Are Sugar-Sweetened Beverage Information Labels Well-Targeted?

NCT05038163

Statistical Analysis Plan

April 1, 2021

Are Sugar-Sweetened Beverage Information Labels Well-Targeted? Evidence and Welfare Implications

Sample Size and Power

The goal of this experiment is not to accept or reject a particular hypothesis, but rather to measure the targeting properties and welfare effects of information labels with statistical precision. Thus, our power calculations are based on 95%-confidence intervals.

We use related preliminary work to estimate our standard errors for this project; a summary of this early work is available at the end of this document. In our preliminary work, each label was assessed with approximately 180 participants, with 90 assigned to a control arm and to each of two labels, and the 95%-confidence intervals for the treatment effects have width of about $0.50 per item, or roughly 25% of the average estimated bias. For our proposed project, a sample size of approximately 3,000 subjects will give us a total sample of 1,000 to assess each label (i.e., 500 from the control and 500 from the treatment group). Assuming the same variances as in our preliminary work, this will give us standard errors that are $\sqrt{1000/180} \approx 2.4$ times smaller than in our preliminary work. Our 95%-confidence intervals for our treatment effects will have width of approximately $0.21 per item, or roughly 10% of the average estimated bias. Under the same assumptions, our 95%-confidence intervals for our targeting measure (i.e., the covariance) will be approximately 2.4 times smaller than in our preliminary work and have width of approximately 0.26 dollars squared.

For the subsample of older adults, we will have approximately 250 participants to assess each label (i.e., 125 from the control and 125 from the treatment group). If we assume the variances on the treatment effects and measure of targeting are the same as for the full sample, then we can again predict power. Under this assumption, we will have standard errors that are $\sqrt{250/180} \approx 1.4$ times smaller than in our preliminary work. Our 95%-confidence intervals for our treatment effects will have width of approximately $0.36 per item, or roughly 20% of the average estimated bias for the full population. Similarly, our 95%-confidence intervals for our targeting measure (i.e., the covariance) will have width of approximately 0.45 dollars squared.

Statistical Design

Theoretical model and framework. There is a unit mass of consumers with binary demand for a good x. We let v denote consumers' true value of the good, and $v+\gamma$ their perceived value. In the absence of an information label, consumers thus buy the good if $v+\gamma \geq p$. The bias γ can covary arbitrarily with true values. We model an information label as a set of treatment effects τ that covary arbitrarily with v and γ, such that consumers buy the good if $v+\gamma+\tau \geq p$. The only assumption we make about the joint distribution $F(v,\gamma,\tau)$ is that it produces smooth demand curves. Even though we write post-intervention perceived valuation as $v+\gamma+\tau$, that does not presuppose that treatment effects are "additive" in the sense of being orthogonal to v and/or γ – they can covary arbitrarily with v and γ. We let $D(p)$ denote the demand curve in the absence of a label, and $D_p(p)$ denote its derivative with respect to p.

Following Weyl (2013), we consider symmetric competition in which the cost of producing quantity q_j for firm j is $c(q_j)$. We limit to symmetric equilibria in which $q_j \equiv q$ in equilibrium, and we let $p(q)$ denote the resulting market price. The elasticity of demand is given by $\varepsilon_D = -\frac{-p}{qp'(q)}$.

Weyl (2013) show that a wide variety of models of firm interactions can be captured by the Lerner index $\theta \coloneqq \frac{p-c'(q)-t}{p}\varepsilon_D$, where t is the tax levied on producers. For example, homogeneous product Bertrand competition corresponds to the case $\theta = 0$. Cournot competition corresponds to the case in which $\theta = 1/n$. Spanning these special cases, we assume that θ is constant. We define ρ, the pass-through rate, to be the degree to which price rises for consumers when a tax t is imposed on producers: $\rho \coloneqq \frac{dp}{dt}$.

