

## Saliency and Taxation: Theory and Evidence

By RAJ CHETTY, ADAM LOONEY, AND KORY KROFT\*

*Using two strategies, we show that consumers underreact to taxes that are not salient. First, using a field experiment in a grocery store, we find that posting tax-inclusive price tags reduces demand by 8 percent. Second, increases in taxes included in posted prices reduce alcohol consumption more than increases in taxes applied at the register. We develop a theoretical framework for applied welfare analysis that accommodates saliency effects and other optimization failures. The simple formulas we derive imply that the economic incidence of a tax depends on its statutory incidence, and that even policies that induce no change in behavior can create efficiency losses. (JEL C93, D12, H25, H71)*

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 that suggests that individuals are inattentive to some types of incentives.<sup>1</sup> 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).

\* Chetty: Department of Economics, University of California, Berkeley, 549 Evans Hall, #3880, Berkeley, CA 94720–3880 (e-mail: [chetty@econ.berkeley.edu](mailto:chetty@econ.berkeley.edu)); Looney: Division of Research and Statistics, Board of Governors of Federal Reserve System, 20th St. and Constitution Avenue, NW, Washington, DC 20551 (e-mail: [adam.looney@frb.gov](mailto:adam.looney@frb.gov)); Kroft: Department of Economics, University of California, Berkeley, 549 Evans Hall, #3880, Berkeley, CA 94720–3882 (e-mail: [kroft@econ.berkeley.edu](mailto:kroft@econ.berkeley.edu)). 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 Köszegi, 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 National Science Foundation grant SES 0452605 and the Hoover Institution. 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. Appendices and stata code for the empirical analysis are available on Chetty's Web site.

<sup>1</sup> Recent studies of inattentive behavior include prices versus shipping fees (Tanjim Hossain and John Morgan 2006; Jennifer Brown, Hossain, and Morgan 2008), financial markets (Stefano DellaVigna and Joshua Pollet 2009), and rebates for car purchases (Meghan Busse, Jorge Silva-Risso, and Florian Zettelmeyer 2006). See DellaVigna (2009) for additional examples and a review of this literature.

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.<sup>2</sup> 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: an experiment in a grocery store, and an observational study of the effect of alcohol 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.<sup>3</sup> To test if people underreact 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 difference-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 United States: an excise tax 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

<sup>2</sup> 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 that depend on the tax rate itself.

<sup>3</sup> 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.



EXHIBIT 1. TAX-INCLUSIVE PRICE TAGS

Photo: Raj Chetty

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 underreact 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 seven out of the eight 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.

Our empirical results contradict the basic assumptions of the neoclassical models currently used to guide tax policy. To understand the implications of the empirical evidence for tax policy, we need a method of characterizing the welfare consequences of taxation when agents do not optimize perfectly relative to taxes. The objective of the second half of the paper is to develop such a method. The main challenge we confront, which is central to behavioral public economics more generally, is the recovery of true preferences when behavior is inconsistent with full optimization. We characterize the welfare consequences of taxation using an approach that does not rely on a specific positive model of behavior, as in B. Douglas Bernheim and Antonio Rangel (2009). Our method relies on two assumptions: taxes affect welfare only through their effects on the consumption bundles chosen by agents; and consumption choices, when prices are perfectly salient, are optimal. Under these assumptions, we derive formulas for the effect of taxes on social surplus (deadweight loss) and distribution (incidence) that depend only on the empirically observed demand function and not on the underlying model that 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 as a function of (nonsalient) taxes, 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 preferences and calculate welfare changes because it is generated by optimizing behavior.

The benefits of this approach to welfare analysis are its simplicity and adaptability. The formulas for deadweight loss and incidence can be derived using supply and demand diagrams and familiar notions of consumer and producer surplus. The formulas differ from Harberger's (1964) widely applied formulas 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.<sup>4</sup> 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 Katusčák 2006) or overestimation 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 about individuals' tax perceptions or rules of thumb.

The results of the welfare analysis 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 an efficiency cost even when demand for the taxed good does not change by distorting budget allocations. 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.

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 (2009) 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 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

<sup>4</sup> Our formulas do not, however, permit errors in optimization relative to salient prices. Such errors can be accommodated by identifying an environment where the true price elasticity is revealed and by applying the same formulas.

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 versus 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 identification problems, which could explain why our use of 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 underreact 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.<sup>5</sup> Normalize the price of  $y$  to one and let  $p$  denote the pretax price of  $x$ . Assume that  $y$  is untaxed and  $x$  is subject to an *ad 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:  $\varepsilon_{x,p} \equiv -(\partial \log x)/(\partial \log p) = \varepsilon_{x,1+\tau^s} \equiv -(\partial \log x)/((\partial \log (1 + \tau^s)))$ . We hypothesize that in practice consumers underreact 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:

$$(1) \quad \log x(p, \tau^s) = \alpha + \beta \log p + \theta_\tau \beta \log (1 + \tau^s).$$

In this equation, the parameter  $\theta_\tau$  measures the degree to which agents underreact to the tax.<sup>6</sup> 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 ( $\varepsilon_{x,p} = -\beta$ ):

$$(2) \quad \theta_\tau = \frac{\partial \log x}{\partial \log (1 + \tau^s)} \bigg/ \frac{\partial \log x}{\partial \log p} = \frac{\varepsilon_{x,1+\tau^s}}{\varepsilon_{x,p}}.$$

<sup>5</sup> 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 because each state accounts for a small share of the national market.

<sup>6</sup> 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 (Chetty, Looney, and 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.



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

**STRATEGY 1: *Manipulate Tax Saliency*.** 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 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  $\theta_\tau$ :

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

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

### 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.<sup>7</sup> 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: products

<sup>7</sup> We estimate that the loss in revenue due to our experiment was about \$300 (8 percent of \$3,900). 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.

with relatively high prices so that the dollar amount of the sales tax is nontrivial; and 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 February 22, 2006, and ending March 15, 2006.<sup>8</sup> 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:  $\$p + \text{Sales Tax} = \$q$ ,” 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 Web Appendix A (available at <http://www.aeaweb.org/articles.php?doi=10.1257/aer.99.4.53>).

*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 Web Appendix Exhibit 1) asking each student to choose two goods and write down “the total bill 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 able only 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

<sup>8</sup> 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.

TABLE 1—SURVEY EVIDENCE: SUMMARY STATISTICS

	Mean	Median	Standard deviation
<i>Panel A. Classroom survey</i>			
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
<i>t</i> -test for equality of means: $p < 0.001$			
$N = 49$			
<i>Panel B. Grocery store survey</i>			
Local sales tax rate	7.48	7.39	0.80
(Actual rate is 7.375 percent)			
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 Web Appendix Exhibit 1 for survey instrument. Panel B reports summary statistics for a survey of 91 customers at the treatment grocery store. See Web Appendix Exhibit 2 for survey instrument.

