Online Appendix to
Attention Variation and Welfare: Theory and Evidence from a Tax Salience Experiment

January 10, 2018

A Appendix to Section 2: Further Results

A.1 Identification from Aggregate Demand Data

Definition A.1. Let $B_\epsilon(p,t) := \{D(p',t')|(p',t') \in [p-\epsilon, p+\epsilon] \times [t-\epsilon, t+\epsilon]\}$. We say that local knowledge of $D(p,t)$ at (p,t) is sufficient to identify the efficiency cost of a small tax change if for each sequence of $\{\Delta_i\}_{i=1}^{\infty}$ converging to zero there is a sequence of $\{\epsilon_i\}_{i=1}^{\infty}$ converging to zero with the property that knowledge of $B_{\epsilon_i}(p,t)$ is sufficient to identify $EB(t+\Delta_i) - EB(t)$.

Proposition A.1. *Suppose that $F(\theta|p,t)$ is degenerate and suppose for simplicity that utility is quasilinear.*

1. *(CLK and Chetty 2009) Suppose that either i) $F(\theta|p,t)$ does not depend on t or that ii) $t=0$. Then local knowledge of $D(p,t)$ is sufficient to identify ΔEB for a small tax change.*

2. *Suppose that $F(\theta|p,t)$ depends on t, and that $t>0$. Then local knowledge of $D(p,t)$ is not sufficient to identify ΔEB for a small tax change. However, full knowledge of $D(p,t)$ is sufficient to identify ΔEB for a small Δt.*

Proposition A.1 shows that when $F(\theta|p,t)$ is degenerate, the demand curve $D(p,t)$ identifies welfare. In fact, when attention does not vary with the tax, the proposition shows that local knowledge of the demand curve is sufficient—a replication of CLK and Chetty et al. (2009) for the case of binary demand. To see the intuition, let θ be the (homogenous) weight placed on the tax. When θ is exogenous to t, it is given by $\frac{D_t}{D_p}$: the extent to which consumers underreact to a change in the tax relative to a change in the posted price.

When θ can depend on t, the ratio $\frac{D_t}{D_p}$ no longer identifies θ. The reason is that a change in the tax also changes attention, so that $\frac{D_t}{D_p} = \theta(t) + \theta'(t)t$. To calculate welfare, however, it is necessary to know both $\theta(t)$ and $\theta'(t)$ separately, as shown in Proposition 5. However, full knowledge of $D(p,t)$ is still sufficient to calculate $\theta(t)$. By definition, to calculate $\theta(t)$, we simply need to find the value Δp such that $D(p+\Delta p, 0) = D(p,t)$. Then $\theta(t) = \Delta p/t$, and $\theta'(t)$ can then be backed out from the demand response.

Proposition A.2. *Suppose for simplicity that utility is quasilinear. Consider $\Delta EB(\Delta t|t) := EB(t+\Delta t) - EB(t)$, and let ΔEB_0 be the value of ΔEB that would be inferred from $D(p,t)$ under the assumption that $F(\theta|p,t)$ is degenerate. Then there exist $\underline{d} \le \Delta EB_0 < \bar{d}$ such that $D(p,t)$ can be consistent with any value* *of $\Delta EB(\Delta t|t)$ in $[\underline{d}, \bar{d}]$. When $t = 0$, and when $\bar{\theta}$ is an upper bound on the possible realizations of θ,*

$$\underline{d} = \frac{1}{2}\left(\frac{D_t(p, \Delta t)}{D_p(p, \Delta t)}\right)^2 (\Delta t)^2 D_p \tag{A1}$$

$$\bar{d} = \frac{1}{2}\left(\frac{D_t(p, \Delta t)}{D_p(p, \Delta t)}\right) \bar{\theta}(\Delta t)^2 D_p \tag{A2}$$

Proposition A.2 shows that when there is heterogeneity in θ, knowledge of the demand curve $D(p,t)$ is not sufficient to calculate the welfare implications of taxation. To see the intuition for this result, consider the case in which $t = 0$. In this case, welfare is proportional to $E[\theta|p,t]^2 + Var[\theta|p,t]$. The mean $E[\theta|p,t]$ is identified by D_t/D_p. However, $Var[\theta|p,t]$ cannot be identified from the demand curve, as aggregate demands do not provide information on the dispersion of the bias, only the extent to which it mutes the response to taxation *on average*. The variance is smallest when consumers are homogeneous, which corresponds to $\underline{d}$ in (A1), and it is largest when all consumers either have $\theta = \bar{\theta}$ or $\theta = 0$, which corresponds to $\bar{d}$ in (A2).

A.2 Income Effects

To generate results for income effects in the presence of bias heterogeneity, we will temporarily focus on a continuous demand model, as modeling income effects in a discrete choice model is somewhat awkward and is not typical even in neoclassical results on efficiency costs of taxation. We assume that producer prices as fixed, as characterizing efficiency costs with both income effects and endogenous producer prices is intractable even in the standard model (CLK).

We suppose that consumers have a utility function $U(x,y) = v(x) + v(y)$, and that they react to a tax as if it were θt. We suppose that U is the same for all consumers for simplicity, but more general results can be obtained for arbitrary distributions of U (analogous to our binary choice model with arbitrary distributions of tastes). We study a budget-adjustment rule where consumers react to the price-inclusive price of x as if it were $p + \theta t$, but overall do not misperceive the size of their budget as a consequence of the tax because they purchase x frequently and in relatively small amounts, and observe their new budget after every purchase of x. This gives rise to choices of x and y characterized by the following conditions

$$\frac{v'(x_\theta^*)}{u'(y_\theta^*)} = p + \theta t \tag{A3}$$

$$y^* + (p+t)x_\theta^* = z \tag{A4}$$

where z is the budget.

Now some routine algebra shows that $\theta = \frac{\frac{d(x_\theta^*)^c}{dt}}{\frac{d(x_\theta^*)^c}{dp}}$ for small t, where the compensated demand responses are defined by $\frac{d(x_\theta^*)^c}{dp} := \frac{d}{dp}x_\theta^* + \frac{d}{dz}x_\theta^*$ and $\frac{d(x_\theta^*)^c}{dt} = \frac{d}{dt}x_\theta^* + \frac{d}{dz}x_\theta^*$. It now follows by Proposition 2 of CLK that the efficiency cost corresponding to a type θ consumer is given by

$$\frac{1}{2}t^2\theta^2\frac{\varepsilon_{D,p}^c}{p+t}, \tag{A5}$$

where $\varepsilon_{D,p}^c$ is the compensated price-elasticity of demand. Aggregating (A5) across consumers yields the following analog to Proposition 2:

Proposition A.3. *If terms of order t^3 are higher are negligible, and if consumers choose x and y according to*

(A3) and (A4), then $EB \approx \frac{1}{2}t^2\left[E[\theta|p,t]\frac{\varepsilon^c_{D,t}}{p+t}+Var[\theta|p,t]\frac{\varepsilon^c_{D,p}}{p+t}\right]$*, where* $\varepsilon^c_{D,t}$ *is the compensated tax elasticity and* $\varepsilon^c_{D,p}$ *is the compensated price elasticity.*

Note that this proposition is essentially identical to Proposition 2 in the text, since the compensated elasticity equals the uncompensated elasticity with quasilinear utility. More generally, all the results in the body of the paper could be re-derived analogously. An important insight of this result is that even with income effects, efficiency costs are still zero when all consumers neglect the tax completely.

Note, however, that this result depends on the budget adjustment rule. If, instead, consumers first purchased large quantities of x and only then spent the remainder of their income on y, then efficiency costs would be positive even with complete neglect of taxes. The reason is that consumers would overspend on x in a way that would generate inefficiently low levels of consumption of y.

Concretely, the choice of x would now be characterized by

$$v'(x^*_\theta) = (p+\theta t)u'(z-(p+\theta t)) \tag{A6}$$

rather than by $v'(x^*_\theta) = (p+\theta t)u'(z-(p+t))$. In this case, the choice of y would would be overly sensitive to change dt of the not-fully-salient commodity tax, in the sense that it would decrease by more than it would with respect to a salient lump sum tax of size dtx^*. This captures the intuition that inattention to taxes may cause consumers to misperceive their effective budget, and thus experience an unpleasant surprise later in time after seeing how much money they have left to purchase y.

Formally, routine algebra now shows that $\theta = \frac{\frac{dx^*_\theta}{dt}}{\frac{dx^*_\theta}{dp}}$ for small t. That is, θ is now the ratio of uncompensated demand responses rather than compensated demand responses. Now define $\theta^c := \frac{\theta\frac{dx^*_\theta}{dp}+x^*_\theta\frac{dx^*_\theta}{dz}}{\frac{d(x^*_\theta)^c}{dp}}$. In this case, Proposition 2 of CLK can now be used to show that when terms of order t^3 and higher are negligible,

$$EB \approx \frac{1}{2}t^2\left[E[\theta^c|p,t]\frac{\varepsilon^c_{D,t}}{p+t}+Var[\theta^c|p,t]\frac{\varepsilon^c_{D,p}}{p+t}\right]. \tag{A7}$$

When $\theta = 0$ for all consumers, $\theta^c < 0$ and $\varepsilon^c_{D,t} < 0$, and thus (A7) shows that excess burden can still be positive, formalizing the intuition that we sketched in the previous paragraph.

Note, however, that if x^*u''/u' is negligible, as is the case when expenditures arising from x are small relative to the total budget, then $\frac{d(x^*_\theta)^c}{dt} \approx \frac{dx^*_\theta}{dt}$ and $\frac{d(x^*_\theta)^c}{dp} \approx \frac{dx^*_\theta}{dp}$. Thus also $\theta^c \approx \theta$, and so efficiency costs can still be well approximated by the formula in Proposition 2 for small t. This captures the intuition that income effects are negligible for small-ticket purchases, and thus the derivations in our paper still hold for such decisions. Modifications to our main formulas are needed only when 1) the purchases are large and 2) the budget adjustment rule does not correspond to the one in (A3) and (A4).

A.3 Welfare with Redistributive Motives

We now consider a policymaker who aims not only to minimize efficiency costs, but also wishes to equalize wealth. We model this setting as simply as possible in this paper, but we refer the interested reader to Lockwood and Taubinsky (2015) and Farhi and Gabaix (2015) for richer models of tax salience with redistributive concerns. Lockwood and Taubinsky (2015), for example, consider a policymaker who has access to both a non-linear income tax and a (non)-salient commodity tax that he can apply to a sin good such as cigarette consumption. Analogous to the results in this section, Lockwood and Taubinsky (2015) also show that the welfare consequences of the less salient commodity tax depend on how attention to the tax covaries with income.[1]

We consider an economy in which consumers start with different levels of wealth $Z_1, \dots Z_N$, indexed by ω. We let F denote the joint distribution of (v, θ, ω), and we let $D^\omega(p,t)$ denote the demand curve of consumers with endowment Z_ω. We assume for simplicity that $D^\omega(p,0)$ and $D^\omega_p(p,0)$ do not depend on i. We let F denote the joint distribution of θ, v, ω and we let H denote the marginal distribution of i. We continue assuming that consumers choose x if $v \geq p + \theta t$.

The government maximizes $W = \int g_\omega(Z_\omega + (v-p-t)\mathbf{1}_x)dF + \lambda D$, where λ is the marginal value of public funds (used for production of a public good, for example), and g_ω is the weight on the utility of consumers with wealth Z_w. Redistributive preferences are captured by g_ω decreasing in Z_ω. Similar results can be obtained by endowing consumers with utility functions $U(Z_\omega + (v-p-t)\mathbf{1}_x)$ instead of assuming exogenous given weights g_ω.

Proposition A.4. *Set $\bar{g} := \int g_\omega$. For a small tax t,*

$$\begin{aligned} W(t) - W(0) \quad &\approx \quad \overbrace{\frac{t^2}{2}\left(\bar{g}\left(E[\theta|p,t]^2 + Var[\theta|p,t]\right)D_p(p,t) + Cov[g_\omega,(\theta-1)^2|p,t]D_p(p,t) - 2\bar{g}D_t(p,t)\right)}^{\textit{Welfare implications of misoptimization}} \\ &+ \quad \underbrace{t(\lambda - \bar{g})D(p,t) + \frac{1}{2}t^2\lambda D_t(p,t)}_{\textit{Impact on public funds net of mechanical income effect}} \end{aligned}$$

Proposition A.4 shows that just as excess burden is increasing in $E[\theta^2]$ and $Var[\theta]$, welfare is similarly decreasing in these two terms. Because the welfare formula reduces to the formula in Proposition 2 when $g_\omega = \lambda = 1$ for all ω, Proposition A.4 is a generalization of our baseline result to the case in which the equalities $g_\omega = \lambda = 1$ do not hold.

The new insight that the more general welfare framework generates is that welfare is also increasing in the covariance between g_ω and the size of the mistake in computing bias. Because $(\theta-1)^2$ attains its minimum at $\theta = 1$, welfare is decreasing in the extent to which the deviation from full rationality, either due to over- or underreaction to taxes, is concentrated on the low income earners.

In short, conditional on $E[\theta|p,t]$ and $Var[\theta|p,t]$, and knowledge of D_p, inferred welfare is lower when the mistake is concentrated on the poor. If consumers are over-spending on x because they are underreacting to the tax, the policymaker prefers that this over-spending is concentrated on consumers with low marginal social benefit from income.