Our main result is as follows:

Proposition *1. Assume that (1) that* $|\gamma|$ *is bounded, (2) that* $|\tau|$ *is bounded by* $\bar{\tau} <$*, and (3) that terms of order* $D_{pp}\bar{\tau}^2$ *are negligible. Let* $\mu = p - c'(q)$ *denote the mark-up. Then:*

1. *The change in welfare from an information label in a market with no pre-existing taxes and subsides is given by:*

$$\Delta W = \left(\frac{1}{2}Var[\tau] + Cov[\tau,\gamma]\right)D_p(p) + \frac{\rho}{2}(E[\tau]^2 + 2E[\tau]E[\gamma])D_p(p) - \rho\mu E[\tau|p]D_p(p)$$

2. *When the government can additionally set an optimal tax or subsidy, the incremental welfare from the information label is:*

$$\left(\frac{1}{2}Var[\tau] + Cov[\tau,\gamma]\right)D_p(p)$$

The proposition states that the welfare effects of the information label depend on two terms: the heterogeneity of treatment effects, $Var[\tau]$, and the extent to which the treatment effects are well-targeted, $Cov[\tau,\gamma]$. Welfare is decreasing in $Var[\tau]$ holding constant $Cov[\tau,\gamma]$, as this large un-targeted heterogeneity generates misallocation in which individuals purchase which products. the extent to which it reduces the variances of bias: $Var[\tau+\gamma] - Var[\gamma]$. To obtain some intuition for this result, note that $\frac{1}{2}Var[\tau] + Cov[\tau,\gamma] = Var[\tau+\gamma] - Var[\gamma]$. Since $\tau+\gamma$ represents the "post-label" bias, the welfare effects of the label are thus proportional to how much the label reduces (or increases) the variance of bias in the population.

In the absence of a financial instrument, the welfare effects of the label also depend on $E[\tau]^2 + 2E[\tau]E[\gamma]$. Note that these terms are functions of $E[\tau]$, the average treatment effect. Intuitively, even if the label does not reduce the variance of bias, and instead simply functions as a blunt policy instrument that affects everyone in the same way, it is still useful because it functions as a substitute for a tax. The term $\rho\mu E[\tau|p]D_p$ simply corresponds to the extent to which labels counteract (or amplify) market power. E.g., if firms price too high, then a modest bias that leads consumers to overestimate the marginal value of the product may be optimal.

Econometric details. Recall that in the experiment, all subjects first make a decision in the absence of any labels (decision set $j = 1$), and then again with the potential presence of a label (decision set $j = 2$). We decompose willingness to pay of person i in decision set $j \in \{1,2\}$ for product k in label treatment T as follows:

$$w_{ijk}(T) = v_{ijk} + \gamma_{ik} + \tau_{ik}(T) \cdot I(j = 2)$$

where $\tau_{ik}(T)$ is the treatment effect on person i for product k, γ_{ik} is the bias, and v_{ijk} is the true cost or benefit of getting the sugary rather than the sugarless drink. Our multiple price lists provide direct estimates of $w_{ijk}(T)$. We obtain estimates of bias, $\widehat{\gamma_{ik}}$, using the methodology of Allcott, Lockwood, and Taubinksy (2019a) and the stage 1 survey questions. Before we detail the methodology of obtaining these estimates, we summarize the approach we take with those estimates in hand.

We make the following identifying assumptions, which essentially state that the measurement error in $\hat{\gamma}$ is independent of the treatment effects.

$E[v_{ijk}|T]$ does not depend on T or on j

$Cov[v_{i2k} - v_{i1k}, \tau_{ik}] = 0$

$Cov[\gamma_{ijk} - \widehat{\gamma_{ijk}}, \tau_{ik}] = 0$

$Cov[\gamma_{ip} - \widehat{\gamma_{ik}}, v_{ijk} - v_{ij'k}] = 0$

Under these assumptions, we obtain the following equation for the average (across the six products) variance of a given label:

$$\frac{1}{6}\sum_j Var[\tau_j|T=t] = \frac{1}{6}\sum_j (Var[\Delta w_{ij}|T=t] - Var[\Delta w_{ij}|T=0])$$

Where $\Delta w_{ij} \coloneqq w_{ij2} - w_{ij1}$is the difference in WTP for the same product-pair \$j\$ before seeing the label versus after seeing the label.