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 Web Appendix Table 1 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 Web Appendix Table 2. Using these two control groups, we implement a standard difference-in-differences methodology to test whether sales of the treated products fell during the intervention relative to the controls.

### 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 13 of 2006. The dataset contains weekly information on price and quantity sold for all toiletry (treatment and control) products in each store. See Web 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 Web 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



TABLE 2—GROCERY EXPERIMENT: SUMMARY STATISTICS

	Treatment store		Control stores		Total
	Treatment products (1)	Control products (2)	Treatment products (3)	Control products (4)	All stores and products (5)
<i>Panel A. Category-level statistics</i>					
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
<i>Panel B. Product-level statistics</i>					
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 (1.7)	\$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 (1.9)	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 Web Appendix Table 2 for complete list. Product price reflects actual price paid, including any discount if product is on sale. See Web Appendix A for data sources and sample definition.

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).

### C. Results

*Comparison of Means.*—We begin our analysis with a cross tabulation of mean quantity sold (see Table 3). The upper panel of the table shows data for the treatment store. The data are split into four cells. The rows split the data by time: preexperiment (week 1 of 2005 to week 6 of 2006) versus the intervention period (weeks 8 to 10 of 2006).<sup>9</sup> The columns split the data by product group: treated versus control categories. Each cell shows the mean quantity sold for the group

<sup>9</sup> In the week before the experiment (week 7 of 2006), the store asked us to conduct a pilot to ensure that our team could place the tags successfully without disrupting business. For a subset of the treated products, we posted tags which said “This product is subject to sales tax” but did not show tax-inclusive prices. To avoid bias, we exclude this pilot

TABLE 3—EFFECT OF POSTING TAX-INCLUSIVE PRICES: DDD ANALYSIS OF MEAN QUANTITY SOLD

Period	Control categories	Treated categories	Difference
<i>Panel A. Treatment store</i>			
Baseline (2005:1–2006:6)	26.48 (0.22) [5,510]	25.17 (0.37) [754]	–1.31 (0.43) [6,264]
Experiment (2006:8–2006:10)	27.32 (0.87) [285]	23.87 (1.02) [39]	–3.45 (0.64) [324]
Difference over time	0.84 (0.75) [5,795]	–1.30 (0.92) [793]	$DD_{TS} = -2.14$ (0.68) [6,588]
<i>Panel B. Control stores</i>			
Baseline (2005:1–2006:6)	30.57 (0.24) [11,020]	27.94 (0.30) [1,508]	–2.63 (0.32) [12,528]
Experiment (2006:8–2006:10)	30.76 (0.72) [570]	28.19 (1.06) [78]	–2.57 (1.09) [648]
Difference over time	0.19 (0.64) [11,590]	0.25 (0.92) [1,586]	$DD_{CS} = 0.06$ (0.95) [13,176]
<i>DDD Estimate</i>			–2.20 (0.59) [19,764]

*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.

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.<sup>10</sup>

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 preperiod 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-differences 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

week throughout the analysis reported in the paper. However, none of the results is affected by extending the preperiod to include this week.

<sup>10</sup> Standard errors are similar when we cluster by category to adjust for serial correlation.

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 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.

We gauge the magnitude of the treatment effect using 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  $\theta_\tau = 1 - (7.6/(1.59 \times 7.375)) = 0.35$ . 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 (see 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):

$$(4) \quad Y = \alpha + \beta_1 TT + \beta_2 TS + \beta_3 TC + \gamma_1 TT \times TC + \gamma_2 TT \times TS + \gamma_3 TS \times TC \\ + \delta TT \times TC \times TS + \xi X + \varepsilon.$$

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.<sup>11</sup> 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  $\times$  quantity) from the treated products.<sup>12</sup>

<sup>11</sup> 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.

<sup>12</sup> 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 versus \$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.

TABLE 4—EFFECT OF POSTING TAX-INCLUSIVE PRICES: REGRESSION ESTIMATES

Dependent variable	Quantity per category (1)	Revenue per category (\$) (2)	Log quantity per category (3)	Quantity per category (4)	Quantity (treat. categories only) (5)
Treatment	-2.20 (0.60)	-13.12 (4.89)	-0.101 (0.03)	-2.27 (0.60)	-1.55 (0.35)
Average price	-3.15 (0.26)	-3.24 (1.74)		-3.04 (0.25)	-15.06 (3.55)
Average price squared	0.05 (0.00)	0.06 (0.03)		0.05 (0.00)	1.24 (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 three-week period before the experiment and the three-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.

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*: weeks 4 to 6 of 2006) and after the intervention ended (*AT*: weeks 11 to 13). 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 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.

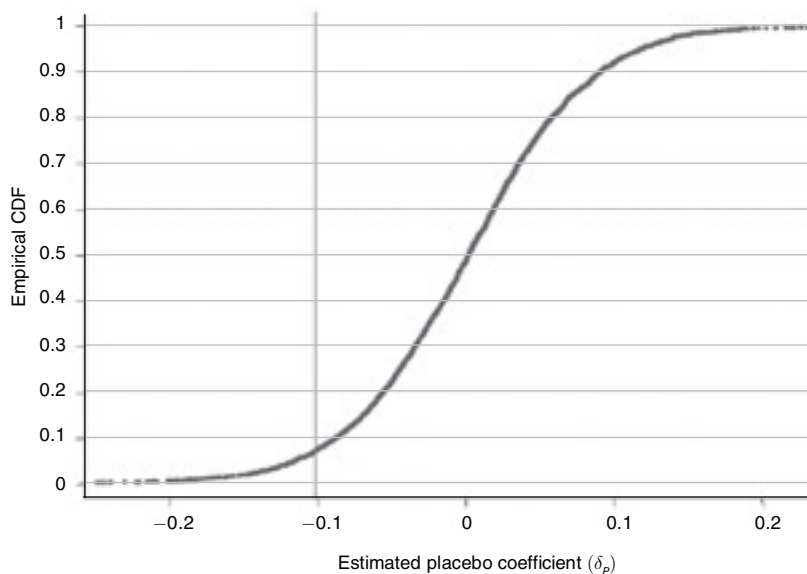


FIGURE 1. DISTRIBUTION OF PLACEBO ESTIMATES: LOG QUANTITY

*Notes:* This figure plots the empirical distribution of placebo effects ( $G$ ) for log quantity. The CDF is constructed from 4,725 estimates of  $\hat{\delta}_p$  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.

A concern in DD analysis is that serial correlation can bias standard errors, leading to over-rejection 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  $G(\hat{\delta}_p)$  to be the empirical cumulative distribution function 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.<sup>13</sup> 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

<sup>13</sup> 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.