B A More General Framework for Optimal Taxes

B.1 Welfare and Optimal Tax Formulas

We now suppose that while a consumer's perceived value from the good x is v, the actual social value from the consumer getting the good is $v - \gamma$. The wedge γ represents either externalities or internalities. For example, γ could correspond to consumers misperceiving the price of the good. We make several simplifying assumptions. First, we assume that we can partition consumers into θ types $j = 1, \dots J$ such that type j

[1]We remind the reader that while the Atkinson-Stiglitz theorem shows that commodity taxation should not be used with neoclassical consumers in the presence of nonlinear income taxation, this theorem does not hold when the income tax and the commodity tax are not equally salient, or when there are other biases that cause consumers to over- or under-consume the good in question (Lockwood and Taubinsky 2015).

consumers reacts to a tax t as if it was $\theta_j(t)t$. Let μ_j be the fraction of type j consumers. Second, we assume that terms of order $t^3 D_{pp}$ are negligible. Third, we assume that γ, v, θ are mutually independent. Fourth, we assume constant returns to scale production technology.

The policymaker's objective function is to maximize

$$W(t) = \int [y - (p+t)\mathbf{1}_x + (v-\gamma)\mathbf{1}_x] + \lambda t D$$

where λ is the marginal value of public funds. We now characterize optimal taxes in this more general model.

Proposition B.1. *Normalize $p = 1$, and define $\bar{\gamma} := E[\gamma|p,t]$ to be the bias of consumers marginal to a tax change, $a(t) := E[\theta|p,t]$ and $b(t) := E[\theta^2|p,t] = E[\theta|p,t]^2 + Var[\theta|p,t]$. Then*

1. $W'(t) = (\lambda-1)D + [(\lambda-1)t - \bar{\gamma}]D_t + \frac{b(t)+b'(t)t/2}{a(t)+a'(t)t} t D_t$
2. *The optimal tax t is implicitly defined by*

$$t = -\frac{(\lambda-1)(a(t)+a'(t)t)D - \bar{\gamma}(a(t)+a'(t))D_t}{(\lambda-1)(a(t)+a'(t)t)D_t + (b(t)+b'(t)t/2)D_t}$$
$$\frac{(\lambda-1)(a(t)+a'(t)t) + \bar{\gamma}(a(t)+a'(t)t)\varepsilon_{D,t}}{(\lambda-1)(a(t)+a'(t)t)\varepsilon_{D,t} + (b(t)+b'(t)t/2)\varepsilon_{D,t} - (\lambda-1)(a(t)+a'(t))} \tag{A8}$$

Proof. Set $\varphi_j = \gamma_j + (1-\theta_j)t$. This is the amount by which a consumer overconsumes a good. There are now three effects from increasing the tax by dt:

1. A mechanical revenue effect, net of the impact on individual's utility, given by $(\lambda-1)Ddt$
2. A revenue loss from the substitution effect, given by $\lambda t D_t dt$
3. A correction effect, given by $-\sum_j \phi_j D_t^j \mu_j dt$, where D_t^j is the demand response of the type j consumer to the tax.

The "correction effect" is given by

$$\begin{aligned}
-\sum_j \phi_j D_t^j \mu_j &= -\sum_j (\gamma_j + (1-\theta_j)t) D_t^j \mu_j \\
&= -\bar{\gamma} D_t - \sum_j (1-\theta_j) t (\theta_j(t) + \theta_j'(t)t) D_p^j \mu_j \\
&= -\bar{\gamma} D_t - t a(t) D_p - t^2 a'(t) D_p + t b(t) D_p + \frac{t^2 b'(t)}{2} D_p \\
&= -\bar{\gamma} D_t - t D_t + t b(t) D_p + \frac{t^2 b'(t)}{2} D_p
\end{aligned}$$

Putting all of the effects together and using $(a(t) + ta'(t))D_p = D_t$ implies that

$$\begin{aligned}
W'(t) &= (\lambda-1)D + [(\lambda-1)t - \bar{\gamma}]D_t + (b(t) + b'(t)t/2) t D_p \\
&= (\lambda-1)D + [(\lambda-1)t - \bar{\gamma}]D_t + \frac{b(t)+b'(t)t/2}{a(t)+ta'(t)} t D_t
\end{aligned}$$

□

The general formula in part 1 of Proposition B.1, which is an analogue of the kinds of general results derived in Farhi and Gabaix (2015) for continuous demand, is a more general manifestation of the forces discussed in our excess burden analysis in Section 2. Keeping in mind that $a(t) := E[\theta|t]$ and $b(t) := E[\theta^2|t] = E[\theta|t]^2 + Var[\theta|t]$, the formula shows that there are four key statistics: the mean, the variance, and how both of those change with respect to the tax. The frictions $\bar{\gamma}$ enter into the formula additively. The higher is $\bar{\gamma}$, the higher is the optimal tax t, and thus the larger the impact that the variance component of $b(t)$ has on welfare.

Part 2 of the proposition partially solves for the optimal tax to present formulas generalizing the usual "inverse elasticity" result from Ramsey taxation. To obtain intuition for the main result, we first focus on a simple case in which $\bar{\gamma} = 0$ and optimal taxes are not large. In this case, the optimal tax formula trades off the deadweight loss computed in Proposition 5 with the revenue gain (net of the mechanical effect on consumers' incomes).

Corollary B.1. *When λ is close to 1 and $\bar{\gamma} = 0$,*

$$t \approx \frac{(\lambda - 1)E[\theta|p,t]}{(E[\theta|p,t]^2 + Var[\theta|p,t])\,\varepsilon_{D,t}}$$

Proof. When $\bar{\gamma} = 0$ and λ is close to 1, the tax t becomes small and thus terms to t in equation (**??**) become negligible. Also the term in the denominator proportional to $(\lambda - 1)$ becomes negligible. This yields the desired result. □

Just as Proposition 2 shows that the deadweight loss is increasing in both the mean and the variance of θ, Corollary B.1 shows that the size of the optimal tax is decreasing in both the mean and the variance of θ.

In the presence of other (small) frictions, the tax must be adjusted to offset the other internalities and/or externalities captured by $\bar{\gamma}$. The extent to which the tax is adjusted depends on both average θ and on the variance. The lower is the average θ, the more the tax needs to be adjusted, as reflected by the $E[\theta|p,t]$ term in the numerator and the $E[\theta|p,t]^2$ in the denominator. On the other hand, the higher is the variance in θ, the greater the misallocation form increasing the tax, and thus the lower is the optimal tax.

Corollary B.2. *When λ is close to 1,*

$$t \approx \frac{(\lambda - 1)E[\theta|p,t]}{(E[\theta|p,t]^2 + Var[\theta|p,t])\,\varepsilon_{D,t}} + \bar{\gamma}\frac{E[\theta|p,t]}{(E[\theta|t]^2 + Var[\theta|p,t])}$$

As a last special case for obtaining intuition, we focus on the case in which $Var[\theta|p,t] = 0$ for all t.

Corollary B.3. *Suppose that $Var[\theta|p,t] = 0$ for all t . Then*

$$t = \frac{\lambda - 1 + \bar{\gamma}\varepsilon_{D,t}}{(\lambda - 1)(\varepsilon_{D,t} - 1) + E[\theta|p,t]\varepsilon_{D,t}} \tag{A9}$$

and when $\bar{\gamma} = 0$,

$$\frac{t}{1+t} = \frac{\lambda - 1}{(\lambda - (1 - E[\theta|t]))\varepsilon_{D,t}} \tag{A10}$$

In this last special case, equation (A10) provides a simple analog to the standard inverse elasticity rule of Ramsey taxation, showing that the rule is simply modified by the bias term $(1 - E[\theta|t])$.

B.2 Implications for Ramsey Taxation

The formulas derived so far are immediately transferable to the canonical Ramsey taxation models. In particular, let y be untaxed leisure, and let $x_1, \ldots x_K$ be the possible products consumers can purchase, and that, for simplicity, utility is separable in the consumption of these goods. Suppose that the government sets taxes $t_1, \ldots t_K$ on the k goods to meet a revenue target R. In this case, the value of public funds λ is determined endogenously. Set $\tau_i = t_i/p_i$ to be the tax rate.

In the standard Ramsey model, the taxes are determined by the inverse elasticity rule

$$\frac{\tau_i/(1+\tau_i)}{\tau_j/(1+\tau_j)} = \frac{\varepsilon_{D^j,t_j}}{\varepsilon_{D^i,t_i}}$$

What are the implications of tax salience? For simple intuition, suppose first that $Var[\theta|t] = 0$ and that $\bar{\gamma} = 0$. Suppose, moreover, that θ depends only on the size of the tax, so that with uniform taxes t_k, it would be identical for across the K goods. In this case, equation (A10) implies that

$$\frac{\tau_i/(1+\tau_i)}{\tau_j/(1+\tau_j)} = \frac{\varepsilon_{D^j,t_j}}{\varepsilon_{D^i,t_i}} \cdot \frac{\lambda - (1 - E[\theta|p_j,\tau_j])}{\lambda - (1 - E[\theta|p_j,\tau_j])}.$$

A key implication here is that if $E[\theta|p,\tau]$ does not depend on p or τ, then the standard inverse elasticity rule continues to hold, and thus *with a fixed revenue requirement* R, taxes are identical to what they are in the standard model. Matters are different, however, if θ is endogenous to the tax. In particular, if $E[\theta|p,\tau]$ is increasing in p and/or τ, then the inverse elasticity rule becomes dampened toward uniform taxation, as consumers will be more attentive to higher taxes, and thus higher taxes generate relatively higher efficiency costs. Additionally, if $E[\theta|p,\tau]$ is increasing in p (because taxes are higher on more expensive items keeping the tax rate constant), then tax rates should be lower on more expensive products. More generally, the inverse elasticity rule is modified by how both the mean and the variance change with respect to the tax.

C Proofs of Propositions

Proof of Proposition 1 Let p_0 be the initial price and let $p(t)$ be the final price set by producers. Let x_1^* be the equilibrium quantity after the tax change, and let x_0^* be the equilibrium quantity before the tax change. The formula for excess burden is given by

$$\begin{aligned} EB(t,F) &= \overbrace{\left[\int_{v \geq p_0} (v - p_0)dF - \int_{v \geq p(t)+\theta t} (v - p(t) - t)dF\right]}^{\text{Equivalent variation in wealth for consumers}} - \overbrace{\int_{v \geq p(t)+\theta t} tdF}^{\text{Change in government revenue}} \\ &+ \overbrace{(p_0 x_0^* - C(x_0^*)) - (p(t)x_1^* - C(x_1^*))}^{\text{Change in producer profits}} \\ &= \int_{v \geq p_0} (v - p_0)dF - \int_{v \geq p(t)+\theta t} (v - p(t))dF + (p_0 x_0^* - C(x_0^*)) - (p(t)x_1^* - C(x_1^*)) \quad \text{(A11)} \end{aligned}$$

Now by the multidimensional Leibniz rule,

$$
\begin{aligned}
\frac{d}{dt}\int_{v\geq p(t)+\theta t}(v-p(t))dF &= -\int\left[(\theta t)\frac{d}{dt}(\theta t+p(t))\right]dF(\theta,v|v=p(t)+\theta t)\\
&+ \int_{v\geq p(t)+\theta t}\left(-\frac{dp(t)}{dt}\right)dF(\theta,v)\\
&= -\int(\theta t)\left(\theta+\frac{dp(t)}{dt}\right)dF(\theta,v|v=p(t)+\theta t)\\
&+ \int_{v\geq p(t)+\theta t}\left(-\frac{dp(t)}{dt}\right)dF\\
&=- \quad tE_F\left[\theta^2+\theta\frac{dp(t)}{dt}|v=p(t)+\theta t\right]\int dF(\theta,v|v=p(t)+\theta t)-x_1^*\frac{dp(t)}{dt}\\
&= -tE_F[\theta]E\left[\theta+\frac{dp(t)}{dt}|v=p(t)+\theta t\right]\int dF(\theta,v|v=p(t)+\theta t)\\
&- tVar_F[\theta|v=p(t)+\theta t]\int dF(\theta,v|v=p(t)+\theta t)-x_1^*\frac{dp(t)}{dt}\\
&= tE_F[\theta|p(t),t]\frac{d}{dt}x_1^*+tVar_F[\theta|p(t),t]D_p(p(t),t)-x_1^*\frac{dp(t)}{dt} \qquad \text{(A12)}
\end{aligned}
$$

To arrive to the final equation in (A12) from the preceding equation, we use the fact that

$$
\begin{aligned}
\frac{d}{dt}x_1^* &= \frac{d}{dt}\int_{v\geq p(t)+\theta t}dF\\
&= -\int\frac{d}{dt}(p(t)+\theta t)dF(\theta,v|v=p(t)+\theta t)\\
&= -\int\left(\theta+\frac{d}{dt}p(t)\right)dF(\theta,v|v=p(t)+\theta t) \qquad \text{(A13)}
\end{aligned}
$$

Next, the Envelope Theorem implies that

$$
\frac{d}{dt}\left(p(t)x_1^*-C(x_1^*)\right)=x_1\frac{dp(t)}{dt}
$$

Putting this together, we thus have that

$$
\frac{d}{dt}EB(t,F)=-tE_F[\theta|v=p(t)+\theta t]\frac{d}{dt}x_1^*-tVar_F[\theta|v=p(t)+\theta t]D_p(p(t),t) \qquad \text{(A14)}
$$

Proof of Proposition 2 Assuming that $E[\theta|p_1,t]$, $Var[\theta|p_1,t]$, D, and x_1^* are smooth, it follows from (A14) that when F_t does not depend on t

$$
\frac{d^2}{dt^2}EB(t,F)=-E_F[\theta|v=p_1+\theta t]\frac{d}{dt}x_1^*-Var_F[\theta|v=p_1+\theta t]D_p(p_1,t)+O(t) \qquad \text{(A15)}
$$

where $O(t)$ represents all terms of order t or higher (as $t\to 0$). A Taylor expansion thus implies that when F_t does not depend on t,

$$EB(t,F) = EB(0,F) + t\frac{d}{dt}EB(t,F)|_{t=0} + \frac{(t)^2}{2}\frac{d^2}{dt^2}EB(t,F)|_{t=0} + O((t)^3)$$
$$= -\frac{1}{2}t^2\left[E_F[\theta|v = p_1 + \theta t]\frac{d}{dt}x_1^* + Var_F[\theta|v = p_1 + \theta t]D_p(p_1,t)\right] + O(t^3) \qquad \text{(A16)}$$

where $O(t^3)$ represents all terms of order t^3 or higher (as $t \to 0$). Finally, because $EB(t,F_t) = EB(t,F_t) - EB(0,F_0) = EB(t,F_t) - EB(0,F_t)$, the assumption that F_t does not depend on F is without loss of generality when computing $EB(t,F_t)$.

