiene | Oral Analgesics | 115.45 |
| 5040 | Oral Hygiene | Portable Oral Care | 52.84 |
| 5201 | Hair Care | Professional Daily Hair Care | 310.75 |
| 5205 | Hair Care | Performance Daily Hair Care | 983.31 |
| 5210 | Hair Care | Value Daily Hair Care | 290.11 |
| 5215 | Hair Care | Dandruff Hair Care | 116.37 |
| 5220 | Hair Care | Therapeutic Hair Care | 20.54 |
| 5225 | Hair Care | Hair Growth | 12.85 |
| 5230 | Hair Care | Kids Hair Care | 46.75 |
| 5235 | Hair Care | Hair Color | 430.18 |
| 5250 | Hair Care | African American Hair Care | 59.91 |
| 5301 | Skin Care | Bar Soap | 395.65 |
| 5305 | Skin Care | Liquid Hand Soap | 138.95 |
| 5308 | Skin Care | Liquid Waterless Sanitizer | 41.00 |
| 5310 | Skin Care | Body Wash | 339.04 |
| 5312 | Skin Care | Bath Care | 29.82 |
| 5314 | Skin Care | Image Bath Boutique | 36.07 |
| 5315 | Skin Care | Acne Prevention | 140.02 |
| 5318 | Skin Care | Acne Treatment | 12.57 |
| 5320 | Skin Care | Basic Facial Care | 427.17 |
| 5322 | Skin Care | Anti-aging/Treatments skin care | 27.99 |
| 5325 | Skin Care | Hand & Body Skin Care | 312.46 |
| 5330 | Skin Care | Lip Care | 91.97 |
| 5335 | Skin Care | Cotton | 169.72 |
| 5340 | Skin Care | Depilatories | 33.61 |

| | | | |
|---|---|---|---|
| 5345 | Skin Care | Adult Skin Care | 172.57 |
| 5350 | Skin Care | Child/Baby Sun Care | 26.06 |
| 5401 | Shave Needs/Men's Personal Care | Razors | 161.13 |
| 5405 | Shave Needs/Men's Personal Care | Cartridges | 389.02 |
| 5410 | Shave Needs/Men's Personal Care | Disposable Razors | 195.95 |
| 5415 | Shave Needs/Men's Personal Care | Shave Preps | 210.23 |
| 5420 | Shave Needs/Men's Personal Care | Men's Skin Care | 14.98 |
| 5601 | Vitamins and Dietary Supplements | Multiple Vitamins | 264.95 |
| 5605 | Vitamins and Dietary Supplements | Joint Relief | 89.57 |
| 5610 | Vitamins and Dietary Supplements | Calcium | 72.59 |
| 5615 | Vitamins and Dietary Supplements | Letters | 120.32 |
| 5620 | Vitamins and Dietary Supplements | Specialty Supplements | 65.91 |
| 5625 | Vitamins and Dietary Supplements | A/O Minerals | 31.65 |
| 5630 | Vitamins and Dietary Supplements | Herbal Supplements | 74.18 |
| 5701 | Pain Relief | Adult Aspirin | 48.23 |
| 5703 | Pain Relief | Enteric/Antacid/Buffered Aspirin | 14.90 |
| 5704 | Pain Relief | Low Strength Aspirin | 62.19 |
| 5705 | Pain Relief | Adult Acetaminophen | 203.24 |
| 5710 | Pain Relief | Ibuprofen Adult | 252.89 |
| 5715 | Pain Relief | Naproxen Sodium | 54.63 |
| 5716 | Pain Relief | Adult Compounds | 86.75 |
| 5718 | Pain Relief | Specialty Indication Pain | 88.92 |
| 5725 | Pain Relief | Children's/Infants Analgesics | 187.25 |
| 5730 | Pain Relief | Sleeping Aids | 64.99 |
| 5735 | Pain Relief | Stimulants | 14.82 |
| 5750 | Pain Relief | Nighttime Pain Relief | 76.19 |
| 5760 | Pain Relief | External Analgesic | 144.08 |
| 5799 | Pain Relief | GM/HBC Trial Size | 66.88 |
| 5801 | Respiratory | Pediatric Cold/Flu/Cough/Allergy/Sinus | 229.73 |
| 5805 | Respiratory | Adult Cough, Cold, Flu | 925.93 |
| 5835 | Respiratory | Adult Allergy/Sinus | 500.74 |
| 5840 | Respiratory | Nasal Products | 269.19 |
| 5845 | Respiratory | Bronchial Asthma | 41.45 |
| 5850 | Respiratory | Cough Drops/Throat Relief | 252.64 |
| 5855 | Respiratory | Thermometers/Covers | 37.72 |
| 5901 | Digestive Health | Acid Neutralizers | 243.37 |
| 5905 | Digestive Health | Acid Combination | 17.21 |
| 5910 | Digestive Health | Acid Blockers | 131.62 |
| 5915 | Digestive Health | Proton Pump Inhibitors (PPI) | 92.82 |
| 5920 | Digestive Health | Multi Symptom Gastro Intestinal Relief | 70.60 |
| 5925 | Digestive Health | Gas Relief | 49.46 |
| 5930 | Digestive Health | Motion Sickness/Anti-Nausea | 24.32 |
| 5935 | Digestive Health | Anti-diarrhea | 82.70 |
| 5940 | Digestive Health | Laxatives | 265.29 |
| 5945 | Digestive Health | Lactose Intolerance | 22.14 |
| 5950 | Digestive Health | Rectal/Hemmorhoidal | 58.79 |
| 5955 | Digestive Health | Pediatric Laxatives | 31.57 |
| 6001 | Eye/Ear Care | Soft Contact Lens Care | 155.16 |
| 6005 | Eye/Ear Care | Rigid Gas Permeable Contact Lens Care | 18.55 |
| 6010 | Eye/Ear Care | General Eye Care | 203.62 |
| 6040 | Eye/Ear Care | Reading Glasses | 71.66 |
| 6042 | Eye/Ear Care | Sunglasses | 43.87 |
| 6045 | Eye/Ear Care | Misc. Eye Glass Accessories | 15.28 |