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 preexperiment sample to the three months immediately before the intervention.

### III. Evidence from Observational Data on Alcohol Sales

#### 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: an *excise* tax that is levied at the wholesale level and is included in the price posted on the shelf or restaurant menu; and 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):

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

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):

$$(6) \quad \Delta \log x_{jt} = \alpha' + \beta \Delta \log(1 + \tau_{jt}^E) + \theta_\tau \beta \Delta \log(1 + \tau_{jt}^S) + X_{jt} \rho + \varepsilon_{jt}.$$

We estimate variants of (6) using OLS and test the hypothesis that  $\theta_\tau = 1$ .<sup>14</sup> The identification assumption is that the changes in sales and excise taxes are uncorrelated with state-specific shocks to alcohol consumption.

#### 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 United States. As we discuss below, the results for total alcohol consumption are similar, because changes in tax rates across the three types of alcohol are very highly correlated.

We use data on aggregate annual beer consumption by state from the National Institute of Alcohol Abuse and Alcoholism (NIAAA) (2006) from 1970 to 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 (University

<sup>14</sup> 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.



TABLE 5— SUMMARY STATISTICS FOR STATE BEER CONSUMPTION, TAXES, AND REGULATION

Per capita beer consumption (cans)	243.2 (46.1)
State beer excise tax (\$/case)	0.51 (0.50)
State beer excise tax (percent)	6.5 (8.2)
Sales tax (percent)	4.3 (1.9)
Drinking age is 21	0.73 (0.44)
Drunk driving standard	0.65 (0.47)
Any alcohol regulation change	0.19 (0.39)
<i>N</i> (number of state-year pairs)	1,666

*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 or implemented a stricter drunk driving standard, an administrative license revocation law, or a zero tolerance youth drunk driving law. See Web Appendix A for data sources and sample definition.

of Michigan Business School 2006), and other sources.<sup>15</sup> 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.<sup>16</sup> 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 Web 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 6.5 percent of the average price, while the mean state sales tax rate is 4.3 percent.<sup>17</sup> 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.

<sup>15</sup> We exclude West Virginia because of problems with the sales tax rate data described in Web Appendix A. Including West Virginia magnifies the difference between the excise and sales tax elasticities.

<sup>16</sup> 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.

<sup>17</sup> 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.

### C. Results

We begin with a graphical analysis (Figures 2A and 2B) to illustrate the relationship between alcohol consumption and taxes. 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  $\pm 0.02$  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 0.88 percent ( $\varepsilon_{x,1+\tau^E} = 0.88$ ).<sup>18</sup> 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  $p = 0.01$ . 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  $\varepsilon_{x,1+\tau^E} = \varepsilon_{x,1+\tau^S}$ .<sup>19</sup>

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 (column 3). The coefficient on the excise tax rate does not change significantly because regulation changes have modest effects on total

<sup>18</sup> 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.

<sup>19</sup> 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.

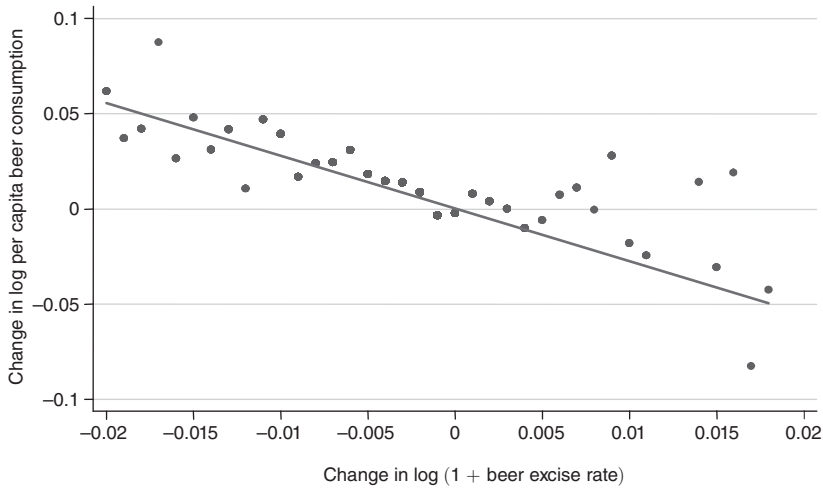


FIGURE 2A. PER CAPITA BEER CONSUMPTION AND STATE BEER EXCISE TAXES

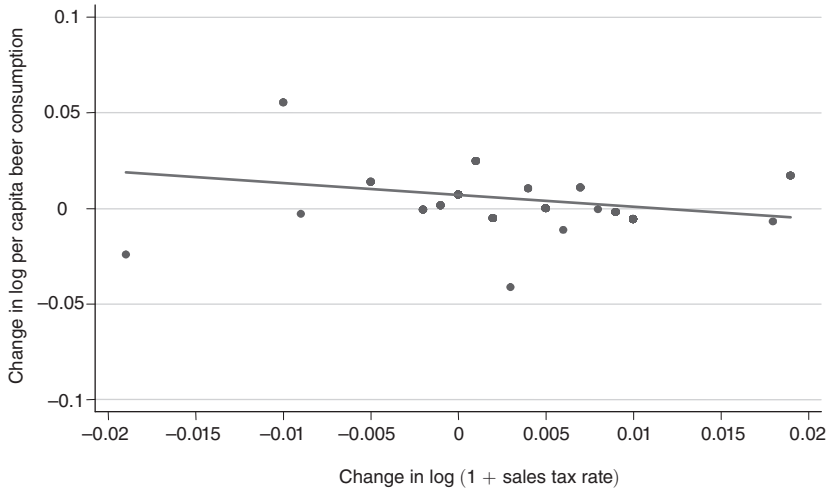


FIGURE 2B. PER CAPITA BEER CONSUMPTION AND STATE SALES TAXES

*Notes:* 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 percent), 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 percent. See Web Appendix A for data sources and sample definition.

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

TABLE 6—EFFECT OF EXCISE AND SALES TAXES ON BEER CONSUMPTION

	Baseline (1)	Business cycle (2)	Alcohol regulations (3)	Region trends (4)
Dependent variable: Change in log (per capita beer consumption)				
$\Delta \log(1 + \text{excise tax rate})$	-0.88 (0.17)	-0.91 (0.17)	-0.89 (0.17)	-0.71 (0.18)
$\Delta \log(1 + \text{sales tax rate})$	-0.20 (0.30)	-0.01 (0.30)	-0.02 (0.30)	-0.05 (0.30)
$\Delta \log(\text{population})$	0.03 (0.06)	-0.07 (0.07)	-0.07 (0.07)	-0.09 (0.08)
$\Delta \log(\text{income per capita})$		0.22 (0.05)	0.22 (0.05)	0.22 (0.05)
$\Delta \log(\text{unemployment rate})$		-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
<i>F</i> -test for equality of tax elasticities (prob > <i>F</i> )	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 are 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.