Proof of Proposition 3. We use the second order Taylor expansion

$$EB(t+\Delta t,F) - EB(t,F) = \Delta t\frac{d}{dt}EB(t,F) + \frac{(\Delta t)^2}{2}\frac{d^2}{dt^2}EB(t,F) + O((\Delta t)^3).$$

We have already computed $\frac{d}{dt}EB(t,F)$ in (A14). To get $\frac{d^2}{dt^2}EB(t,F)$ we differentiate the expression in (A14):

$$\begin{aligned}\frac{d^2}{dt^2}EB(t,F) &= -E_F[\theta|v = p(t) + \theta t]\frac{d}{dt}x^* - Var_F[\theta|v = p_1 + \theta t]D_p(p_1,t)\\ &- t\frac{d}{dt}E_F[\theta|v = p(t) + \theta t]_{t=t_1}\frac{d}{dt}x^* - t\frac{d}{dt}Var_F[\theta|v = p_1 + \theta t]|_{t=t_1}D_p\\ &- tE_F[\theta|v = p_1 + \theta t]\frac{d^2}{dt^2}x^* - tVar_F[\theta|v = p_1 + \theta t]\frac{d}{dt}D_p\end{aligned}$$

The crux of the proof is to show that the terms $t(\Delta t)^2D_{pt}$, $t_1(\Delta t)^2\frac{d}{dt}E[\theta|v = p+\theta t_1]$ and $t_1(\Delta t)^2\frac{d}{dt}Var_F[\theta|v = p_1 + \theta t_1]$ are negligible. This will establish that $t(\Delta t)^2E_F[\theta|v = p_1 + \theta t]\frac{d^2}{dt^2}x^*$ is negligible, and thus that

$$\begin{aligned}\frac{d^2}{dt^2}EB(t,F) &= -E_F[\theta|v = p(t) + \theta t]\frac{d}{dt}x^* - Var_F[\theta|v = p_1 + \theta t]D_p(p_1,t)\\ &+ \text{terms that are negligible when mulitplied by } (\Delta t)^2\end{aligned}$$

Once that is established, the second-order Taylor expansion will yield the statement of Proposition 3.

We first show that $t(\Delta t)^2\frac{d^2}{dt^2}x^*$ is negligible when $t(\Delta t)^2D_{pt}$, $t_1(\Delta t)^2\frac{d}{dt}E[\theta|v = p+\theta t_1]$ and $t_1(\Delta t)^2\frac{d}{dt}Var_F[\theta|v = p_1 + \theta t_1]$ are negligible. First, note that $\frac{d}{dt}D_p = D_{pp}p'(t) + D_{pt}$, and thus the term $t(\Delta t)^2\frac{d}{dt}D_p = t(\Delta t)^2D_{pp}p'(t) + t(\Delta t)^2D_{pt}$ is negligible. Next, note that equation (A13) implies that $\frac{d}{dt}x^* = E[\theta|p,t]D_p + p'(t)D_p$. To derive $p'(t)$, note that it satisfies $D(p(t),t) = S(p(t))$, and thus $D_pp'(t) + D_t = S_pp'(t)$, which implies that $p'(t) = \frac{D_t}{S_p - D_p} = \frac{E[\theta|p,t]D_p}{S_p - D_p}$. Thus $\frac{d}{dt}x^* = E[\theta|p,t]\left(\frac{D_p}{S_p-D_p}D_p + D_p\right)$. As before, the condition that $t(\Delta t)^2\frac{d}{dt}D_p$ is negligible will imply that that the term $t(\Delta t)^2\frac{d}{dt}\left(\frac{D_p}{S_p-D_p}D_p + D_p\right)$ is negligible. Thus by the chain rule:

$$\begin{aligned}t(\Delta t)^2E_F[\theta|v = p_1 + \theta t]\frac{d^2}{dt^2}x^* &\approx t(\Delta t)^2\left(\frac{d}{dt}E[\theta|p,t]\right)\left(\frac{D_p}{S_p - D_p}D_p + D_p\right)\\ &+ t(\Delta t)^2\left(E[\theta|p,t]\right)\frac{d}{dt}\left(\frac{D_p}{S_p - D_p}D_p + D_p\right)\end{aligned}$$

Thus, since $t(\Delta t)^2 \frac{d}{dt} E[\theta|v = p + \theta t_1]$ and $t(\Delta t)^2 \frac{d}{dt}\left(\frac{D_p}{S_p - D_p} D_p + D_p\right)$ are negligible, so is $t(\Delta t)^2 E_F[\theta|v = p_1 + \theta t] \frac{d^2}{dt^2} x^*$.

We now establish that $t(\Delta t)^2 D_{pt}$ is negligible and also that $t_1(\Delta t)^2 \frac{d}{dt} E[\theta|v = p + \theta t_1]$ and $t_1(\Delta t)^2 \frac{d}{dt} Var_F[\theta|v = p_1 + \theta t_1]$ are negligible. Let $H(\theta)$ be the marginal distribution of θ and let $G_\theta(v)$ be the marginal distribution of v conditional on θ, with differentiable density function g_θ. First, we show that $t(\Delta t)^2 D_{pt}$ is negligible. Letting M be the upper bound on θ, we have

$$\begin{aligned}
\left|t(\Delta t)^2 \frac{d}{dt} D_p\right| &= \left|t(\Delta t)^2 \frac{d}{dt} \int g_\theta(p(t) + \theta t) dH(\theta)\right| \\
&= \left|\int (t(\Delta t)^2)(\theta + p'(t)) g'_\theta(p(t) + \theta t)) dH(\theta)\right| \\
&\leq \left|\int (t(\Delta t)^2)(M + |p'(t)|) g'_\theta(p(t) + \theta t)) dH(\theta)\right| \\
&= \left|(t(\Delta t)^2)(M + p'(t)) D_{pp}\right| \\
&\approx 0
\end{aligned}$$

Similarly, we can show that $t_1(\Delta t)^2 \frac{d}{dt} \int \theta g_\theta(p(t) + \theta t) dH(\theta) \approx 0$. Thus, since

$$E[\theta|p, t] = \frac{\int_{v = p + \theta t} \theta g_\theta(p(t) + \theta t) dH(\theta)}{\int_{v = p + \theta t} g_\theta(p(t) + \theta t) dH(\theta)},$$

it follows that $t_1(\Delta t)^2 \frac{d}{dt} E[\theta|p(t), t]|_{t=t_1} \approx 0$. Analogous reasoning establishes that $t_1(\Delta t)^2 \frac{d}{dt} E[\theta^2|p(t), t]|_{t=t_1} \approx 0$. From this it then also follows that

$$t(\Delta t)^2 \frac{d}{dt} Var[\theta|p(t), t]_{t=t_1} = t(\Delta t)^2 \frac{d}{dt}\left(E[\theta^2|p(t), t] + E[\theta|p(t), t]^2\right)_{t=t_1} \approx 0$$

Putting this all together, we thus have

$$\begin{aligned}
\frac{d^2}{dt^2} EB(t, F) &= -E_F[\theta|v = p(t) + \theta t] \frac{d}{dt} x^* - Var_F[\theta|v = p_1 + \theta t] D_p(p_1, t) \\
&+ \text{ terms that are negligible when mulitplied by } (\Delta t)^2
\end{aligned}$$

The result in the proposition now follows from the second order Taylor expansion $EB(t + \Delta t, F) - EB(t, F) = \Delta t \frac{d}{dt} EB(t, F) + \frac{(\Delta t)^2}{2} \frac{d^2}{dt^2} EB(t, F) + O((\Delta t)^3)$.

Proof of Proposition 4 We write the equilibrium price as a function of n here, suppressing the dependency on t.

Let $H_n(\theta)$ denote the marginal distribution of θ given nudge parameter n. Let $G_n(v|\theta)$ denote the conditional distribution of v, with a differentiable density function denoted by g_n. Under assumptions A, B, and C:

$$
\begin{aligned}
\frac{d}{dn}\int_{v\geq p+\theta t}(v-p)dF_n &= \frac{d}{dn}\int_{v\geq p+\theta t}(v-p)dG_n(v|\theta)dH_n(\theta)\\
&= \frac{d}{dn}\int_{v\geq p+h(\theta,n)t}(v-p)dG_0(v|\theta)dH_0(\theta)\\
&= \int_\theta\left[h(\theta,n)t\frac{d}{dn}(h(\theta,n)t+p)\right]dG_0(p+h(\theta,n)t|\theta)dH_0(\theta)\\
&= t^2\int_\theta h(\theta,n)\frac{d}{dn}h(\theta,n)dG_0(p+h(\theta,n)t|\theta)dH_0(\theta)
\end{aligned}
$$

Next,

$$
\begin{aligned}
\frac{d}{dn}E_{F_n}[\theta^2|p,t] = &\frac{d}{dn}\frac{\int_\theta h(\theta,n)^2 g_0(p+h(\theta,n)t|\theta)dH_0(\theta)}{\int_\theta g_0(p+h(\theta,n)t|\theta)dH_0(\theta)}\\
= &\frac{2\int\left(h(\theta,n)\frac{d}{dn}h(\theta,n)\right)g_0(p+h(\theta,n)t|\theta)dH_0(\theta)}{D_p}\\
+ &\frac{t\int_\theta h(\theta,n)^2\frac{d}{dn}h(\theta,n)g_0'(p+h(\theta,n)t)dH_0(\theta)}{D_p}\\
- &\frac{\int_\theta h(\theta,n)^2 g_0(p+h(\theta,n)t|\theta)dH_0(\theta)\cdot t\int_\theta h(\theta,n)^2\frac{d}{dn}h(\theta,n)g_0'(p+h(\theta,n)t|\theta)dH_0(\theta)}{D_p^2}
\end{aligned}
$$

Let M_1 be a bound on $|h|$ and M_2 be a bound on $|\frac{\partial}{\partial n}h(\theta,n)|$. By the triangle inequality,

$$
\begin{aligned}
\left|\frac{1}{2}t^2\frac{d}{dn}E_{F_n}[\theta^2|p,t]D_p - \frac{d}{dn}\int_{v\geq p+\theta t}(v-p)dF_n\right| =&\frac{1}{2}t^2\left|t\int_\theta h(\theta,n)^2\frac{d}{dn}h(\theta,n)g_0'(p+h(\theta,n)t)dH_0(\theta)\right|\\
&+\frac{1}{2}t^2\left|\frac{\int_\theta h(\theta,n)^2 g_0(p+h(\theta,n)t|\theta)dH_0(\theta)\cdot t\int_\theta h(\theta,n)^2\frac{d}{dn}h(\theta,n)g_0'(p+h(\theta
}{D_p^2}\right.\\
\leq&\frac{1}{2}t^3M_1^2M_2\left|\int_\theta g_0'(p+h(\theta,n)t)dH_0(\theta)\right|\\
&+\frac{1}{2}t^3\frac{(M_1^2|D_p|)\cdot\left(M_1^2M_2\left|\int_\theta g_0'(p+h(\theta,n)t)dH_0(\theta)\right|\right)}{D_p^2}\\
\leq&\frac{1}{2}t^3|D_{pp}|\left(M_1^2M_2+\frac{M_1^4M_2}{|D_p|}\right)
\end{aligned}
$$

But since terms proportional to t^3D_{pp} are negligible (assumption C). We thus have that

$$
\frac{d}{dn}\int_{v\geq p+\theta t}(v-p)dF_n \approx \frac{1}{2}t^2\frac{d}{dn}E_{F_n}[\theta^2|p,t]D_p
$$

Since producer profits are fixed, $\frac{d}{dn}EB(t,F_n) = -\frac{d}{dn}\int_{v\geq p+\theta t}(v-p)dF_n$, which yields the first statement in the proposition.

