Effects of "Fat Taxes" on Package Sizes, and
Welfare Distribution

Joseph V. Balagtas, Purdue University

Jose G. Nuno-Ledesma, Purdue University

Steven Y. Wu, Purdue University

Effects of "Fat Taxes" on Package Sizes, and Welfare Distribution

Joseph V. Balagtas, José G. Nuno-Ledesma & Steven Y. Wu [*]

June 2016

Abstract

We analytically study the pricing strategies a food retailer can execute under taxation when there is adverse selection. The questions are: how does the implementation of a tax regime affect the seller's ability to screen the market, and what effects do taxes have on package sizes and welfare distribution. We develop a parsimonious screening model in which the retailer offers a divisible good and does not know the buyers' willingness to pay. Under the more likely marketing strategies, we find that only consumers with high willingness to pay for the food see a reduction in their welfare; all package sizes are smaller, and the retailer sees her expected profit unambiguously diminished. These general results hold regardless of the type of implemented tax. Additionally, we briefly discuss how changes in *ad valorem* and specific taxes impact final prices.

JEL Codes: D04, D18, D82, H20, I18

PLEASE, DO NOT CITE THIS VERSION WITHOUT AUTHORS'
PERMISSION

1 Introduction

In this document, we aim to understand how a food retailer alters her menu of choices following the imposition of a tax regime. We also look at the effects that the introduction of taxation has on package sizes and distribution of surplus between consumers and the seller.

We leverage nonlinear pricing theory (second-degree price discrimination) to model a retailer's (the principal) marketing behavior and how she endogenously modifies the screening contracts offered to buyers (the agents) after a tax regime comes into effect. Nonlinear pricing is a sorting mechanism that stems from the mechanism design literature where the principal relies on the Revelation Principle (Myerson, 1979) to arrange a price schedule that does not depend on verifiable agent characteristics and, as a result, variations in prices across contracts are not explained solely by differences in marginal cost of production. Every agent is offered the same price schedule, from which they choose their consumption. The price-size combinations are designed in such a way that one group of buyers are left without consumer surplus, while other customers enjoy quantity discounts. By using a nonlinear pricing scheme, the principal is able to extract more surplus compared to other price-discriminating techniques.

Stylized observations suggest that retailers operating in a monopolistic competitive market sort their customers in a way consistent to what nonlinear pricing theory predicts. These market conditions are common in the food retailing industry. [1] In this context, the agents' hidden feature is their

---

[1]For instance, in Mexico, where a tax on sugar-sweetened beverages is in effect, the market share of the most popular soda brand is 70 % and the share of the second place brand is 15 % (Rivera-Dommarco et al., 2013)

willingness to pay for the food. Some customers (H-type buyers) enjoy consuming large quantities and are willing to pay more for the product, while others (L-type buyers) prefer smaller packages. The retailer offers a set of customized price-size combinations and the selection of one of these reveals the consumer type. The effects that a governmental intervention (in the form of a tax) could have on the set of offered price-size combinations and their resulting surplus distributions are unclear. This document analytically tackles this problem. We incorporate taxation into a parsimonious screening model and analyze direct and indirect effects of this regulation.

Although generalizable to settings where a monopolist applies second-degree price discrimination, our research is relevant to the public policy debate around proposed taxes and other regulations on certain foods and ingredients. The so-called "fat taxes" repeatedly appear in the public sphere as an alternative to regulate food consumption. Commonly, these taxes are suggested to be applied to edibles associated with weight gain, and/or containing ingredients judged to have deleterious effects on health. Examples of interventions of this nature in the United States include the penny-per-ounce tax on sugar-sweetened beverages (SSB) in Berkeley, passed in 2014 and in effect since 2015, and the proposed 1.5 cent-per-ounce tax on SSB in Philadelphia (Cohen, 2016). The rationale behind these policies is that it is possible to reduce the population's levels of obesity via a reduction of caloric intake when SSB become either more expensive or cannot be purchased in large containers. Critics argue that these measures hurt consumer welfare (Nestle, 2012) because they limit the space of options available to consumers. Setting aside the efficacy of these measures on reducing over-

2

weight rates and its health consequences among SSB consumers, the effects that these policies have on retailer's pricing schemes, overall welfare, and final consumption decisions is unclear. We address these issues.