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: policy changes in the nominal tax rate, which produce sharp jumps in tax rates; and gradual erosion of the nominal value of the tax by inflation, which creates differential changes in excise tax rates across states because they have different initial tax rates.<sup>20</sup> 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.

<sup>20</sup> 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  $(1.50 - 1.25)/1.50 = 17$  percent in state *A* but is unchanged in state *B*.

TABLE 7—EFFECT OF EXCISE AND SALES TAXES ON BEER CONSUMPTION: ROBUSTNESS CHECKS

	Dependent variable: Change in log (per capita beer consumption)				Dep. var.
	IV for excise w/ policy (1)	3-Year differences (2)	IV for ACCRA beer price (3)	Food exempt (4)	Share ethanol from beer (5)
$\Delta \log(1 + \text{excise tax rate})$	-0.63 (0.28)	-1.10 (0.47)		-0.91 (0.22)	0.16 (0.13)
$\Delta \log(\text{beer price})$			-0.88 (0.42)		
$\Delta \log(1 + \text{sales tax rate})$	-0.03 (0.30)	-0.00 (0.33)	0.10 (0.59)	-0.14 (0.30)	0.25 (0.22)
$\Delta \log(\text{population})$	-0.06 (0.07)	-1.24 (0.33)	-0.06 (0.15)	0.03 (0.07)	0.09 (0.05)
$\Delta \log(\text{income per capita})$	0.22 (0.05)	0.08 (0.05)	0.23 (0.10)	0.22 (0.05)	0.01 (0.03)
$\Delta \log(\text{unemployment rate})$	-0.01 (0.01)	-0.00 (0.01)	0.00 (0.01)	-0.01 (0.01)	0.00 (0.01)
Alcohol regulation controls	x	x	x	x	x
Year fixed effects	x	x	x	x	x
<i>F</i> -test for equality of tax elasticities (prob > <i>F</i> )	0.15	0.05	0.12	0.04	0.73
Sample size	1,487	1,389	825	937	1,487

*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.

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 underreact to taxes that are not salient, even in the long run.

The premise underlying our analysis is that firms pass changes in the state excise tax through to consumers so that they are reflected in the posted price of beer. We expect full pass through of state-level tax changes because each state constitutes a small share of the national market, effectively making the state-level supply curves flat. 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 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 as expected—the coefficient on the excise tax variable in the first-stage regression is approximately one. 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 versus 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.<sup>21</sup> 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 30 states that fully exempted all food items from the sales tax in 2000.<sup>22</sup> 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  $(1.01/1.004) - 1 \simeq 0.6$  percent. It follows that the effect of a 1 percent increase in the tax  $\tau^A$  that applies solely to alcohol is given by  $\varepsilon_{x,1+\tau^A} = (1/0.6)\varepsilon_{1,1+\tau^S}$ . Scaling up the largest estimated response to the sales tax in Table 6 of  $-0.20$  by  $5/3$  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 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 changes in beer excise taxes on *total* ethanol consumption is much larger than the effect of changes in sales tax rates. 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  $5/3$ , 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  $\theta_\tau = \varepsilon_{x,1+\tau^A}/\varepsilon_{x,1+\tau^E} = 0.06$ .

<sup>21</sup> In 2004, sales tax revenues were 2.1 percent 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.

<sup>22</sup> We do not have historical data on which goods are subject to the sales tax. Case studies of some states suggest, however, that the set of items subject to sales tax is fairly stable over time.



#### IV. Why Do Consumers Underreact to Taxes?

There are two potential explanations for the finding that consumers underreact 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 preexperiment levels after the intervention ended, suggesting that there were no persistent learning effects.<sup>23</sup> Third, the finding that sales and excise tax elasticities for alcohol demand do not converge over time suggests that the underreaction 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 Web Appendix A for details on survey implementation and Appendix Exhibit 2 for the survey instrument. We asked individuals what the local sales tax rate was 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. Seventy-five 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. Coca Cola and cookies accounted for more than half the mistakes among respondents who got one question wrong.

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 underreaction 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  $x_0 = \$1,000$  and 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  $\varepsilon_{x,p} \gg \varepsilon_{x,1+\tau^s}$ .

More generally, agents with limited attention may use heuristics to achieve a consumption allocation that approximates the fully optimal bundle, but leads them to underreact to taxes. For

<sup>23</sup> We cannot rule out another equally plausible explanation of this finding: the set of individuals who shop for these durable goods is likely to vary substantially across weeks, so customers in the weeks after the experiment may have been untreated.

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 do not allow us to distinguish between these models. We therefore proceed to analyze the welfare consequences of taxation in a manner that does not depend on a specific positive theory of underreaction to taxes.

## V. Welfare Analysis

This section explores the implications of our empirical results for tax policy. In particular, we generalize Harberger's (1964) canonical partial-equilibrium formulas for incidence and dead-weight loss to allow for salience effects and other optimization errors with respect to taxes. The formulas we develop can be used to analyze the effects of taxes in the specific commodity markets that we analyzed empirically, as well as other policies such as labor and capital income taxation.

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).<sup>24</sup> The tools developed below can be adapted to analyze Pigouvian taxes intended to correct behavior, but we defer that analysis to future work.

### A. Setup

We use the same two-good model as in Section I, but assume from this point onward that the sales tax levied on good  $x$  is a specific (unit) tax  $t^s$  rather than an *ad valorem* tax for consistency with the theoretical literature on commodity taxation.<sup>25</sup> Let  $p$  denote the pretax price of  $x$  and  $q = p + t^s$  denote the tax-inclusive price of  $x$ . Good  $y$ , the numeraire, is untaxed. As is standard in partial equilibrium analysis, assume that 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  $u(x) + v(y)$ . Let  $(x^*(p, t^s, Z), y^*(p, t^s, Z))$  denote the bundle chosen by a fully optimizing consumer as a function of the posted price, tax, and wealth. Full optimization implies  $\partial x^*/\partial p = \partial x^*/\partial t^s$ , contradicting our empirical findings. Let  $(x(p, t^s, Z), y(p, t^s, Z))$  denote the empirically observed demand functions, which permit  $\partial x/\partial p \neq \partial x/\partial t^s$ . We do not place structure on the positive model that generates  $(x(p, t^s, Z), y(p, t^s, Z))$  other than to assume that the demand functions are smooth and that the choices are feasible:

$$(p + t^s)x(p, t^s, Z) + y(p, t^s, Z) = Z.$$

<sup>24</sup> We focus 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).