To prove the second part of the proposition, we now compute the second derivative of excess burden. We first establish that $t(\Delta n)^2\frac{d}{dn}D_p$ is negligible.

Under Assumption A,

$$\begin{aligned}
\left|\frac{d}{dn}D_p\right| &= \left|\frac{d}{dn}\int g(p+h(\theta,n)t)dH_0(\theta)\right| \\
&= \left|\int\left(\frac{\partial}{\partial n}h\right)g'(p+h(\theta,n)t)dH_0(\theta)\right| \\
&\leq M_2\left|\int g'(p+h(\theta,n)t)dH_0(\theta)\right| \\
&\leq M_2\left|D_{pp}\right|
\end{aligned}$$

Thus $t(\Delta n)^2\frac{d}{dn}D_p$ is negligible if $t(\Delta n)^2 D_{pp}$ is negligible, and

$$(\Delta n)^2\frac{d^2}{dn^2}EB \approx -\frac{t^2}{2}\frac{d^2}{dn^2}E_{F_n}[\theta^2|p,t]D_p.$$

Now a second-order expansion of EB around n yields

$$\begin{aligned}
EB(t,F_{n+\Delta n})-EB(t,F_n) &= (\Delta n)\frac{d}{dn}EB(t,F_n)+\frac{(\Delta n)^2}{2}\frac{d^2}{dn^2}EB(t,F_n)+O((\Delta n)^3) \\
&= -\frac{t^2}{2}\left((\Delta n)\frac{d}{dn}E_{F_n}[\theta^2|p,t]+\frac{(\Delta n)^2}{2}\frac{d^2}{dn^2}E_{F_n}[\theta^2|p,t]\right)D_p+O((\Delta n)^3).
\end{aligned}$$

To complete the proof, note that a second-order Taylor expansion shows that:

$$E_{F_{n+\Delta n}}[\theta^2|p,t]-E_{F_n}[\theta^2|p,t] = (\Delta n)\frac{d}{dn}E_{F_n}[\theta^2|p,t]+\frac{(\Delta n)^2}{2}\frac{d^2}{dn^2}E_{F_n}[\theta^2|p,t]+O((\Delta n)^3)$$

Proof of Proposition 5 Combining Propositions 3 and 4, we have

$$\frac{d}{dt}EB(t,F_t) \approx -E[\theta^2|p,t]tD_p-\frac{t^2}{2}\frac{d}{dt}E[\theta^2|p,t]D_p.$$

Taking the second derivative, and ignoring the terms proportional to $\frac{d}{dt}D_p$ since those are negligible (when multiplied by $t(\Delta t)^2$ by the reasoning in Propositions 3 and 4, we have

$$\begin{aligned}
\frac{d^2}{dt^2}EB(t,F_t) &\approx -E[\theta^2|p,t]D_p-t\frac{d}{dt}E[\theta^2|p,t]D_p \\
&- t\frac{d}{dt}E[\theta^2|p,t]D_p-\frac{t^2}{2}\frac{d^2}{dt^2}E[\theta^2|p,t]D_p
\end{aligned}$$

Next, note that

$$\begin{aligned}
(\Delta t)\frac{d}{dt}E[\theta^2|p,t]|_{t=t_1}+\frac{(\Delta t)^2}{2}\frac{d^2}{dt^2}E[\theta^2|p,t]|_{t=t_1} &= E[\theta^2|p,t_2]-E[\theta^2|p,t_1]+O((\Delta t)^3) \\
(\Delta t)E[\theta^2|p,t_1]+(\Delta t)^2\frac{d}{dt}E[\theta^2|p,t]_{t=t_1} &= (\Delta t)E[\theta^2|p,t_2]+O((\Delta t)^3) \\
(\Delta t)^2E[\theta^2|p,t_1] &= (\Delta t)^2E[\theta^2|p,t_2]+O((\Delta t)^3)
\end{aligned}$$

Using these identities, we have:

$$
\begin{aligned}
EB(t_1+\Delta t, F_{t_2}) - EB(t_1, F_{t_1}) &= \Delta t \frac{d}{dt} EB(t,F_t)|_{t=t_1} + \frac{(\Delta t)^2}{2}\frac{d^2}{dt^2} EB(t,F_t)|_{t=t_1} + O((\Delta t)^3) \\
&= -(\Delta t)\left(t_1 E[\theta^2|p,t] + \frac{t_1^2}{2}\frac{d}{dt}E[\theta^2|p,t]|_{t=t_1}\right) D_p \\
&+ -\frac{(\Delta t)^2}{2}\left(E[\theta^2|p,t] + t_1 \frac{d}{dt}E[\theta^2|p,t]|_{t=t_1}\right) D_p \\
&+ -\frac{(\Delta t)^2}{2}\left(t_1 \frac{d}{dt}E[\theta^2|p,t]|_{t=t_1} + \frac{t_1^2}{2}\frac{d^2}{dt^2}E[\theta^2|p,t]|_{t=t_1}\right) D_p + O((\Delta t)^3) \\
&= -t_1(\Delta t) E[\theta^2|p,t_1] D_p - \frac{(\Delta t)^2}{2} E[\theta^2|p,t_1] D_p \\
&+ -\frac{t_1^2}{2}\left((\Delta t)\frac{d}{dt}E[\theta^2|p,t]|_{t=t_1} + \frac{(\Delta t)^2}{2}\frac{d^2}{dt^2}E[\theta^2|p,t]|_{t=t_1}\right) D_p \\
&+ -\frac{t_1(\Delta t)^2}{2}\left(\frac{d}{dt}E[\theta^2|p,t]|_{t=t_1} + \frac{d}{dt}E[\theta^2|p,t]|_{t=t_1}\right) D_p + O((\Delta t)^3) \\
&= -t_1\left((\Delta t) E[\theta^2|p,t_1] + (\Delta t)^2 \frac{d}{dt}E[\theta^2|p,t]|_{t=t_1})\right) D_p \\
&+ -\frac{(\Delta t)^2}{2} E[\theta^2|p,t_1] D_p \\
&+ -\frac{t_1^2}{2}\left(E[\theta^2|p,t_2] - E[\theta^2|p,t_1]\right) D_p + O((\Delta t)^3) \\
&= -\left(t_1(\Delta t) + \frac{(\Delta t)^2}{2}\right) E[\theta^2|p,t_2] D_p \\
&+ -\frac{t_1^2}{2}\left(E[\theta^2|p,t_2] - E[\theta^2|p,t_1]\right) D_p + O((\Delta t)^3)
\end{aligned}
$$

Proof of Corollary 2 Using $\frac{d}{dt}D = \left(E[\theta|p,t] + t\frac{d}{dt}E[\theta|p,t]\right) D_p$, we have

$$
\begin{aligned}
\frac{d}{dt}EB(t,F) &\approx -E[\theta^2|p,t] t D_p - \frac{t^2}{2}\frac{d}{dt}E[\theta^2|p,t] D_p \\
&= -E[\theta|p,t]^2 t D_p - Var[\theta|p,t] t D_p - \frac{t^2}{2}\frac{d}{dt}\left(E[\theta|p,t]^2 + Var[\theta|p,t]\right) D_p \\
&= -E[\theta|p,t]^2 t D_p - Var[\theta|p,t] t D_p - t^2 E[\theta|p,t]\frac{d}{dt}E[\theta|p,t] D_p - \frac{t^2}{2}\frac{d}{dt}Var[\theta|p,t] D_p \\
&= -E[\theta|p,t] t \left(E[\theta|p,t] + t\frac{d}{dt}E[\theta|p,t]\right) D_p - t\left(Var[\theta|p,t] + \frac{t}{2}\frac{d}{dt}Var[\theta|p,t]\right) D_p \\
&= -E[\theta|p,t] t D_t - t\left(Var[\theta|p,t] + \frac{t}{2}\frac{d}{dt}Var[\theta|p,t]\right) D_p
\end{aligned}
$$

Next,

$$\begin{aligned}\frac{d^2}{dt^2}EB(t,F) &\approx -E[\theta|p,t]D_t - \left(Var[\theta|p,t] + t\frac{d}{dt}Var[\theta|p,t]\right)D_p \\ &\quad -t\frac{d}{dt}E[\theta|p,t]D_t - E[\theta|p,t]t\frac{d}{dt}\left(E[\theta|p,t] + t\frac{d}{dt}E[\theta|p,t]\right)D_p \\ &\quad -t\left(\frac{d}{dt}Var[\theta|p,t] + \frac{t}{2}\frac{d^2}{dt^2}Var[\theta|p,t]\right)D_p + O(tD_{pp}) \\ &= -E[\theta|p,t]D_t - \left(Var[\theta|p,t] + t\frac{d}{dt}Var[\theta|p,t]\right)D_p \\ &\quad -t\frac{d}{dt}E[\theta|p,t]D_t - \frac{t}{2}\frac{d}{dt}E[\theta|p,t]^2 D_p - t\frac{d}{dt}Var[\theta|p,t]D_p \\ &\quad -tE[\theta|p,t]\frac{d^2}{dt^2}\left(E[\theta|p,t] + \frac{t}{2}Var[\theta|p,t]\right)D_p + O(tD_{pp})\end{aligned}$$

where we use $2E[\theta|p,t]\frac{d}{dt}E[\theta|p,t] = \frac{d}{dt}\left(E[\theta|p,t]^2\right)$.

Now since the approximations $E[\theta|p,t_2] - E[\theta|p,t_1] \approx \Delta t\frac{d}{dt}E[\theta|p,t]|_{t=t_1}$ and $Var[\theta|p,t_2] - Var[\theta|p,t_1] \approx \Delta t\frac{d}{dt}Var[\theta|p,t]|_{t=t_1}$ are valid, the terms $(\Delta t)^2\frac{d^2}{dt^2}E[\theta|p,t]$ and $(\Delta t)^2\frac{d^2}{dt^2}Var[\theta|p,t]$ are negligible. So

$$\begin{aligned}EB(t+\Delta t, F_{t_2}) - EB(t, F_{t_1}) &= \Delta t\frac{d}{dt}EB(t,F)|_{t=t_1} + \frac{(\Delta t)^2}{2}\frac{d^2}{dt^2}EB(t,F)|_{t=t_1} + O((\Delta t)^3) \\ &\approx -t_1(\Delta t)E[\theta|p,t_1]D_t - t_1(\Delta t)Var[\theta|p,t_1]D_p - \frac{t_1^2(\Delta t)}{2}\frac{d}{dt}Var[\theta|p,t]_{t=t_1}D_p \\ &\quad -\frac{(\Delta t)^2}{2}E[\theta|p,t_1]D_t - \frac{(\Delta t)^2}{2}Var[\theta|p,t_1]D_p - \frac{(\Delta t)^2}{2}t_1\frac{d}{dt}Var[\theta|p,t_1]|_{t=t_1}D_p \\ &\quad -\frac{(\Delta t)^2}{2}t_1\frac{d}{dt}E[\theta|p,t_1]D_t - t_1\frac{(\Delta t)^2}{2}\frac{d}{dt}Var[\theta|p,t]|_{t=t_1}D_p \\ &\quad -\frac{t_1(\Delta t)^2}{4}\frac{d}{dt}E[\theta|p,t]^2|_{t=t_1}D_p + O((\Delta t)^3)\end{aligned}$$

Next, note that

$$(\Delta t)E[\theta|p,t_1] + \frac{1}{2}(\Delta t)^2\frac{d}{dt}E[\theta|p,t]_{t=t_1} = (\Delta t)\frac{E[\theta|p,t_1] + E[\theta|p,t_2]}{2} + O((\Delta t)^3) \quad \text{(A17)}$$

$$(\Delta t)Var[\theta|p,t_1] + \frac{1}{2}(\Delta t)^2\frac{d}{dt}Var[\theta|p,t]_{t=t_1} = (\Delta t)\frac{Var[\theta|p,t_1] + Var[\theta|p,t_2]}{2} + O((\Delta t)^3) \quad \text{(A18)}$$

$$(\Delta t)D_p\left(E[\theta^2|p,t_2] - E[\theta^2|p,t_1]\right) = (\Delta t)^2\frac{d}{dt}E[\theta^2|p,t]|_{t=t_1} + O((\Delta t)^3) \quad \text{(A19)}$$

$$(\Delta t)D_p\left(Var[\theta|p,t_2] - Var[\theta|p,t_1]\right) = (\Delta t)^2\frac{d}{dt}Var[\theta|p,t]|_{t=t_1} + O((\Delta t)^3) \quad \text{(A20)}$$