We also obtain the following formula for the average covariance of the label with the bias:

$$\frac{1}{6}\sum_j Cov[\tau_j, \gamma_j|T=t] = \frac{1}{6}\sum_j (Cov[\Delta w_{ij}, \widehat{\gamma_i}|T=t])$$

From this, it we can also immediately obtain the following statistics:

$$Var[\tau+\gamma] - Var[\gamma] = Var[\tau] + 2Cov[\tau,\gamma]$$

$$E[(\tau+\gamma)^2] - E[\gamma^2] = Var[\tau] + 2Cov[\tau,\gamma] + E[\tau]^2 + 2E[\tau]E[\gamma]$$

Estimating bias. We estimate each participant's bias using methods from Allcott, Lockwood, and Taubinksy (2019a). Specifically, we assume bias is driven by a lack of nutritional knowledge and/or a lack of self-control.

To estimate each individual's nutritional knowledge, we compute a nutrition score, $nutrition_i$. This score ranges from 0 to 1, where higher values indicate more nutritional knowledge. To get this score, in the first stage of the experiment we ask participants 24 questions about their nutritional knowledge. Some of these questions include multiple parts, leading to a total of 59 sub-questions. As in Allcott, Lockwood, and Taubinksy (2019a), we compute their nutritional knowledge score by calculating how many of the 59 sub-questions they answered correctly, and converting this into a percentage. In our preliminary work, the average score is 68%, the 10th percentile is 44%, the 25th is 61%, the 50th is 69%, the 75th is 78%, and the 90th is 85%.

To estimate each individual's self-control $selfcontrol_i$, we ask participants to answer the following question: "Please indicate how much the following statement reflects how you typically are: I drink soda pop or other sugar-sweetened beverages more often than I should." We then convert this response to a 0 to 1 scale (0 = Definitely, 1/3 = Mostly, 2/3 = Somewhat, 1 = Not at all).

We convert the proxy variables into a proxy for bias via equation (26) fromAllcott, Lockwood, and Taubinksy (2019a):

$$\widehat{\gamma_i} = [\widehat{\tau_1} \cdot (0.92 - nutrition_i) + \widehat{\tau_2} \cdot (1 - selfcontrol_i)] \cdot p_i / \widehat{\zeta_i^c}$$

Here 0.92 is the average score among nutrition experts to the nutrition knowledge quiz (per Allcott, Lockwood and Taubinsky, 2019a) and 1 represents perfect self-control.

While we eventually plan to estimate using the Homescan data, for now we use the point estimates from Allcott, Lockwood and Taubinsky (2019a). Specifically, we use $\widehat{\tau_1} = 0.854$ and $\widehat{\tau_2} = 0.825$, as reported in Table 5, column 1 of their paper. We also set $p = 3.63$ and \$$\widehat{\zeta^c} = 1.39$, the averages reported in Table 6 of their paper. Finally, we rescale so that bias is in terms of \$ per 144 ounces (i.e., \$ per 12-pack, since all drinks in our sample come in 12 oz containers).