<sup>25</sup> The incidence and excess burden of an *ad valorem* tax  $\tau^s$  can be calculated by replacing  $t^s$  by  $\tau^s$  and  $\partial x/\partial t^s$  by  $\partial x/\partial \tau^s$  in Propositions 1 and 2.

Define the degree of underreaction to the specific tax  $t^s$  as

$$\theta = \frac{\partial x(p, t^s, Z)}{\partial t^s} \bigg/ \frac{\partial x(p, t^s, Z)}{\partial p} = \frac{\varepsilon_{x,q|t^s}}{\varepsilon_{x,q|p}},$$

where  $\varepsilon_{x,q|t^s} = (-\partial x/\partial t^s)(q/x(p, t^s, Z))$  measures the percentage change in demand caused by a 1 percent increase in the total price of good  $x$  through a tax change, and  $\varepsilon_{x,q|p} = (-\partial x/\partial p) \times (q/x(p, t^s, Z))$  represents the analogous measure for a 1 percent increase in  $q$  through a change in  $p$ .<sup>26</sup> 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  $\partial x/\partial t^s$  to differ from  $\partial x/\partial p$  for any reason, not just inattention.<sup>27</sup> The formulas derived below therefore account for any errors that consumers may make when optimizing with respect to taxes.

*Production.*—Price-taking firms use  $c(S)$  units of the numeraire  $y$  to produce  $S$  units of  $x$ . The marginal cost of production is weakly increasing:  $c'(S) > 0$  and  $c''(S) \geq 0$ . The representative firm's profit at pretax price  $p$  and level of supply  $S$  is  $pS - c(S)$ . Assuming that firms optimize perfectly, the supply function for good  $x$  is implicitly defined by the marginal condition  $p = c'(S(p))$ .<sup>28</sup> Let  $\varepsilon_{S,p} = (\partial S/\partial p)(p/S(p))$  denote the price elasticity of supply.

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

$$(7) \quad D(p, t^s, Z) = S(p).$$

Implicit differentiation of (7) yields the following results.

**PROPOSITION 1:** *The incidence on producers of increasing  $t^s$  is*

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

<sup>26</sup> The empirical estimates of  $\theta_\tau$  can be directly mapped to values for  $\theta$  using the equation  $\theta_\tau = \theta(1 + t^s)/(1 + \theta t^s)$ . The reason that  $\theta_\tau < \theta$  is that agents underreact 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  $\theta$ .

<sup>27</sup> 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 overestimate tax rates on wealth (Slemrod 2006).

<sup>28</sup> 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 2009). 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  $\partial S/\partial p$  and  $\partial S/\partial t^s$ .

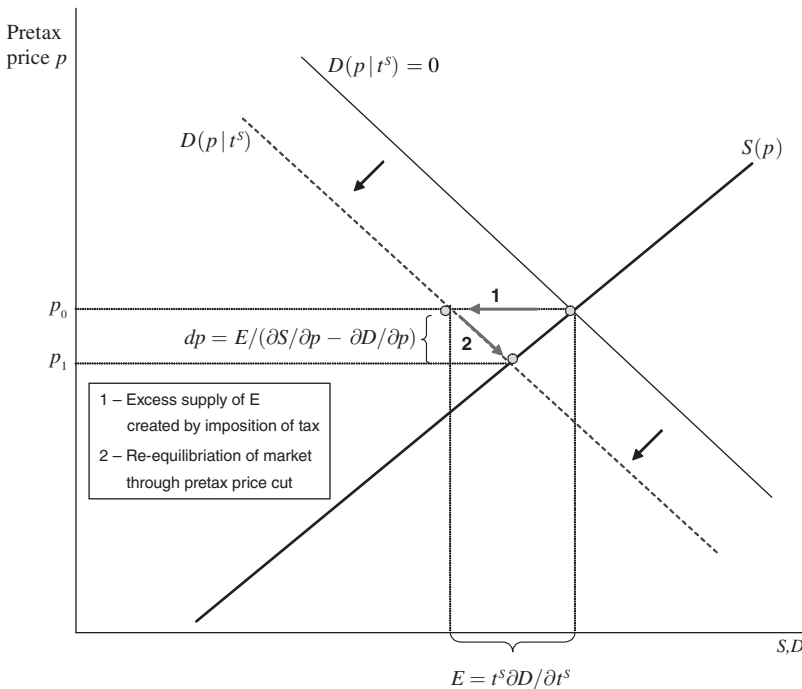


FIGURE 3. INCIDENCE OF TAXATION

Notes: 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 pretax 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 reequilibrate the market, producers cut the pretax price by  $E/(\partial S/\partial p - \partial D/\partial p)$  units, implying  $dp/dt^s = (\partial D/\partial t^s)/(\partial S/\partial p - \partial D/\partial p)$ .

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  $(p, t^s, Z)$ , 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 reequilibrate 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.

**Attenuated Incidence on Producers.** Incidence on producers is attenuated by  $\theta = (\partial D/\partial t^s)/(\partial D/\partial p)$  relative to the traditional model. Intuitively, producers face less pressure to reduce the pretax price when consumers underreact 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  $\varepsilon_{D,q|p} = \theta\varepsilon_{D,q|p}$  in the same way, their effects on incidence are not equivalent. To see this, consider two markets, *A* and *B*, where  $\varepsilon_{s,p}^A = \varepsilon_{s,p}^B = 0.1$ . In market *A*, demand is inelastic and consumers are fully attentive to taxes:  $\varepsilon_{D,q|p}^A = 0.3$  and  $\theta^A = 1$ . In market *B*, demand is elastic but consumers are inattentive:  $\varepsilon_{D,q|p}^B = 1$  and  $\theta^B = 0.3$ . An econometrician would estimate the same tax elasticity in both markets:  $\varepsilon_{D,q|p}^A = \varepsilon_{D,q|p}^B = 0.3$ . However,  $[dp/dt^s]^A = -0.75$  whereas  $[dp/dt^s]^B = -0.27$ . In market *A*, suppliers bear most of the incidence since demand is three times more elastic to price than supply. In market *B*, even though demand is ten 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 reequilibrate 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  $dp/dt$  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,  $dq/dt^s$  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.

**No Tax Neutrality.** Taxes that are included in posted prices—such as the alcohol excise tax—have greater incidence on producers because they are fully salient ( $\theta = 1$ ). 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 imply 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.<sup>29</sup>

**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 ( $\partial[dp/dt^s]/\partial\varepsilon_{D,q|p} > 0$ ). 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  $\partial[dp/dt^s]/\partial\varepsilon_{D,q|p} < 0$  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  $\varepsilon_{D,q|p}$ .

<sup>29</sup> 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.