Thus

$$\begin{aligned}&t_1(\Delta t)E[\theta|p,t_1]D_t + \frac{(\Delta t)^2}{2}t_1\frac{d}{dt}E[\theta|p,t_1]D_t + \frac{(\Delta t)^2}{2}E[\theta|p,t_1]D_t \\ &= \left(t_1(\Delta t) + \frac{(\Delta t)^2}{2}\right)\frac{E[\theta|p,t_1] + E[\theta|p,t_2]}{2}D_t + O((\Delta t)^3)\end{aligned}$$

$$t_1(\Delta t)Var[\theta|p,t]D_p + t_1\frac{(\Delta t)^2}{2}\frac{d}{dt}Var[\theta|p,t_1]D_p + t_1\frac{(\Delta t)^2}{2}Var[\theta|p,t_1]D_p$$
$$= \left(t_1(\Delta t) + \frac{(\Delta t)^2}{2}\right)\frac{Var[\theta|p,t_1] + Var[\theta|p,t_2]}{2}D_p + O((\Delta t)^3)$$

$$\frac{t_1^2(\Delta t)}{2}\frac{d}{dt}Var[\theta|p,t]_{t=t_1}D_p + t_1\frac{(\Delta t)^2}{2}\frac{d}{dt}Var[\theta|p,t]|_{t=t_1}D_p$$
$$= \frac{1}{2}t_1(\Delta t + t_1)\left(Var[\theta|p,t_2] - Var[\theta|p,t_1]\right)D_p + O((\Delta t)^3)$$

$$\frac{t_1(\Delta t)^2}{4}\frac{d}{dt}E[\theta|p,t]^2|_{t=t_1}D_p = \frac{t(\Delta t)}{4}\left(E[\theta|p,t_2]^2 - E[\theta|p,t_1]^2\right)D_p + O((\Delta t)^3)$$

Collecting terms using the above identities shows that

$$\begin{aligned} EB(t_2,F_{t_2}) - EB(t_1,F_{t_1}) \quad &\approx \quad \frac{\left(t_1(\Delta t) + \frac{(\Delta t)^2}{2}\right)D}{p+t_1}\left(\frac{E[\theta|p,t_1] + E[\theta|p,t_2]}{2}\varepsilon_{D,t} + \frac{Var[\theta|p,t_1] + Var[\theta|p,t_2]}{2}\varepsilon_{D,p}\right) \\ &+ \quad \frac{1}{2}t(\Delta t + t)\frac{D}{p+t_1}\left(Var[\theta|p,t_2] - Var[\theta|p,t_1]\right)\varepsilon_{D,p} \\ &+ \quad \frac{t(\Delta t)}{4}\frac{D}{p+t_1}\left(E[\theta|p,t_2]^2 - E[\theta|p,t_1]^2\right)\varepsilon_{D,p} \end{aligned}$$

Proof of Proposition 6 *Step 1.* We first show that $Var[\theta|p,\tau] \geq Var\,[\phi|p,\tau]$, where $\phi = \frac{\log(1+\theta\tau)}{\tau}$. Define $\mathcal{J}(\theta) := \theta - \phi(\theta)$. Note that $\mathcal{J}$ is increasing since $\mathcal{J}' = 1 - 1/(1+\theta\tau) > 0$. Now consider

$$\mathcal{F}(k) = Var[\theta - k\mathcal{J}(\theta)|p,t] = E\left[(\theta - k\mathcal{J}(\theta) - E[\theta - k\mathcal{J}(\theta)|p,t]))^2\,|p,t\right]$$

The derivative with respect to k is

$$\begin{aligned} \frac{d}{dk}\mathcal{F}(k) \quad &= \quad -2E\left[((\theta - k\mathcal{J}(\theta) - E[\theta - k\mathcal{J}(\theta)|p,t])\left(\mathcal{J}(\theta) - E[\mathcal{J}(\theta)|p,t]\right)\right] \\ &= \quad -2Cov\left[(\theta - k\mathcal{J}(\theta) - E[\theta - k\mathcal{J}(\theta)|p,t], \mathcal{J}(\theta) - E[\mathcal{J}(\theta)|p,t]|p,t\right] \\ &< \quad 0 \end{aligned}$$

where the inequality follows because the two random variables in the covariance operator are both increasing in θ and thus must have positive covariance. This shows that $\mathcal{F}(k)$ is decreasing in k. But note that by definition, $\mathcal{F}(0) = Var[\theta|p,\tau]$ and $\mathcal{F}(1) = Var[\phi|p,\tau]$ since $\theta - \mathcal{J}(\theta) = \phi(\theta)$. Thus $Var[\theta|p,\tau] \geq Var\,[\phi|p,\tau]$.

Intuitively, we are taking a distribution of θ, and we are modifying it by subtracting from each outcome θ a function $J(\theta)$ that is increasing in θ. Thus we are modifying the distribution in a way that pulls in the highest realizations the most toward zero—exactly the kind of operation that reduces variances.

Step 2. For each consumer marginal at price p and tax τ, and with survey response $R = r$, define $\bar{\phi}(r,p,\tau) = E[\phi|r,p,\tau]$. Note that for each pair (p,τ), the distribution of ϕ is a mean preserving spread of the distribution of $\bar{\phi}$. Thus $E\left[Var\left[\phi|p,\tau\right]\right] \geq E[Var[\bar{\phi}|p,\tau]]$.

Step 3. Let $\mu(p,\tau) = E[\phi(\theta)|p,\tau]$, and let G be the induced distribution of (p,τ). Then

$$\begin{aligned} E[Var[\bar{\phi}|p,\tau]] &= \int \left[\sum_r Pr(R=r|p,\tau)(\bar{\phi}(r,p,\tau)-\mu(p,\tau))^2 \right] dG(p,\tau) \\ &= \sum_r \int Pr(R=r|p,\tau)(\bar{\phi}(r,p,\tau)-\mu(p,\tau))^2 dG(p,\tau) \end{aligned}$$

Now

$$Pr(R=r)\int Pr(R=r|p,\tau)(\bar{\phi}(r,p,\tau)-\mu(r,p,\tau))^2 dG(p,\tau) \tag{A21}$$

$$=E_G\left[\left(Pr(R=r|p,\tau)^{1/2}\right)^2\right] E_G\left[\left(Pr(R=r|p,\tau)^{1/2}(\bar{\phi}-\mu)\right)^2\right] \tag{A22}$$

$$\geq \left(E_G\left[Pr(R=r|p,\tau)^{1/2}\cdot Pr(R=r|p,\tau)^{1/2}(\bar{\phi}-\mu)\right]\right)^2 \tag{A23}$$

$$=\left[\int (\bar{\phi}(r,p,\tau)-\mu(p,\tau))Pr(R=r|p,\tau)dG\right]^2 \tag{A24}$$

$$=\left[Pr(R=r)E[\bar{\phi}|R=r]-Pr(R=r)E[\mu|R=r]\right]^2 \tag{A25}$$

$$=Pr(R=r)^2(E[\bar{\phi}|R=r]-E[\mu|R=r])^2. \tag{A26}$$

In the computations above, line (A22) follows from line (A21) by definition. Line (A23) follows from line (A22) by the Cauchy-Schwarz inequality. And line (A25) follows from line (A24) because by definition,

$$\begin{aligned} E[\bar{\phi}|R=r] &= \frac{\int \bar{\phi}Pr(R=r,p,\tau)dG}{\int Pr(R=r,p,\tau)dG} \\ &= \frac{\int \bar{\phi}Pr(R=r,p,\tau)dG}{Pr(R=r)} \end{aligned}$$

and

$$\begin{aligned} E[\mu|R=r] &= \frac{\int \mu Pr(R=r,p,\tau)dG}{\int Pr(R=r,p,\tau)dG} \\ &= \frac{\int \mu Pr(R=r,p,\tau)dG}{Pr(R=r)} \end{aligned}$$

This implies

$$\int Pr(R=r|p,\tau)(\bar{\phi}-\mu)^2 dG \geq Pr(R=r)(E[\bar{\phi}|R=r]-E[\mu|R=r])^2$$

and thus

$$E[Var[\bar{\phi}|p,\tau]] \geq \sum_r Pr(R=r)(E[\phi(\theta)|R=r]-E[\mu|R=r])^2. \tag{A27}$$

Proof of Proposition A.1 With minor abuse of notation, we let $\theta(t)$ denote the (homogeneous) θ, as a function of t.

Part 1. Note that $\theta = D_t/D_p$. Now

$$\begin{aligned} EB'(t) &= (1-\theta)tD_t(p,t) \\ &= [1 - D_t(p,t)/D_p(p,t)]tD_t(p,t) \end{aligned}$$

Thus if $D(p,t')$ is known for all values $t' \in [t, t+\Delta t]$, $EB(t+\Delta) - EB(t)$ is identified by $\int_{t'=t}^{t'=t+\Delta} EB'(t')dt'$.

Part 2. We show that $EB'(t)$ cannot be identified if $D(p,t)$ is known only in a small neighborhood around (p,t). Because $EB'(t) = (1-\theta(t))tD_t(p,t)$, it is necessary to identify $\theta(t)$. Concretely, suppose that we observe D in the neighborhood $\mathbb{R}^+ \times [t_1, t_2]$, with $t_1 > 0$. The data is rationalized if there exist functions ψ and $\theta(t)$ such that $D(p,t) = \psi(p + \theta(t)t)$ for all p and $t \in [t_1, t_2]$. Now consider one such pair of functions ψ and θ. We show that these are not uniquely determined. In particular, consider $\tilde{\theta}(t) = \theta(t) - \epsilon t_1/t$, and $\tilde{\psi}(x) = \psi(x + \epsilon t_1)$. Then $\psi(p + \theta(t)t) = \tilde{\psi}(p + \tilde{\theta}(t)t)$ for $t \in [t_1, t_2]$, and thus θ is not uniquely identified by the data. In particular, note that while $\tilde{\theta}(t_1) < \theta(t_2)$, it is also true that $\tilde{\theta}'(t) > \theta'(t)$ for $t > t_1$.

These two different rationalizations have different welfare implications. In the first case, the marginal consumers' valuations belong to the interval $[v_1, v_2]$ such that $v_1 = p + \theta(t_1)t_1$ and $v_2 = p + \theta(t_1 + \Delta t)(t_1 + \Delta t)$. In the second case, the marginal consumers' valuations belong to the interval $[v_1', v_2']$ such that

$$v_1' = p + t_1 \cdot (\theta(t_1) - \epsilon t_1) = v_1 - t_1^2 \epsilon$$

and

$$v_2' = p + (t_1 + \Delta) \cdot [\theta(t_1 + \Delta t) - t_1 \epsilon] = v_2 - t_1(t_1 + \Delta)\epsilon.$$

Because the consumers on the margin are different, while the change in demand is the same, the efficiency costs generated by eliminating sales to these consumers must be different.

On the other hand, full knowledge of D is sufficient to identify $\theta(t)$ for each t. Simply let $\Delta p(t)$ be the value for which $D(p + \Delta p, 0) = D(p,t)$. Then $\theta(t) = \Delta p/t$.

Proof of Proposition A.2 First, we show that every demand curve $D(p,t)$ can be rationalized by assuming that $F(\theta|v,t)$ is degenerate. In particular, consider a function $\psi(p,t)$ implicitly defined by satisfying $D(p + \psi(p,t)t, 0) = D(p,t)$. Now since $v = p + \theta t$, a homogeneous $\theta(v,t)$ that rationalizes choices simply needs to satisfy $\theta(p + \psi(p,t)t, t) = \psi(p,t)$.

Alternatively rationalize $D(p,t)$ by a distribution in which a consumer has $\theta = \bar{\theta}$ with probability $q(v,t)$, and $\theta = 0$ with probability $1 - q(v,t)$. Set $\tilde{q}(p,t)$ to satisfy $\tilde{q}(p,t)D(p + \bar{\theta}t, 0) + (1 - \tilde{q}(p,t))D(p,0) = D(p,t)$. Note that because $D(p,0) \geq D(p,t) \geq (p + \bar{\theta}t, 0)$ by definition, $\tilde{q}(p,t) \in [0,1]$. Now a consumer with $\theta = \bar{\theta}$ is marginal at (p,t) if $v = p + \bar{\theta}t$. Thus $q(v,t)$ rationalizes $D(p,t)$ if $q(p + \bar{\theta}t, t)D(p + \bar{\theta}t, 0) + (1 - q(p + \bar{\theta}t, t))D(p,0) = D(p,t)$. In this case $EB'(t) = -\bar{\theta}tD_t$. Now by construction, $\psi(p,t) < \bar{\theta}$, and thus $EB'(t)$ is higher in the case with heterogeneous θ.

Finally, to establish the bounds for $t = 0$ and $\Delta t \to 0$, note that $EB(\Delta t) \to -\frac{1}{2}t^2(E[\theta|p,0]^2 + Var[\theta|p,0])D_p$ as $\Delta t \to 0$. Now $E[\theta|p,0]$ is pinned down by $D_t(p,0)/D_p(p,0)$. But the variance is highest when all consumers are either $\theta = \bar{\theta}$ or $\theta = 0$.