Because as in 2012, 34.9 per cent of Americans over 20 years old were overweight (Ogden et al., 2014) and considering that the obesity epidemic results in over $ 148 billion dollar in deadweight losses to the United States (MacEwan et al., 2014), we expect campaigns promoting taxes on calorie-dense foods and beverages to arise in other cities in the nation, and the welfare-restriction argument to be brandished by their criticizers; hence the importance of developing research on the economic consequences of these policies. We consider important to mention that we are neutral towards these food policies. We do not intend to model the fine details of any "fat tax" policy, either already implemented or suggested, in the United States or abroad. We aim to provide a general explanation of the possible economic consequences that these interventions can produce.

The rest of this paper is organized as follows: in the next section we discuss relevant related prior research; in the third part we present our benchmark model; in the fourth section we incorporates taxation into the model and present the different strategies the retailer can follow; in the fifth part we discuss the major policy implications of the model (package size, distributional effects and tax incidence); lastly we give our concluding remarks.

2 Prior Research

A growing literature explores the impacts of "fat taxes" on consumption using observational data. Typically, these documents explore de association between these taxes and health and weight outcomes. However, most of these studies ignore the underlying market mechanisms that determine the amount of beverage at sale, the pricing strategies retailers use to appeal certain groups of consumers, and the final purchase decisions. In this sense, most of these documents are atheoretical, meaning that they do not account for the marketing strategies behind the design of size-price combinations encountered in the market.

Fletcher et al. (2010) examine the impact that taxes have had on soft drinks consumption and weight outcomes among children and adolescents in the United States. They conclude that taxation is associated with a moderate reduction in soft drinks consumption among adolescents and children, however, the authors do not find evidence supporting that these taxes have had a positive impact on weight outcomes, since young consumers have substituted soft drinks with other high calorie foods. This last finding coupled with the possibility of low tax revenue (due to the moderate reduction in consumption of taxed beverages) leads the authors to resolve that the health benefits and revenue generation of taxes on soft drinks may be be weaker than expected.

Sturm et al. (2010) also did not find a substantial effect of small taxes on obesity rates, but unlike Fletcher et al. (2010), Sturm et al also fail to discover a significant impact on soda consumption. The authors describe differentiated effects across groups of consumers. They indicate that low-

income families and other groups at high risk of becoming obese may be more sensitive to soda taxes. These authors suggest that per unit taxes could be preferable over *ad valorem* mainly because of easiness of implementation.

Miao et al. (2012) compare the outcomes of taxation on SSB versus taxing the sweetener ingredient of such drinks. Miao et al claim that a tax on the sweetener ingredient results in a much smaller consumer surplus loss, compared to a tax on the final product.

Colchero et al. (2016) find that the excise tax on SSB in Mexico is associated with a reduction in purchases of these products. Grogger (2015) also studies the Mexican "soda tax" and found that prices of SSB rose by more than the amount of the tax and consumers did not substitute SSB with other calorie dense drinks.

Bourquard and Wu (2016) is a document closer in spirit to our research project. In a setting where the retailer executes nonlinear pricing strategies, they model potential consequences of package size restrictions (think, for example about the sugary beverages portion cap rule proposed in New York City around 2013). They conclude that small to moderate package size restrictions do not affect consumer welfare, although producer welfare does get negatively impacted.

3 Model Setup without taxes in effect

In this section, we present a standard nonlinear pricing model in which a seller offers a divisible good, say soda, to one potential customer. We think about each transaction as a contracting situation where the retailer (the principal)

offers a menu of take-it-or-leave-it contracts to the buyer (the agent). Each contract is composed by two arguments: price and quantity. For simplicity, we assume that there exists only two types of consumers (H and L). The seller cannot observe the type of the buyer but she knows the distribution of types in the population. Therefore, the seller cannot engage in first-price discrimination. The customers differ in their willingness to pay for the good. H-type customers consume larger quantities of the good and have higher willingness to pay for them. L-types have a low willingness to pay for the good and therefore are less inclined to purchase large quantities of it.

We assume that the seller is risk neutral and thus she seeks to maximize her expected profit. The retailer offers a menu of size-price combinations from which the buyer decides his consumption. Let q represent size and $p(q)$ price. The seller's production cost function is $c(q)$, for simplicity we assume $c(q) = cq$ where $c'(q) = c > 0$ is a constant, so that the cost function is strictly increasing and quasi-convex. The profit obtained by the seller after selling a nonnegative quantity q to an i-type buyer is $p(q_i) - cq_i$.