### C. Efficiency Cost

In the interest of space, we formally characterize the excess burden of introducing a sales tax in a market where there are no preexisting taxes and production is constant-returns-to-scale ( $c'' = 0$ ). In this case, the pretax price of  $x$  is fixed at  $p = c'(0)$ . Moreover, since firms earn zero profits, only consumer welfare matters for excess burden. At the end of this section, we briefly discuss the effects of endogenous producer prices and preexisting taxes.

*Definitions.*—Let  $V(p, t^s, Z) = u(x(p, t^s, Z)) + v(y(p, t^s, Z))$  denote the agent's indirect utility as a function of the posted price of good  $x$ , the sales tax, and wealth. Let  $e(p, t^s, V)$  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  $R(p, t^s, Z) = t^s x(p, t^s, Z)$  denote tax revenue.

Following Herbert Mohring (1971) and Alan J. Auerbach (1985), we measure the excess burden (deadweight cost) of a tax using the concept of equivalent variation. When  $p$  is fixed, the excess burden of introducing a sales tax  $t^s$  in a previously untaxed market is

$$(9) \quad EB(t^s) = Z - e(p, 0, V(p, t^s, Z)) - R(p, t^s, Z).$$

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 a simple expression for (9) in terms of empirically estimable elasticities.

*Preference Recovery.*—The efficiency cost of a tax policy depends on two elements: the change in behavior induced by the tax; and the effect of that change in behavior on the consumer's utility. The first element is observed empirically. The second element is the key challenge for behavioral welfare economics. How do we compute indirect utility  $V(p, t^s, Z)$  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  $x(p, t^s, Z)$ .

**A1:** Taxes affect utility only through their effects on the chosen consumption bundle. The agent's indirect utility given a tax of  $t^s$  is

$$V(p, t^s, Z) = u(x(p, t^s, Z)) + v(y(p, t^s, Z)).$$

**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.<sup>30</sup>

Assumption A2 requires that the agent behave 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 Kőszegi, 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  $x(p, t^s, Z)$  to calculate the agent's indirect utility  $V(p, t^s, Z)$  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 (2009) 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 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 specific structure on the model that generates  $x(p, t^s, Z)$ .

*Formula for Excess Burden.*—We derive a formula for excess burden using quadratic approximations analogous to those used by Harberger (1964) and Edgar K. Browning (1987). 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  $\partial x^c / \partial p < 0$  through the Slutsky condition. Let  $\varepsilon_{x,q|p}^c = -(\partial x^c / \partial p)(q/x)$  and  $\varepsilon_{x,q|t^s}^c = -(\partial x^c / \partial t^s)(q/x)$  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*

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

<sup>30</sup> 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.

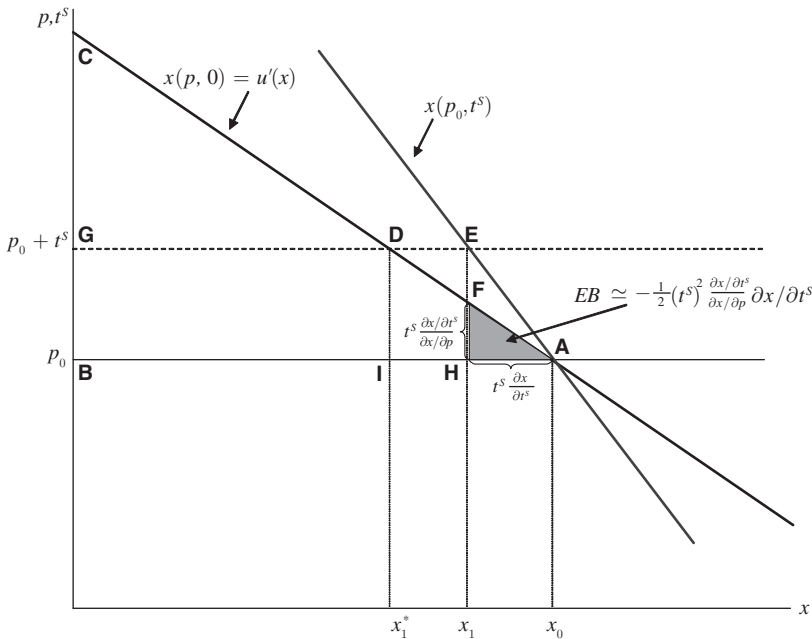


FIGURE 4. EXCESS BURDEN WITH NO INCOME EFFECT FOR GOOD  $x$  ( $\partial x/\partial Z = 0$ )

Notes: This figure illustrates the deadweight cost of introducing a tax  $t^s$  levied on consumers when  $\partial x/\partial Z = 0$  and producer prices are fixed. The figure plots two demand curves: the price-demand curve  $x(p, 0)$ , which shows how demand varies with the pretax price of the good; and 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  $AFH$ .

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

**PROOF:**

See Web Appendix B. Chetty (2009) gives an instructive proof for the case without income effects ( $\partial x/\partial Z = 0$ ).

Figure 4 illustrates the calculation of deadweight loss for the case without income effects ( $\partial x/\partial Z = 0$ ), which implies utility is quasilinear. 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  $u'(x)$  under Assumption A2. The second,  $x(p_0, t^s)$  represents how demand varies with the tax on  $x$ . This tax-demand curve is drawn assuming  $\partial x/\partial p < \partial x/\partial t^s$ , 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). 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  $AFH$ , whose area is given by (10).<sup>31</sup>

When there are income effects ( $\partial x/\partial Z > 0$ ), 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 the difference between the actual behavioral response ( $\partial x/\partial t^s$ ) and the socially optimal response given the reduction in net-of-tax income ( $-x(\partial x/\partial Z)$ ), which is  $(\partial x^c/\partial t^s) - (-x(\partial x/\partial Z)) = \partial x^c/\partial t^s$ .

Like the Harberger formula, (10) ignores the third- and higher-order terms in the Taylor expansion for  $EB$ . Hence, it provides an accurate measure of excess burden for small tax changes. In addition, note that  $\partial x^c/\partial p$  must be 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 ( $d^2x^c/dpdt^s = 0$ ), and using an estimate of  $dx^c/dp(p, t^s, Z) = dx^c/dp(p, 0, Z)$ .

*Discussion.*—The only difference between (10) and the canonical Harberger formula ( $EB^*(t^s) = -(\frac{1}{2})(t^s)^2(\partial x^c/\partial t^s)$ ) is the introduction of the parameter  $\theta^c = (\partial x^c/\partial t^s)/(\partial x^c/\partial p)$ . Three general lessons about excess burden emerge from this new parameter.