Proof of Proposition A.4

$$W(t) = \int_{v<p+\theta t} g_\omega Z_\omega d\tilde{F} + \int_{v\geq p+\theta t} g_\omega (Z_\omega - p - t + v) dF + \int_{v \geq p+\theta t} t\lambda dF$$

Analogous to the strategy for excess burden, define $\tilde{W}(t, \tilde{F})$ to be the welfare at a tax t given a distribution $\tilde{F}(\theta, v, \omega)$ that does not depend on t. Let $\tilde{F}(\theta, v, \omega) = F(\theta, v, \omega|t)$ here. Then

$$\begin{aligned}
\frac{d}{dt}\tilde{W} &= \int g_\omega \left[\theta Z_\omega - \theta(Z_\omega + \theta t - t)\right] d\tilde{F}(v, \theta, \omega | v = p + \theta t) \\
&- \int_{v \geq p+\theta t} g_\omega d\tilde{F} + t\lambda D_t(p,t) + \lambda D(p,t) \\
&= t \int g_\omega \theta(1-\theta) d\tilde{F}(v, \theta, \omega | v = p+\theta t) \\
&- \int_{v \geq p+\theta t} g_\omega d\tilde{F} + t\lambda D_t(p,t) + \lambda D(p,t) \\
&= -t \sum_\omega g_\omega \mu(\omega) E[\theta(1-\theta)|p,t,\omega] D_p^\omega(p,t) \\
&- \int_{v \geq p} g_\omega d\tilde{F} + \int_{p \leq v \leq p+\theta t} g_\omega d\tilde{F} + t\lambda D_t(p,t) + \lambda D(p,t) \\
&= -t \sum_\omega g_\omega \mu(\omega) E[\theta(1-\theta)|p,t,\omega] D_p^\omega(p,t) \\
&- \sum_\omega g_\omega \mu(\omega) D^\omega(p,0) - t \sum_\omega g_\omega \mu(\omega) E[\theta|p,t,\omega] D_p^\omega(p,t) + t\lambda D_t(p,t) + \lambda D(p,t) + O(t^2) \\
&= -\sum_\omega g_\omega \mu(\omega) D^\omega(p,0) - t \sum_\omega g_\omega \mu(\omega) E[2\theta - \theta^2|p,t,\omega] D_p^\omega(p,t) + t\lambda D_t(p,t) + \lambda D(p,t) + O(t^2) \\
&= -\bar{g} D(p,0) - tCov[g_\omega, 2\theta - \theta^2] D_p - 2t\bar{g} D_t + t\bar{g} E[\theta^2|p,t] D_p + t\lambda D_t + \lambda D + O(t^2) \\
&= -\bar{g} D + t(\lambda - 2\bar{g}) D_t + t\bar{g} E[\theta^2|p,t] D_p + \lambda D + tCov[g_\omega, (\theta-1)^2] D_p + O(t^2)
\end{aligned}$$

Thus

$$\frac{d^2}{dt^2}\tilde{W} = (\lambda - \bar{g}) D + (\lambda - 2\bar{g}) D_t + \bar{g} E[\theta^2|p,t] D_p + Cov[g_\omega, (\theta-1)^2] D_p + O(t)$$

A second order Taylor expansion thus implies that

$$\begin{aligned}
W(t) - W(0) &= t(\lambda - \bar{g}) D(p,0) + \frac{t^2}{2}\bar{g}\left(E[\theta|p,t]^2 + Var[\theta|p,t]\right) D_p - \frac{t^2}{2} Cov[g_\omega, 2\theta - \theta^2] D_p + t^2(\lambda - 2\bar{g}) D_t + O(t^3) \\
&= t(\lambda - \bar{g}) D(p,t) + \frac{t^2}{2}\bar{g}\left(E[\theta|p,t]^2 + Var[\theta|p,t]\right) D_p - \frac{t^2}{2} Cov[g_\omega, 2\theta - \theta^2] D_p + \frac{t^2}{2}(\lambda - 2\bar{g}) D_t + O(t^3)
\end{aligned}$$

D Additional Econometric Results

D.1 Appendix to Section 4.7: Robustness to Selection on Subject Comprehension

Let $\pi \in \{0,1\}$ denote whether the person passes the quiz question or not. Let η denote the characteristics associated with passing. Continue letting X denote the vector of covariates of θ. Let $\phi = \frac{\log(1+\theta\tau)}{\tau}$

Proposition D.1. *(Behaghel et al., 2009) Assume that* $Pr(\pi = 1|\eta, C{=}3x) \leq Pr(\pi = 1|\eta, C{=}1x)\ \forall\ \eta$*. Then*

$$\int E[\phi|\eta, \pi = 1, C{=}3x, X]dF(\eta|\pi = 1, C{=}1x, X) - \int E[\phi|\eta, \pi = 1, C{=}1x, X]dF(\eta|\pi = 1, C{=}1x, X)$$
$$\geq \frac{Pr(\pi = 1|C{=}3x)}{Pr(\pi = 1|C{=}1x)} E[\phi|\pi = 1, C{=}3x, X] - E[\phi|\pi = 1, C{=}1x, X] \tag{A28}$$

The proposition—which is derived in Jones and Mahajan (2015) and Behaghel et al. (2009)—deals with the question of how to compare average θ across conditions. Here, we use an additional monotonicity condition to derive a lower bound for the difference in average θ between conditions C=3x and C=1x. In essence, the monotonicity condition states that any subject who did not pass the comprehension check in the standard-tax arm also would not pass the comprehension check in the triple-tax arm.

Intuitively, the "worst case scenario" for the lower bound is when the study participants who pass in condition C=1x but not in condition C=3x have $\theta = 0$. The lower bound corresponds to this scenario, in which case $E[\theta|\pi = 1, \text{C=3x}, X]$ must be deflated by the ratio $\frac{Pr(\pi=1|\text{C=3x})}{Pr(\pi=1|\text{C=1x})}$ to derive the treatment effect of higher taxes for the types of study participants who pass in condition 1x. Again, the treatment effect here is the average treatment effect on the types of study participants who pass in condition C=1x in the experiment, rather than the average treatment effect on all types in the experiment.

Implementation:

To implement the lower-bound estimate (A28), we estimate three moment conditions: The first two are the moment conditions (5) and (6) for study participants who pass the comprehension questions—these give us estimates of $E[\theta|\pi = 1, \text{C=3x}]$ and $E[\theta|\pi = 1, \text{C=1x}]$. The third moment condition employs the full sample to estimate $\frac{Pr(\pi=1|\text{C=3x})}{Pr(\pi=1|\text{C=1x})}$. We use these estimates to derive the lower-bound (A28), and we use the delta method to obtain standard errors.

D.2 Within-Subject Estimation of Endogenous Attention

Let X^1_{ik} denote whether $p_2 \in [5, 10)$ for consumer i's kth product. Similarly, define X^2_{ik} to be an indicator for $p_2 \in [5, 10)$ for consumer i's kth product. For $\phi_{ik} = \frac{\log(1+\theta_{ik}\tau_i)}{\tau_i}$, we model

$$E[\phi_{ik} \quad |\mathbf{1}_{p_2\in[5,10)}, \alpha_{p_2\geq 10}\mathbf{1}_{p_2\geq 10}] = \quad \alpha_i + \alpha_{p_2\in[5,10)}X^1_{ik} + \alpha_{p_2\geq 10}X^2_{ik}$$

and

$$E[y_{ik}] \quad = \quad b_i + b_1 X^1_{ik} + b_2 X^2_{ik} + \phi_{ik}.$$

We set $\bar{\phi}_i = \frac{1}{20}\sum_k \phi_{ik}$, $\bar{\bar{y}}_i = \frac{1}{20}\sum_k y_{ik}$, $\overline{X^h_i} = \frac{1}{20}\sum_k X^h_{ik}$. From this, it follows that

$$E\left[\frac{y_{ik}-\bar{\bar{y}}_i}{1-\mathbf{1}_{tax}+\tau_i\mathbf{1}_{tax}}\right] = \frac{b_1X^1_{ik}-\overline{b_1X^1_i}}{1-\mathbf{1}_{tax}+\tau_i\mathbf{1}_{tax}}+\frac{b_2X^2_{ik}-\overline{b_2X^2_i}}{1-\mathbf{1}_{tax}+\tau_i\mathbf{1}_{tax}} + \alpha_{p_2\in[5,10)}(X^1_{ik}-\overline{X^1_i})\mathbf{1}_{tax}+\alpha_{p_2\geq 10}(X^2_{ik}-\overline{X^2_i})\mathbf{1}_{tax} \quad \text{(A29)}$$

To estimate the parameters, we proceed as before with method of moments, replacing the theoretical moment in (A29) with the empirical moment. Note that equation (A29) does not contain any of the terms α_i, and simply identifies the terms $\alpha_{p_2\in[5,10)}$ and $\alpha_{p_2\geq 10}$ using only within-consumer variation. This is analogous to estimating a linear fixed-effects model with the standard demeaning fixed-effects estimator.

D.3 Further Details for the Lower-Bound Estimation

The statistic with which we approximate the lower bound from Proposition 6 is

$$\sum_{r\in\{L,M,H\}} Pr(R=r)(\bar{\theta}_r - E[\tilde{\mu}|R=r])^2 \quad \text{(A30)}$$

To estimate (A30), we estimate each $\bar{\theta}_r$ using the empirical moment version of the left-hand-side of (7). We estimate $\tilde{\mu}(p_1,\tau)$ using the empirical moment counterpart of

$$E\left[\frac{y_{ik}-E[y_{ik}|\text{no-tax arm}, p^{ik}_1\in\mathbf{p}(p_1), \tau_i\in\boldsymbol{\tau}(\tau)]}{\tau_i}|p^{ik}_1\in\mathbf{p}(p_1),\tau_i\in\boldsymbol{\tau}(\tau)\right] \quad \text{(A31)}$$

where $E[y_{ik}|C=0\text{x}]$ denotes the average change in valuations that occurs between module 1 and module 2, and is identified from the no-tax arm. We estimate $E[\tilde{\mu}|R=r]$ by computing the empirical average over all pairs (p,τ) associated with $R=r$ in the dataset. For concreteness, we construct the estimator for the standard-tax arm. The estimator for the triple-tax arm is analogous.

We estimate $\Pr(R=r)$ by $\widehat{\Pr(R=r)} := \frac{1}{N}\sum \mathbf{1}_{R_i=r}$, where N is the number of participants in the standard-tax arm, and $\mathbf{1}_{R_i=r}$ is an indicator that consumer i's response was r. We estimate $\bar{\theta}_r$ by

$$\hat{\bar{\theta}}_r = \frac{1}{N_r}\sum_{i,k}\left[\frac{y_{ik}-E[y_{ik}|\text{no-tax arm}]}{\tau_i}|R_i=r\right].$$

where N_r is the number of consumer-product pairs associated with $R_i=r$. We estimate $E[y_{ik}|\text{no-tax arm}, \mathbf{p}\times\boldsymbol{\tau}]$, the average order effect in the no-tax arm, by

$$m(\mathbf{p}\times\boldsymbol{\tau}) := \frac{1}{|\mathbf{p}\times\boldsymbol{\tau}|_{no\ tax}}\sum_{(p^{ik}_1,\tau_i)\in\mathbf{p}\times\boldsymbol{\tau}} y_{ik}\mathbf{1}_{\text{no-tax}}$$

where $|\mathbf{p}\times\boldsymbol{\tau}|_{no\ tax}$ is the number of observations (p_1,τ) in the interval $\mathbf{p}\times\boldsymbol{\tau}$ in the no-tax arm. We estimate $\tilde{\mu}$ by

$$\widehat{\tilde{\mu}(p_1,\tau)} := \frac{1}{|\mathbf{p}(p_1)\times\boldsymbol{\tau}(\tau)|}\sum_{(p^{ik}_1,\tau_i)\in\mathbf{p}\times\boldsymbol{\tau}}\frac{y_{ik}-m(\mathbf{p}\times\boldsymbol{\tau})}{\tau_i} \quad \text{(A32)}$$

Clearly, $\widehat{\tilde{\mu}(p_1,\tau)}$ is an unbiased estimate of $E\left[\frac{y_{ik}-E[y_{ik}|\text{no-tax arm},\mathbf{p}(p_1),\boldsymbol{\tau}(\tau)]}{\tau_i}|\mathbf{p}(p_1),\boldsymbol{\tau}(\tau)\right]$. We now end by showing that this is an unbiased estimate of $\tilde{\mu}$. To see this, note that assumption A2 implies that in the standard-tax arm,

$$E\left[y_{ik}|\theta_{ik}, R_j, \mathbf{p}, \boldsymbol{\tau}\right] = E[y_{ik}|\text{no-tax arm}, \mathbf{p}, \boldsymbol{\tau}] + E[\log(1 + \theta_{ik}\tau_i)|\mathbf{p}, \boldsymbol{\tau}]$$

from which the conclusion follows by rearrangement.

Finally, to estimate $E[\mu|R = r]$, we simply take the average of $\widehat{\tilde{\mu}_{ik}}$ over all observations associated with $R = r$ in the standard-tax arm. We will call this $E[\widehat{\mu|R = r}]$. Our estimate of the variance bound is now

$$\sum_{r \in \{L,M,H\}} \widehat{Pr(R = r)} \left(\widehat{\bar{\theta}_r} - E[\widehat{\mu_{ik}|R = r}] \right)^2 \tag{A33}$$

By construction, our estimates of $Pr(R = r)$,$\bar{\theta}_r$, and $\tilde{\mu}$ are all unbiased. Note, however, that (A33) is not an unbiased estimate of the lower bound because any residual noise terms in our estimates of the moments are squared and then averaged. We estimate this mean bias with the same bootstrap procedure that we use to compute the standard errors.