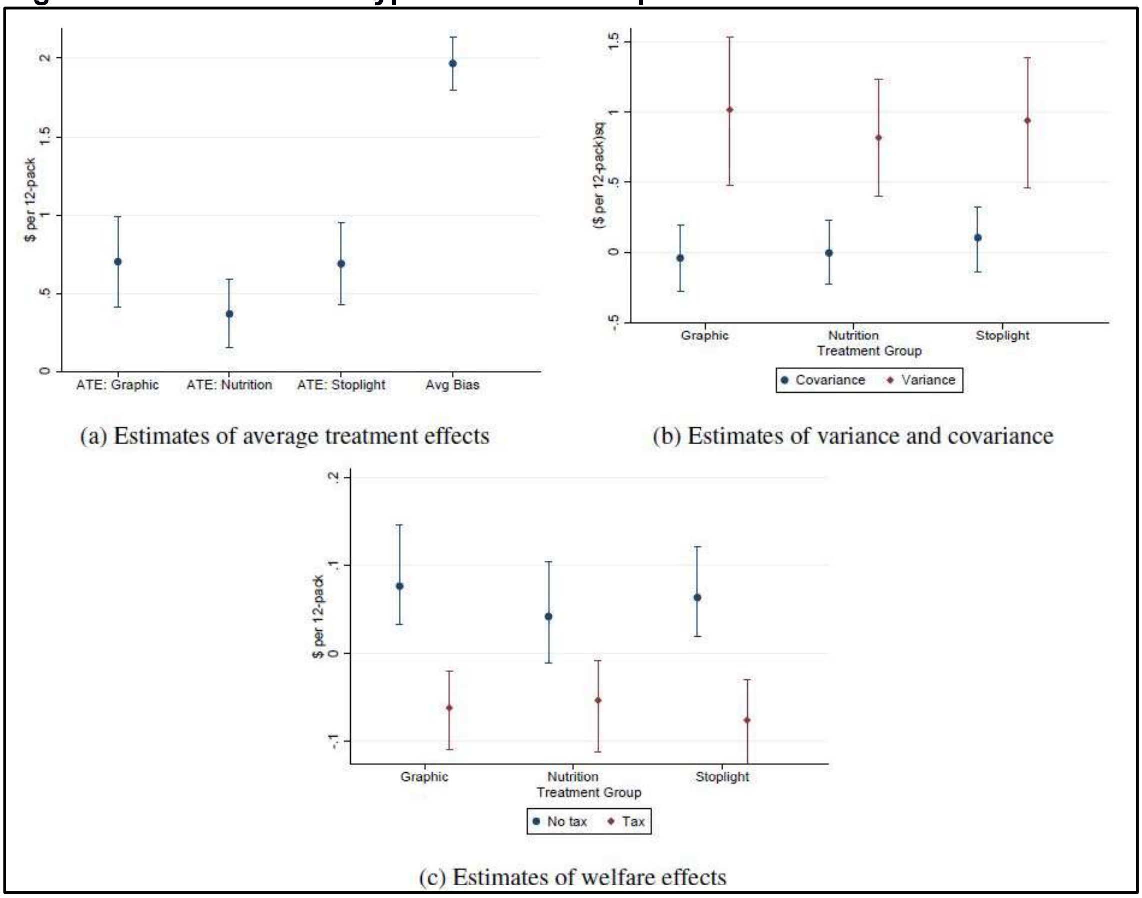
In our sample, $E[\widehat{\gamma_i}] = 1.95$, with a standard deviation of 1.26.

<u>Related Preliminary Work</u>

To ensure that our statistical methods our sufficiently well-powered, we have already piloted the soft drinks experiment with 270 Amazon Mechanical Turk (MTurk) participants making hypothetical choices. As shown in the top left panel of Figure A.8 below, we estimate that on average, consumers underweight the health costs of sugary drinks by about \$2 per 12-pack of 12-oz drinks. The labels have an average treatment effect (ATE) of decreasing WTP for the sugary drinks by about \$0.50. In an aggregate sense, these effects look like they are "in the right direction," partially counteracting consumer bias. Unfortunately, as the top right panel shows, while the treatment effects of these labels are heterogeneous (the variance is significant and positive), they are not targeted to the consumers making the biggest mistakes (the covariance is approximately zero).

When these statistics are combined with our theoretical formula (assuming a pass-through parameter of ρ = 1 for simplicity), the bottom panel shows that in the absence of any taxes and subsidies, the labels do increase welfare by about $0.50 per pack, as on aggregate they reduce over-consumption of the drinks. However, the bottom panel also shows that at the optimal sugary drinks tax, these labels in fact decrease social welfare. This is because the average decrease in consumption generated by the labels is no longer beneficial, and thus their highly heterogeneous but poorly targeted effects simply create "noise" in consumer choices. Our preliminary finding of differential benefits of the labels in the presence and absence of a sugary drinks tax illustrates how the combination of our theoretical formula and measurement techniques can be used to quantify the nuanced economic effects of these instruments.

Figure A.8: Results from hypothetical choice pilot



(a) Estimates of average treatment effects

(b) Estimates of variance and covariance

(c) Estimates of welfare effects

Note: Estimates are calculated from a hypothetical choice pilot of 270 participants on MTurk.

We plan to keep using hypothetical choice pilots to test new labels, and plan to have a set of five labels in the incentivized choice experiment.

One major caveat in interpreting the results of this preliminary work is that participants were not actually shopping for beverages, and thus may not have been properly incentivized when making decisions. In our NBER Roybal Center pilot, we plan to all participants receive a $7 shopping budget, randomly choose one row of one of the MPLs, and have participants receive a 12-pack of their stated choice at the stated price. For example, if in the randomly-chosen row of the randomly-chosen MPL the participant chose to receive a 12-pack of Pepsi at $4 over a 12-pack of LaCroix Cola at $4, then the participant would receive a 12-pack of Pepsi, shipped to their address, and keep the remainder of their shopping budget (here, $3).