**Inattention Reduces Excess Burden if  $\partial x/\partial Z = 0$ .** When there are no income effects, the tax  $t^s$  generates deadweight loss equivalent to that created by a perfectly salient tax of  $\theta t^s$ . If agents ignore the tax 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.<sup>32</sup> As the degree of attention to the tax 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  $\partial x/\partial Z = 0$ .

**Inattention Can Raise Excess Burden if  $\partial x/\partial Z > 0$ .** When there are income effects, making a tax less salient to reduce  $\partial x/\partial t^s$  can *increase* deadweight loss. In fact, a tax can create deadweight cost even if the agent completely ignores it and demand for the taxed good does not

<sup>31</sup> 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: subtracting the additional revenue earned by the government because the agent underreacts to the tax (rectangle  $HIDE$ ); and adding the *private* welfare loss due to the optimization error (triangle  $FED$ ).

<sup>32</sup> 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  $\partial x/\partial Z = 0$ .

change, i.e.,  $\partial x/\partial t^s = 0$ . This result contradicts the canonical 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  $\partial x/\partial p < 0$ , since  $\partial x/\partial p = 0 \Rightarrow \partial x^c/\partial p = 0$  given  $\partial x/\partial Z > 0$ . This reasoning fails when the tax demand is not the outcome of perfect optimization, because there is no Slutsky condition for  $\partial x^c/\partial t^s$ . A zero uncompensated tax elasticity does not imply that the compensated tax elasticity is zero. Instead, when  $\partial x/\partial t^s = 0$ ,  $\partial x^c/\partial t^s = \partial x/\partial Z$  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 quasi-linear 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 quasi-linear, 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  $\partial x^c/\partial t^s = \partial x/\partial Z$  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 ( $\partial x/\partial t^s = 0$ ) 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  $\partial x/\partial Z$  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 that "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  $\partial x/\partial t^s = 0$ . The effect of inattention on  $\partial x/\partial t^s$  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:  $\partial x/\partial t^s + \partial y/\partial t^s = -x$ . 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  $\partial x/\partial t^s = \theta = 0$  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 response  $\partial x/\partial t^s = -x(\partial x/\partial Z)$ , 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  $\partial x^c/\partial t^s$ .

**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  $\varepsilon_{x,q|p}$ .

It is straightforward to extend the preceding results to allow for preexisting taxes and endogenous producer prices; see Chetty (2009) for a complete analysis and discussion of these cases. When  $p$  is fixed and the initial sales tax rate is  $t_0^s$ , the excess burden of a sales tax increase  $\Delta t$  is approximately

$$(11) \quad EB(\Delta t | t_0^s) \simeq \theta^c x_0 \frac{\varepsilon_{x,q|t^s}^c}{q_0} \left( \frac{1}{2} (\Delta t)^2 + t_0^s \Delta t \right),$$

where  $x_0$  denotes the initial demand and  $q_0 = p + t_0^s$  denotes the initial price. This expression, which is simply the Harberger “trapezoid” formula multiplied by  $\theta^c$ , shows that tax increases can have a first-order (large) deadweight cost when there are preexisting taxes. The first-order deadweight cost due to  $t_0^s$  is attenuated by  $\theta^c$  because the deviation from the socially optimal level of  $x$  caused by  $t_0^s$  is proportional to  $\theta^c$ . When  $p$  is endogenous (i.e., supply is upward sloping) and  $\partial x / \partial Z = 0$ , (11) holds with the elasticity  $\varepsilon_{x,q|t^s}$  replaced by  $\varepsilon_{x,q|t^s}^{TOT} = -(dx/dt^s)(q/x(p, t^s))$ . The elasticity  $\varepsilon_{x,q|t^s}^{TOT}$  measures the total change in demand caused by a 1 percent increase in the price  $q = p + t^s$  through an increase in  $t^s$ , taking into account the effect of the endogenous price response.

## VI. Conclusion

This paper has shown empirically that commodity taxes that are included in posted prices reduce demand significantly more than taxes that are not included in posted prices. Individuals appear to be well informed about commodity taxes when their attention is drawn to the topic, suggesting that *salience* is an important determinant of behavioral responses to taxation. The finding that individuals make systematic optimization errors even with respect to relatively simple, linear commodity taxes suggests that more complex policies such as income taxes or transfers could generate very different behavioral responses from those predicted by standard models.<sup>33</sup> Moreover, the standard method of using variation in tax rates as instruments to estimate wage and price elasticities cannot be applied unless the tax is perfectly salient.

Our empirical results contradict the basic assumptions of the canonical theory of taxation used for policy analysis. As an alternative, we have proposed a method of welfare analysis that does not rely on a specific positive model of how agents make choices when faced with taxes. This approach accommodates salience effects as well as other optimization errors with respect to taxes. The formulas we obtain for the incidence and excess burden of commodity taxes are simple variants of those in introductory textbooks and can be easily adapted to analyze other tax policies, such as income or capital taxation. Much as Harberger (1964) identified the compensated price elasticity as the key parameter to be estimated in subsequent work, our analysis identifies the compensated tax and price elasticities ( $\varepsilon_{x,q|t^s}^c$  and  $\varepsilon_{x,q|p}^c$ ) as “sufficient statistics” for empirical studies in behavioral public economics.

A natural next step would be to characterize optimal taxation when agents optimize imperfectly, generalizing the results of Ramsey (1927) and Mirrlees (1971). For this purpose, it will be important to extend the welfare analysis to a general equilibrium model with more than two markets. Combining the formulas developed here with a positive theory of tax salience could be useful in characterizing the optimal structure of the tax system. For example, Chetty, Looney,

<sup>33</sup> In a follow-up study, Chetty and Saez (2009) document similar optimization errors in income taxation and labor supply decisions.

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 suggest 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 and Social Security) have complex features and may induce suboptimal behaviors. One can characterize the welfare consequences of these programs more accurately by estimating behavioral responses to analogous programs whose incentives are more salient. 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