E Additional Empirical Analyses and Robustness Checks

E.1 Further Tests of Module 2 Differences

Table A1: Testing for Module 2 Differences by Experimental Arm

	(1)	(2)	(3)	(4)
	OLS	0.25 Quantile	0.5 Quantile	0.75 Quantile
1x Arm	–0.09	–0.03	–0.03	–0.09
	(0.10)	(0.12)	(0.08)	(0.09)
3x Arm	–0.03	–0.05	–0.04	0.21
	(0.10)	(0.12)	(0.08)	(0.16)
Observations	59960	59960	59960	59960

Notes: This table tests for differences in module 2 willingness to pay for products by experimental arm. Column 1 reports estimates from an OLS regression. Columns (2)-(4) report 0.25, 0.5, and 0.75 quantile regressions. Standard errors, clustered at the subject level, reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

E.2 Demographic covariates

We analyze how θ covaries with standard demographics provided by ClearVoice—race, age, educational attainment, marital status—as well as the three additional covariates that we collected in our experiment, described below:

Household Income. Participants were also asked to state their household income. We analyze the data by income quartiles, the cutoffs for which are 28k, 50k, and 82k, which match almost exactly to the 2010 US census data.[2]

Ability to compute taxes / Numeracy. Immediately after the survey question about the sales tax rate, consumers were asked to compute the sales tax (in absolute terms) on an $8 (non-tax-exempt) item. We code answers as correct if consumers provide the correct answer using their perceived sales tax rate. For example, if the true sales tax rate is 6%, but the consumer thinks that it is 7%, then an answer is coded as being correct if it is less than 1 cent from $0.56. Consumers were asked to answer this question in the format of $0.56. However, as with the question about sales tax beliefs, not all consumers followed the instructions. Some

[2] According to the 2010 census, the quartile thresholds are 25k, 50k, 90k.

consumers seemed to have entered their answers in the format of $8.56 instead of $0.56. Other consumers seem to have entered their answers as 56 instead of $0.56. For consumers whose answers are between 8 and 12 (about 10% of consumers), we recode answers by subtracting 8, as we think it is implausible that anyone would think that the tax on an $8 item would be greater than $8. For consumers whose answers are above 20, we recode their answers by dividing by 100, as these consumers most likely entered their answers in number of cents rather than dollars. Our results are robust to simply excluding consumers with answers above 8. Overall, accuracy was very high, with 73% of consumers giving the right answer. That underreaction persists despite this high level of accuracy shows that consumers are either deliberately choosing not to compute the taxes, or are simply forgetting to think about taxes when determining their willingness to pay. This idea that consumers seem to make "bad" decisions despite knowing how to make "good" ones is broadly consistent with the results reported in Ambuehl et al. (2016) and Zimmermann and Enke (2015) for other types of financial decisions.

Financial Sophistication. We use the "Big Three" financial literacy questions (Lusardi and Mitchell, 2008, 2014). The three multiple choice questions test for understanding of interest rates, inflation, and risk diversification.[3] We code participants as financially sophisticated if they answer all three questions correctly. Overall, 49% of consumers in our final sample answered all three questions correctly.[4] The three questions are as follows:

1. Suppose you had $100 in a savings account and the interest rate was 2 percent per year. After 5 years, how much do you think you would have in the account if you left the money to grow? a) More than $102 b) Exactly $102 c) Less than $102 d) Do not know

2. Imagine that the interest rate on your savings account was 1 percent per year and inflation was 2 percent per year. After 1 year, would you be able to buy more than, exactly the same as, or less than today with the money in this account? a) More than today b) Exactly the same as today c) Less than today d) Do not know

3. Do you think that the following statement is true or false? "Buying a single company stock usually provides a safer return than a stock mutual fund." a) True b) False c) Do not know

[3]Previous work has shown that financial literacy is associated with mistakes in other domains, including incurring overdraft fees (Stango and Zinman, 2014), incorrectly valuing annuities (Brown et al., forthcoming), and not saving enough for retirement (Lusardi and Mitchell, 2007a,b).

[4]Our measure of tax numeracy and financial sophistication are strongly correlated. Financially sophisticated consumers have a 12 percentage point grater likelihood of correctly answering the tax computation question ($p < 0.01$).

Table A2: Average θ (Weight Placed on Tax) by Demographics

	(1) Standard	(2) Triple	(3) Pooled
Compute Tax Correctly	0.123	0.132	0.138 *
	(0.204)	(0.084)	(0.084)
Financially Sophisticated	0.439 *	0.206 **	0.201 **
	(0.224)	(0.090)	(0.090)
Income Quartile 2	–0.058	0.006	–0.003
	(0.277)	(0.115)	(0.115)
Income Quartile 3	0.237	0.123	0.127
	(0.291)	(0.121)	(0.120)
Income Quartile 4	0.066	0.247 *	0.239 *
	(0.295)	(0.134)	(0.132)
Age	–0.022 ***	–0.010 ***	–0.009 ***
	(0.007)	(0.003)	(0.003)
Male	–0.026	–0.044	–0.040
	(0.210)	(0.085)	(0.085)
Married	–0.183	–0.130	–0.135
	(0.230)	(0.092)	(0.092)
College Degree	–0.159	0.046	0.037
	(0.218)	(0.095)	(0.094)
Asian	–0.607	–0.101	–0.092
	(0.525)	(0.235)	(0.237)
Caucasian	0.453	0.008	0.018
	(0.321)	(0.147)	(0.145)
Hispanic	0.737	–0.303	–0.227
	(0.517)	(0.253)	(0.252)
African American	0.664	0.008	0.001
	(0.458)	(0.191)	(0.190)
Observations	38010	36790	54643

Notes: This table displays method of moments estimates of the relationship between average θ and demographic covariates. θ is defined as the "weight" that consumers place on the sales tax, with $\theta = 0$ corresponding to complete neglect of the tax and $\theta = 1$ corresponding to full optimization. Standard errors, clustered at the subject level, reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

E.3 Average θ of Marginal Consumers

We estimate a linear model for average θ given by $E[\theta] = \theta_A + \alpha \mathbf{1}_{\text{marginal}}$, where $\theta_A \in \{\theta_{1X}, \theta_{3X}\}$ is the constant for either the triple-tax or standard-tax arm, $\mathbf{1}_{\text{marginal}}$ is an indicator for whether the consumer is labeled "marginal" at the actual Amazon.com price, and α is the scalar corresponding to how that affects average θ. We label consumers are "marginal" if the the module 2 price at which they are marginal is within Δ of the actual Amazon.com price, where Δ takes on values of \$1, \$2, \$3.

Table A3: Difference in average θ between marginal and non-marginal consumers

	(1) Within \$1	(2) Within \$2	(3) Within \$3
Std. tax base θ	0.216**	0.191**	0.189**
	(0.092)	(0.091)	(0.094)
Triple tax base θ	0.460***	0.442***	0.422***
	(0.041)	(0.042)	(0.044)
Difference for marginal consumers	0.102**	0.121***	0.148***
	(0.047)	(0.045)	(0.045)
Observations	58478	58478	58478

Notes: The first column defines consumers as "marginal" if their module 2 price is within \$1 of the actual Amazon.com price. Columns 2 and 3 are constructed analogously for \$2 and \$3. Standard errors, clustered at the subject level, reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

E.4 Further Tests of Endogenous Attention Using Amazon.com Prices

We estimate a linear model for average θ given by $E[\theta] = \theta_A + \alpha p_1$, where $\theta_A \in \{\theta_{1X}, \theta_{3X}\}$ is the constant for either the triple-tax or standard-tax arm, p_1 is the module 1 price, and α is the scalar corresponding to how that effects average θ. We instrument p_1 using the prices on Amazon.com for the corresponding products at the time that the experiment was run.

Table A4: Average θ as a function of prices, using Amazon.com prices as instruments

	(1) Pooled	(2) Standard	(3) Triple
Std. tax cons.	0.123	0.157	
	(0.131)	(0.263)	
Triple tax cons.	0.357***		0.356***
	(0.092)		(0.093)
Impact of price	0.020	0.015	0.019
	(0.014)	(0.040)	(0.014)
Observations	58478	40651	39378

Notes: Standard errors, clustered at the subject level, reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

E.5 Replication of Table 7 Without Excluding Study Participants Failing Comprehension Questions

Table A5: Average θ (Weight Placed on Tax) Conditional on Self-Classifying Survey Response

	(1) Standard	(2) Triple
"Yes" average θ	0.616***	0.625***
	(0.229)	(0.081)
"A little" average θ	0.289***	0.410***
	(0.097)	(0.040)
"No" average θ	–0.246*	–0.027
	(0.126)	(0.045)
Observations	54503	54988

Notes: This table replicates Table 7, but does not exclude study participants who failed comprehension checks. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

E.6 Replication of Main Results with OLS Regressions

As an alternative to our GMM approach, one could instead proceed with a simple OLS framework given by:

$$E[y] = \alpha_0 + \alpha_1\tau + \alpha_2\tau X + \beta_1 X. \tag{A34}$$

Because the linear model is misspecified when underreaction is endogenous to the tax rate, an estimate $\hat{\alpha}_1$ obtained from equation (A34) is not a consistent estimate of $E[\theta_{ik}]$. In particular, the OLS estimate $\hat{\alpha}_1$ will depend on how much less consumers underreact to large taxes than to small taxes. To take a concrete illustration, we estimate an average θ of approximately 0.25 and 0.48 in the standard- and triple-tax arms, respectively, and thus obtain an average θ of 0.37 in the pooled sample. If we simply estimate (A34) using the OLS estimator, however, we get an $\hat{\alpha}_1$ of 0.49—an estimate that is higher than either average and does not have a clear economic interpretation. While this problem can be reduced by allowing for different coefficients on the tax rate across the standard- and triple-tax arms, there remains natural variation in tax rates within each experimental arm. This variation induces the same endogeneity concern—although the variation in natural tax rates is smaller, and thus the bias induced by this variation is less dramatic in magnitude.

While these concerns lead us to prefer the GMM approach as a primary specification, our qualitative results are also obtained when estimating (A34) using OLS, as we show below.

Table A6: Estimates of Average θ (Weight Placed on Tax) by Experimental Arm
Dependent variable: $y_{ik} = \log(p_2^{ik}) - \log(p_1^{ik})$

	(1) All	(2) $p_2 \geq 1$	(3) $p_2 \geq 5$
tax × *standard*	0.254**	0.252***	0.202**
	(0.105)	(0.088)	(0.082)
tax × *triple*	0.487***	0.479***	0.532***
	(0.043)	(0.038)	(0.039)
Observations	59960	58478	32810

Notes: This table replicates Table 2, implemented through an analogous OLS procedure. Standard errors, clustered at the subject level, reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A7: Average θ (Weight Placed on Tax) for Different Product Valuations
Dependent variable: $y_{ik} = \log(p_2^{ik}) - \log(p_1^{ik})$

	(1) Standard	(2) Triple	(3) Pooled	(4) Standard	(5) Triple	(6) Pooled
tax × Middle p_2 bin	–0.127	0.102**	0.109**	–0.059	0.085**	0.093***
	(0.133)	(0.051)	(0.051)	(0.097)	(0.036)	(0.036)
tax × High p_2 bin	0.055	0.135*	0.133*	0.074	0.061	0.050
	(0.164)	(0.070)	(0.071)	(0.133)	(0.048)	(0.048)
tax × std. tax arm	0.286**		0.177*			
	(0.133)		(0.094)			
tax × triple tax arm		0.415***	0.412***			
		(0.052)	(0.051)			
Fixed effects	No	No	No	Yes	Yes	Yes
Observations	40651	39378	58478	40651	39378	58478

Notes: This table replicates Table 3, implemented through an analogous OLS procedure. Standard errors, clustered at the subject level, reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A8: Average θ (Weight Placed on Tax) Conditional on Self-Classifying Survey Response
Dependent variable: $y_{ik} = \log(p_2^{ik}) - \log(p_1^{ik})$

	(1) Standard	(2) Triple
tax × Yes	1.068***	0.923***
	(0.281)	(0.096)
tax × A little	0.435***	0.627***
	(0.104)	(0.047)
tax × No	–0.182	0.047
	(0.122)	(0.052)
Observations	40651	39378

Notes: This table replicates Table 7, implemented through an analogous OLS procedure. Standard errors, clustered at the subject level, reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

E.7 Replication of Main Results Excluding Study Participants Not Understanding the BDM Mechanism

Table A9: Estimates of Average θ (Weight Placed on Tax) by Experimental Arm

	(1) All	(2) $p^2 \geq 1$	(3) $p^2 \geq 5$
Std. tax avg. θ	0.277**	0.274***	0.262***
	(0.126)	(0.103)	(0.094)
Triple tax avg. θ	0.524***	0.513***	0.595***
	(0.052)	(0.044)	(0.046)
Observations	46540	45372	25658

Notes: This table replicates Table 2, dropping study participants who failed comprehension checks about the BDM mechanism.* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A10: Average θ (Weight Placed on Tax) for Different Product Valuations