| | | | |
|---|---|---|---|
| 6050 | Eye/Ear Care | Ear Care/Ear Plugs | 33.25 |
| 6101 | Foot Care | Insoles/Inserts | 75.90 |
| 6105 | Foot Care | Corns/Callous/Padding/Bunion/Blister | 28.88 |
| 6110 | Foot Care | Odor/Wetness Control | 19.64 |
| 6115 | Foot Care | Anti-Fungal/Athlete's Foot | 107.49 |
| 6120 | Foot Care | Jock Itch | 20.22 |
| 6130 | Foot Care | Wart Removers | 37.76 |
| 6190 | Foot Care | Grooming and Misc. Foot Care | 12.70 |

Note: Weekly revenue statistics based on sales in calendar year 2005.

**Appendix Figure 1**
Excess Burden of Taxation with Pre Existing Taxes




NOTE–This figure depicts the excess burden of increasing the sales tax by $\Delta t$ starting from initial tax rates of $(t_0^E, t_0^S)$ when utility is quasilinear and prices are endogenous (Proposition 4ii). The figure plots three Marshallian demand curves as a function of the pre-tax price: (1) $x(p,0)$ – the price-demand curve absent taxes, which allows us to recover true preferences; (2) $x(p + t_0^E, t_0^S)$ – the initial demand curve prior to the tax increase; and (3) $x(p + t_0^E + \Delta t, t_0^S)$ – the demand curve after the tax increase. The figure also depicts demand with only a pre-existing excise tax $x_0^E = x(p_0^E + t_0^E, 0)$, which is the point from which the approximations are made to calculate the area of the trapezoid.

## APPENDIX EXHIBIT 1: CLASSROOM SURVEY

### Short survey on spending patterns

Major: _____

Year: _____

Gender: _____

Choose two items from the image projected on the screen.

Number of item #1: _____

Number of item #2: _____

Total bill due at the register for these two items:

$_____

# APPENDIX EXHIBIT 2: TAX SURVEY

University of California, Berkeley
Department of Economics

This survey is part of a project about taxes being conducted by researchers at UC Berkeley. Your identity will be kept strictly confidential and will not be used in the research. If you have any questions about your rights or treatment as a participant in this research project, please contact UC-Berkeley's Committee for Protection of Human Subjects at (510) 642-7461, or e-mail: subjects@berkeley.edu.

| **Gender**: | **Age**: | **Marital Status**: | **Education**: | **Years You Have Lived in California:** |
|---|---|---|---|---|
| ☐ Male | | ☐ Married | ☐ High School | |
| ☐ Female | | ☐ Unmarried | ☐ College Degree | |
| | | | ☐ Graduate Degree | |

| Is tax added **at the register** (in addition to the price posted on the shelf) for each of the following items? | | | | Have you purchased these items within the last month? | | | |
|---|---|---|---|---|---|---|---|
| milk | Y N | toothpaste | Y N | milk | Y N | toothpaste | Y N |
| magazines | Y N | soda | Y N | magazines | Y N | soda | Y N |
| beer | Y N | cookies | Y N | beer | Y N | cookies | Y N |
| potatoes | Y N | cigarettes | Y N | potatoes | Y N | cigarettes | Y N |

What is the sales tax rate in [city]? _____%

What is the California **state** income tax rate in the highest tax bracket? _____ %

What percentage of families in the US do you think pay the federal estate tax when someone dies?

< 2% 2-10% 10-25% 25-50% > 50%

Thank you for your time!