- American Chamber of Commerce Researchers Association.** 1982–2000. *ACCRA Cost of Living Index Quarterly Report, 1982–2000*. Louisville, KY: American Chamber of Commerce Researchers Association.
- 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.
- Atkinson, Anthony B., and Joseph E. Stiglitz.** 1976. “The Design of Tax Structure: Direct versus Indirect Taxation.” *Journal of Public Economics*, 6(1–2): 55–75.
- Auerbach, Alan J.** 1985. “The Theory of Excess Burden and Optimal Taxation.” In *Handbook of Public Economics, Vol. 1*, ed. Alan J. Auerbach and Martin S. Feldstein, 67–127. Amsterdam: Elsevier Science.
- Bernheim, B. Douglas, and Antonio Rangel.** 2009. “Beyond Revealed Preference: Choice-Theoretic Foundations for Behavioral Welfare Economics.” *Quarterly Journal of Economics*, 124(1): 51–104.
- Bertrand, Marianne, Esther Dufo, and Sendhil Mullainathan.** 2002. “How Much Should We Trust Differences-in-Differences Estimates?” National Bureau of Economic Research Working Paper 8841.
- Beer Institute.** Various years. *Brewer's Almanac*. <http://www.beerinstitute.org/statistics.asp?bid=200>.
- Blendon, Robert J., Stephen R. Pelletier, Marcus D. Rosenbaum, and Mollyann Brodie.** 2003. “Tax Uncertainty: A Divided America's Unformed View of the Federal Tax System.” *Brookings Review*, 21: 28–31.
- Blumkin, Tomer, Bradley J. Ruffle, and Yosi Ganun.** 2008. “Are Income and Consumption Taxes Ever Really Equivalent? Evidence from a Real-Effort Experiment.” Munich Personal RePEc Paper 6479.
- Brown, Jennifer, Tanjim Hossain, and John Morgan.** 2008. “Shrouded Attributes and Information Suppression: Evidence from Field Experiments.” Unpublished.
- Browning, Edgar K.** 1987. “On the Marginal Welfare Cost of Taxation.” *American Economic Review*, 77(1): 11–23.
- Busse, Meghan, Jorge Silva-Risso, and Florian Zettelmeyer.** 2006. “\$1,000 Cash Back: The Pass-through of Auto Manufacture Promotions.” *American Economic Review*, 96(4): 1253–70.
- Chetty, Raj.** 2006. “A New Method of Estimating Risk Aversion.” *American Economic Review*, 96(5): 1821–34.
- Chetty, Raj.** 2009. “The Simple Economics of Salience and Taxation.” Unpublished.
- Chetty, Raj, Adam Looney, and Kory Kroft.** 2007. “Salience and Taxation: Theory and Evidence.” National Bureau of Economic Research Working Paper 13330.
- Chetty, Raj, and Emmanuel Saez.** 2009. “Teaching the Tax Code: Earnings Responses to an Experiment with EITC Recipients.” National Bureau of Economic Research Working Paper 14836.
- Commerce Clearing House.** Various years. *State Tax Handbook*. Chicago: Commerce Clearing House Inc.
- 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 Working Paper 11138.



- de Bartolome, Charles A. M.** 1995. "Which Tax Rate Do People Use: Average or Marginal?" *Journal of Public Economics*, 56(1): 79–96.
- DellaVigna, Stefano.** 2009. "Psychology and Economics: Evidence from the Field." *Journal of Economic Literature*, 47(2): 315–72.
- DellaVigna, Stefano, and Joshua M. Pollet.** 2009. "Investor Inattention and Friday Earnings Announcements." *Journal of Finance*, 64(2): 709–49.
- Feldman, Naomi E., and Peter Katusčák.** 2006. "Should the Average Tax Rate be Marginalized?" Center for Economic Research and Graduate Education Economics Institute Working Paper 304.
- Finkelstein, Amy.** 2007. "E-Z Tax: Tax Salience and Tax Rates." National Bureau of Economic Research Working Paper 12924.
- Fisher, Ronald A.** 1922. "On the interpretation of  $\chi^2$  from contingency tables, and the calculation of P. J. Roy." *Journal of the Royal Statistical Society*, 85(1): 87–94.
- Gabaix, Xavier, and David Laibson.** 2006. "Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets." *Quarterly Journal of Economics*, 121(2): 505–40.
- Gruber, Jonathan.** 1994. "The Incidence of Mandated Maternity Benefits." *American Economic Review*, 84(3): 622–41.
- Gruber, Jonathan, and Emmanuel Saez.** 2002. "The Elasticity of Taxable Income: Evidence and Implications." *Journal of Public Economics*, 84(1): 1–32.
- Harberger, Arnold C.** 1964. "The Measurement of Waste." *American Economic Review*, 54(3): 58–76.
- 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.
- Hossain, Tanjim, and John Morgan.** 2006. ". . . Plus Shipping and Handling: Revenue (Non) Equivalence in Field Experiments on eBay." *B E Journal of Economic Analysis and Policy: Advances in Economic Analysis and Policy*, 6(2): 1–27.
- Kerschbamer, Rudolf, and Georg Kirchsteiger.** 2000. "Theoretically Robust but Empirically Invalid? An Experimental Investigation into Tax Equivalence." *Economic Theory*, 16(3): 719–34.
- Kotlikoff, Laurence J., and Lawrence H. Summers.** 1987. "Tax Incidence." In *Handbook of Public Economics*, Vol. 2, ed. Alan J. Auerbach and Martin Feldstein, 1043–92. Amsterdam: Elsevier Science.
- 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.
- Liebman, Jeffrey B., and Richard J. Zeckhauser.** 2004. "Schmeduling." Unpublished.
- Mazar, Nina, Botond Köszegi, and Dan Ariely.** 2008. "Price-Sensitive Preferences." Unpublished.
- Meyer, Bruce D.** 1995. "Natural and Quasi-Experiments in Economics." *Journal of Business and Economic Statistics*, 13(2): 151–61.
- Mill, John S.** 1848. *Principles of Political Economy*. Oxford: Oxford University Press.
- Mirrlees, James A.** 1971. "An Exploration in the Theory of Optimum Income Taxation." *The Review of Economic Studies*, 38(2): 175–208.
- Mohring, Herbert.** 1971. "Alternative Welfare Gain and Loss Measures." *Western Economic Journal*, 9: 349–68.
- National Institute of Alcohol Abuse and Alcoholism (NIAAA).** 2006. "Per Capita Alcohol Consumption." <http://www.niaaa.nih.gov/Resources/DatabaseResources/QuickFacts/AlcoholSales>.
- 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 04–5563.
- Ramsey, Frank P.** 1927. "A Contribution to the Theory of Taxation." *Economic Journal*, 37(145): 47–61.
- Rosen, Harvey S.** 1976. "Taxes in a Labor Supply Model with Joint Wage-Hours Determination." *Econometrica*, 44(3): 485–507.
- Rosenbaum, Paul R.** 1996. "Observational Studies and Nonrandomized Experiments." In *Design and Analysis of Experiments*, ed. S. Ghosh and C. R. Rao, 181–97. Amsterdam: Elsevier Science.
- Slemrod, Joel B.** 2006. "The Role of Misconceptions in Support for Regressive Tax Reform." *National Tax Journal*, 59(1): 57–75.
- Tax Foundation.** Various years. *Special Report: State Tax Rates and Collections*. Washington, DC: Tax Foundation. <http://www.taxfoundation.org>.
- University of Michigan Business School.** 2006. *World Tax Database*. <http://www.bus.umich.edu/OTPR/otpr/introduction.htm>.