There is a fraction $\beta \in [0, 1]$ of L-type customers and a proportion $(1 - \beta)$ of H-types buyers. The utility gained by an i-type customer after purchasing a nonnegative quantity q and paying a price p is $\theta_i v(q) - pq$. The parameter θ_i reflects the valuation that the i-type buyer has for the good. $\theta_H > \theta_L$, so that the Spence-Mirrlees single crossing condition is satisfied. We assume that $v(0) = 0$, $v'(q) > 0$ and $v''(q) < 0 \ \forall \ q \geq 0$, i.e. buyer's utility is strictly increasing and strictly concave. We also assume $\theta_H v'(0) > c$ and $\lim_{q \to \infty} \theta_H v'(q) < c$, so that at least H-type consumers have an incentive to engage in trade and that the retailer cannot offer an infinite large quantity

of the good.

The buyer knows $v(q)$ and his own type θ_i. The utility function of the i-type consumer is $U(\theta_i, q) = \theta_i v(q) - p$, where p denotes the price paid by the consumer. The seller does not know θ_i but she does know $v(q)$; thus the retailer faces uncertainty in one single dimension, namely the type of the potential buyer. The retailer is the first mover, sets the menu of screening contracts and she has full commitment power. We require voluntary participation at the interim level: buyers learn their types before the seller offers the menu of price-size options.

It is well known that, by the Revelation Principle, the seller's problem can be stated as a direct mechanism that is incentive-compatible and individually rational. Without loss of generality, the retailer restricts her price schemes $p(q)$ to the optimal choices made by each type of consumer. Let $p(q_i) = p_i$, for $i = L, H$; then the retailer's problem can be written as follows:

$$\underset{\langle (p_L, q_L), (p_H, q_H) \rangle}{\text{maximize}} \quad \mathbb{E}\left[\pi\right] = (\beta)[p_L - cq_L] + (1 - \beta)[p_H - cq_H]$$

subject to:

$$ICH : \theta_H v(q_H) - p_H \geq \theta_H v(q_L) - p_L$$

$$ICL : \theta_L v(q_L) - p_L \geq \theta_L v(q_H) - p_H \tag{1}$$

$$PCH : \theta_H v(q_H) - p_H \geq \bar{v}_H$$

$$PCL : \theta_L v(q_L) - p_L \geq \bar{v}_L$$

Where $\bar{v}_i$ is the i-type's reservation utility. Without loss of generality, we will assume $\bar{v}_i = 0$.

In the optimization problem (1), H-type consumers can always select the

bundle designed for L-types, thus the participation constraint for L-type buyers (PCL) is automatically satisfied. It is well known that, by the single-crossing condition, the incentive compatibility constraint for L-types (ICL) can be omitted from the program. Letting the two remaining constraints to hold with equality, the principal's relaxed problem is:

$$\underset{\langle (p_L,q_L),(p_H,q_H) \rangle}{\text{maximize}} \quad \mathbb{E}\left[\pi\right] = (\beta)[p_L - cq_L] + (1-\beta)[p_H - cq_H]$$

subject to:

$$ICH : \theta_H v(q_H) - p_H = \theta_H v(q_L) - p_L$$

$$PCL : \theta_L v(q_L) - p_L = 0$$

(2)

Using ICH and PCL from program (2), we obtain the following pricing rules:

$$p_L = \theta_L v(q_L) \tag{3}$$

$$p_H = \theta_H v(q_H) - (\theta_H - \theta_L)v(q_L) \tag{4}$$

Substituting (3) and (4) into the objective function we can express the problem as follows:

$$\max_{q_L, q_H} \quad \mathbb{E}\left[\pi\right] = \beta[\theta_L v(q_L) - cq_L] + (1-\beta)[\theta_H v(q_H) - (\theta_H - \theta_L)v(q_L) - cq_H] \tag{5}$$

The quantities in square brackets in (5) are "virtual surpluses" generated by transactions between the seller and an i-types consumer. When the re-

tailer and an L-type buyer engage in transaction, virtual surplus is equal to the actual surplus.[2] On the other hand, the virtual surplus generated by a transaction with an H-type consumer is smaller than the actual surplus due to the information rent he receives as an incentive for truthfully revelation of his type. The first order Kuhn-Tucker conditions of this problem are:

$$\text{FOC}[q_H] : \frac{\partial \mathbb{E}[\pi]}{\partial q_H} = (1 - \beta)[\theta_H v'(q_H) - c] \leq 0$$

where

$$q_H \geq 0 \quad \text{and} \quad \frac{\partial \mathbb{E}[\pi]}{\partial q_H} \cdot q_H = 0$$