	(1) Standard	(2) Triple	(3) Pooled	(4) Standard	(5) Triple	(6) Pooled
Middle p_2 bin	–0.153	0.156**	0.172***	–0.078	0.123***	0.132***
	(0.155)	(0.061)	(0.061)	(0.115)	(0.041)	(0.041)
High p_2 bin	0.253	0.229***	0.256***	0.183	0.095*	0.102*
	(0.206)	(0.083)	(0.082)	(0.171)	(0.054)	(0.054)
Std. tax cons.	0.296*		0.103			
	(0.159)		(0.105)			
Triple tax cons.		0.410***	0.394***			
		(0.061)	(0.060)			
Fixed effects	No	No	No	Yes	Yes	Yes
Observations	31319	30363	45372	31319	30363	45372

Notes: This table replicates Table 3, dropping study participants who failed comprehension checks about the BDM mechanism. All specifications condition on $p_2 \geq 1$. Standard errors, clustered at the subject level, reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A11: Average θ (Weight Placed on Tax) Conditional on Self-Classifying Survey Response

	(1) Standard	(2) Triple
"Yes" average θ	0.844***	0.984***
	(0.298)	(0.113)
"A little" average θ	0.572***	0.661***
	(0.127)	(0.054)
"No" average θ	–0.219*	0.066
	(0.131)	(0.064)
Observations	31319	30363

Notes: This table replicates Table 7, dropping study participants who failed comprehension checks about the BDM mechanism. Column (1) provides estimates for the standard-tax arm, Column (2) provides estimates for the triple-tax arm. All specifications condition on $p_2 \geq 1$. Standard errors, clustered at the subject level, reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

F Items Used in the Study

Product	**Amazon.com price (as of Feb 2015)**	**Amazon.com Product Description**
RainStoppers 68-Inch Oversize Windproof Golf Umbrella	$12.61	This RainStoppers 68" oversize golf umbrella is large enough to cover three or more people. Umbrella frame constructed with fiberglass shaft and ribs for maximum stability. Canopy is made of 190T Nylon fabric. Complete with a foam non slip handle. Matching sleeve included. Length when closed is 43".
Energizer AA Batteries max Alkaline 20-Pack	$11.15	energizer AA max alkaline batteries 20 pack super fresh, Expiration Date: 2024 or better. Packed in original Energizer small box 4 batteries per box x 5 boxes total 20 batteries.
Glad OdorShield Tall Kitchen Drawstring Trash Bags, Fresh Clean, 13 Gallon, 80 Count	$12.79	Glad OdorShield Tall Kitchen Drawstring Trash Bags backed by the power of Febreze are tough, reliable trash bags that neutralize strong and offensive odors for lasting freshness. These durable bags are great for use in the kitchen, home office, garage, and laundry room.
Admiral Blue 100% Cotton Bath Towel - 27 x 52 Inches	$14.99	There isn't much that's better than stepping out of a refreshing shower and wrapping yourself in the soft, Luxury Bath Towels. Now you can have that feeling every single day. It won't just be a treat anymore; it'll be your way of life. These extra-absorbent 100% cotton towels can be just hanging around waiting for you, ready to fulfill their duty in making you feel pampered. Not only practical but also stylish, these towels will also add a fashionable and luxurious touch to your bathroom.

Martex Egyptian Cotton Hand Towel with Dry-Fast (French Blue)	$6.79	Martex is one of the oldest and most trusted names in bath products. This towel is made of loops of 100% Egyptian cotton which offers the absorbency and quality of this fine extra-long-staple fiber. The towel offers DryFast Technology. Enjoy a broad color palette to compliment any bathroom decor.
Pilot G2 Retractable Premium Gel Ink Roller Ball Pens, Fine Point, Black Ink, Dozen Box (31020)	$11.89	Discover the smooth writing and comfortable G2, America's #1 Selling Gel Pen*. G2 gel ink writes 2X longer than the average of branded gel ink pens**. The G2 product line includes four point sizes, fifteen color options, and multiple barrel styles to suit every situation and personality. It is the only gel pen that offers this level of customization–because after all, pens aren't one size fits all.
Scotch-Brite Heavy Duty Scrub Sponge 426, 6-Count	$7.73	O-Cel-O™ sponges and Scotch Brite scrubbers are truly a fashion-meets-function success story. The highly absorbent and durable sponges come in different sizes and scrub levels for the various surfaces around the home. Their assorted colors and patterns follow the current fashion trends to create the perfect accent in any room.
Febreze Fabric Refresher Spring & Renewal Air Freshener, 27 Fluid Ounce	$4.94	When it comes to your home, you should never settle for less than fresh. Febreze Fabric Refresher is the first step to total freshness in every room. The fine mist eliminates odors that can linger in fabrics and air, leaving behind nothing but a light, pleasing scent. With Febreze Fabric Refresher, uplifting freshness is a simple spray away.
Microban Antimicrobial Cutting Board Lime Green - 11.5x8 inch	$8.99	The Microban cutting board from Uniware is the perfect cutting board for the health conscious. The cutting board has a soft grip with handle and is dishwasher safe. The cutting board can be reversible, use on both sides, and is non-porous, non absorbent. The rubber grips prevents slipping on countertop. Doesn't dull knives, juice-collecting groove. Microban is the most trusted antimicrobial product protection in the world. Built-In defense that inhibits the growth of stain and odor causing bacteria, mold, and mildew. Always works to keep the cutting board cleaner between cleanings. Lasts throughout the lifetime of the cutting board. Size: 11.5"x8" Color: Lime Green.
Nordic Ware Natural Aluminum Commercial Baker's Half Sheet	$11.63	Nordic Ware's line of Natural Commercial Bakeware is designed for commercial use, and exceeds expectations in the home. The durable, natural aluminum construction bakes evenly and browns uniformly, while the light color prevents overbrowning. The oversized edge also makes getting these pans in and out of the oven a cinch. Proudly made in the USA by Nordic Ware
Gain with FreshLock HE Original Liquid Detergent, 100 Fl Oz	$9.97	The scent of Gain Original liquid laundry detergent brings a lively scent to your laundry room. Powerful Lift & Lock Technology lifts away dirt and stains so you can lock in the amazing scent you love. With bursts of citrus, a green twist, and just enough floral fragrance, you'll wish laundry day came more often.

Rubbermaid Configurations Folding Laundry Hamper, 23-inch, Natural (FG4D0602NATUR)	$12.99	Rubbermaid Configurations Folding Laundry Hamper, 23-inch, Natural (FG4D0602NATUR). Makes it easy to add hamper space to any Rubbermaid Configurations Kit. Collapses for easy storage. Neutral two-tone canvas is breathable and stylish. Coordinates with other items in Rubbermaid Configurations collection. For nearly 80 years, Rubbermaid has represented innovative, high-quality products that help simplify life. Recognized as a “Brand of the Century” for its impact on the American way of life.
Scotch Precision Scissor, 8-Inches (1448)	$5.44	Scotch Precision 8" Scissors come with the finest quality stainless steel blades for a sharp edge and long cutting life. These scissors also comes with a soft grip handles for ease of use. Great for everyday cutting needs. Comes with a limited lifetime warranty.
Clorox Company 00450 Gw All Purpose Cleaner, 32-Ounce	$8.09	Cuts through grease, grime and dirt as well as traditional cleaners. Spray on counters, appliances, stainless steel, sealed granite, chrome, cook top hoods, sinks and toilets. Made from plants and minerals, 99 percent natural, so it leaves no harsh chemical fumes or residue.
Rubbermaid Easy Find Lid Medium Value Pack Food Storage Containers	$10.20	The Rubbermaid Easy Find Lids Medium Value Pack includes (2) 3.0 cup Easy Find Lid containers measures 7" x 7" x 2.3"and (1) 5.0 cup Easy Find Lid container measures 7" x 7" x 3.4". The number one unmet need for food storage is container and lid organization. With Rubbermaid’s new Easy Find Lids you’ll find storage and organization a breeze! The Easy Find Lids snap together as well as snap to the bases for easy storage. The Easy Find Lids and bases also nest together making storage in a cabinet or a drawer much more efficient. Easy Find Lids are square in shape and allow for easy of stacking when placed on shelves or in the refrigerator. Rubbermaid Easy Find Lids also feature a super clarified base which takes the guessing out of what’s inside and allows you to see what’s inside quickly and easily. Rubbermaid Easy Find Lids and bases are also microwave, freezer, and dishwasher safe.
Rubbermaid Lunch Blox medium durable bag - Black Etch	$10.47	The Rubbermaid 1813501 Lunch Blox medium durable bag - Black Etch is an insulated lunch bag designed to work with the Rubbermaid Lunch Blox food storage container system. The bag is insulated to achieve the maximum benefit of Blue Ice blocks and keep your food cold. The bag features a bottle holder, side pocket, comfort-grip handle and removable shoulder strap. The lunch Blox bag is durable and looks good for both the professional bringing their lunch to work or the kid taking their lunch to school.
Libbey 14-Ounce Classic White Wine Glass, Clear, 4-Piece	$12.99	Great for any party, this set includes four 14-ounce clear classic white wine glasses which match perfectly with the classic collection by libbey. The glasses are dishwasher safe and made in the USA.
Fulcrum 20010-301 Multi-Flex LED Task Light and Book Light	$9.47	The Multi Flex Light is an all-purpose book light, task light or travel light

Envision Home Microfiber Bath Mat with Memory Foam, 16 by 24-Inch, Espresso	$10.82	Enjoy spa luxury at home with the Envision Home Microfiber Bath Mat, featuring memory foam! Designed to absorb water like a sponge and help protect floors from damaging puddles of water, your feet will love stepping on to this soft cushion of memory foam encased in super-absorbent microfiber. The Microfiber Bath Mat starts with fibers that are split down to microscopic level, resulting in tiny threads that love to absorb every drop of water. Because of this increased surface area, this microfiber mat can collect more water than an ordinary bath mat. Plus, it dries unbelievably fast.The soft memory foam interior provides a comfortable and warm place to stand, or when kneeling to bathe a child or pet, preventing aches and pains. The seams across the mat allow for it to be easily folded for storage, or simply hang it from the convenient drying loop. It is available in three colors to compliment your personal décor and style – Cream, Celestial and Espresso.Caring for your Microfiber Bath Mat is easy; simply toss it in the washing machine with cold water and a liquid detergent and then place in the dryer on a low heat setting. The Microfiber Bath Mat is just one of the many impressive items offered in the Envision Home Collection. Designed to make it easier to take care of the home, our innovative, high-value and superior-quality products provide cleaning, kitchen, bath, laundry and pet solutions to solve life's little dilemmas.
Carnation Home Fashions Hotel Collection 8-Gauge Vinyl Shower Curtain Liner with Metal Grommets, Monaco Blue	$8.99	Protect your favorite shower curtain with our top-of-the-line Hotel Collection Vinyl Shower Curtain Liner. This standard-sized (72" x 72") liner is made with an extra heavy (8 gauge), water repellant vinyl that easily wipes clean. With metal grommets along top of the liner to prevent tearing. Here in Monaco Blue, this liner is available in a variety of fashionable colors. With its wonderful features and fashionable colors, this liner could also make a great shower curtain

Note: These are Amazon.com prices as they were displayed to, and documented by, our research assistant in February 2015. Prices may vary over time or by geographic region.

References

Ambuehl, Sandro, B. Douglas Bernheim, and Annamaria Lusardi, "The effect of financial education on the quality of decision making," *working paper*, 2016.

Behaghel, Luc, Bruno Crépon, Marc Gurgand, and Thomas Le Barbanchon, "Sample Attrition Bias in Randomized Experiments: A Tale of Two Surveys," *Paris School of Economics Working Paper No. 2009-15*, 2009.

Brown, Jeffrey R., Arie Kapteyn, Erzo F.P. Luttmer, and Olivia S. Mitchell, "Cognitive Constrains on Valuing Annuities," *Journal of the European Economic Association*, forthcoming.

Chetty, Raj, Adam Looney, and Kory Kroft, "Salience and Taxation: Theory and Evidence," *American Economic Review*, 2009, *99* (4), 1145–1177.

Jones, Damon and Aprajit Mahajan, "Time-Inconsistency and Saving: Experimental Evidence from Low-Income Tax Filers," *NBER Working Paper No. 21272*, 2015.

Lockwood, Benjamin B. and Dmitry Taubinsky, "Regressive Sin Taxes," *Working Paper*, 2015.

Lusardi, Annamaria and Olivia S. Mitchell, "Baby Boomer Retirement Security: The Roles of Planning, Financial Literacy, and Housing Wealth," *Journal of Monetary Economics*, 2007, *51* (1), 205–224.

_ **and** _ , "Financial Literacy and Retirement Preparedness: Evidence and Implications for Financial Education," *Business Economics*, 2007, *42* (1), 35–44.

_ **and** _ , "Planning and Financial Literacy: How do Women Fare?," *American Economic Review*, 2008, *98* (2), 413–417.

_ **and** _ , "The Economic Importance of Financial Literacy: Theory and Evidence," *Journal of Economic Literature*, 2014, *52* (1), 5–44.

Stango, Victor and Jonathan Zinman, "Limited and Varying Consumer Attention: Evidence from Shocks to the Salience of Bank Overdraft Fees," *Review of Financial Studies*, 2014, *27* (4), 990–1030.

Zimmermann, Florian and Benjamin Enke, "Correlation Neglect in Belief Formation," *Working Paper*, 2015.