(6)

$$\text{FOC}[q_L] : \frac{\partial \mathbb{E}[\pi]}{\partial q_L} = \beta[\theta_L v'(q_L) - c] + (1 - \beta)[-(\theta_H - \theta_L)v'(q_L)] \leq 0$$

where

$$q_L \geq 0 \quad \text{and} \quad \frac{\partial \mathbb{E}[\pi]}{\partial q_L} \cdot q_L = 0$$

(7)

In the coming subsection, we will discuss our base case: when the retailer sells to both types of consumers and there is no tax on the product she offers.

3.1 Pre-Tax Case I-A: The seller serves both types of consumers

This is our benchmark. The textbook results obtained from this first scenario will be compared to other relatively more complicated scenarios. In this

---

[2]The "actual surplus" generated by a transaction with a type i consumer equals the buyer's utility of consuming quantity q minus the retailer's cost of selling quantity q: actual surplus $= \theta_i v(q) - c(q)$, i.e. the consumer's utility minus the cost of production.

case, the retailer offers a menu of two price-size options with strictly positive quantities $q_H > 0$ and $q_L > 0$. She expects the consumers to reveal their true types via their purchase decisions. The two First Order Conditions FOC$[q_H]$ and FOC$[q_L]$ in (6) and (7), respectively, bind with strict equality. This implies the following about marginal utilities received by each type of customer:

$$\theta_H v'(q_H) = c \tag{8}$$

$$\theta_L v'(q_L) = c + \left(\frac{1-\beta}{\beta}\right)(\theta_H - \theta_L)v'(q_L) \tag{9}$$

It follows from (8) that there is no distortion at the top: the H-type consumer is supplied with his first-best quantity, where his marginal willingness to pay equals the marginal cost of production. On the other hand, an L-type consumer gets a quantity that is lower than first-best. For the seller, it is more costly to serve an L-type consumer since on top of marginal cost of production, the retailer incurs in an additional cost $(\frac{1-\beta}{\beta})(\theta_H - \theta_L)v'(q_L)$, associated with the information rent transferred to H-types.

H-types get their first-best quantities and virtual surplus is maximized during transactions between this type of customers and the retailer. L-types get a quantity lower than their first-best because the seller distorts the quantity supplied downwards and equates marginal cost to virtual marginal benefit to the consumer:

$$v'(q_L)\left[\theta_L - \left(\frac{1-\beta}{\beta}\right)(\theta_H - \theta_L)\right] = c \tag{10}$$

10

Let $q_L^{ia} > 0$ and $q_H^{ia} > 0$ be the quantities that solve (8) and (10). If it exists, the unique interior solution to program (2) is characterized by equations (11) and (12). These imply $q_L^{ia} < q_H^{ia}$. Therefore, the ICH restriction in the original problem (1) is satisfied.

$$\theta_H v'(q_H^{ia}) = c \tag{11}$$

$$\theta_L v'(q_L^{ia}) = \frac{c}{[1 - (\frac{1-\beta}{\beta})(\frac{\theta_H - \theta_L}{\theta_L})]} > c \tag{12}$$

Final pricing rules are $p_H^{ia} = \theta_H v(q_H^{ia}) - (\theta_H - \theta_L)v(q_L^{ia})$ and $p_L^{ia} = \theta_L v(q_L^{ia})$. The retailer's expected profit is:

$$\pi^{ia} = (\beta)[\theta_L v(q_L^{ia}) - cq_L^{ia}] + (1 - \beta)[\theta_H v(q_H^{ia}) - (\theta_H - \theta_L)v(q_L^{ia}) - cq_H^{ia}] \tag{13}$$

The value functions of the consumers are:

$$U_L^{ia} = 0$$
$$U_H^{ia} = (\theta_H - \theta_L)v(q_L^{ia}) \tag{14}$$

All the discussion above and the mentioned results are summarized in the following proposition:

Proposition 1. *Suppose that the retailer decides to screen the market by offering a menu of two price-size combinations. Assume the sale of the product is untaxed. Then:*

1. *$\theta_H v'(q_H^{ia}) = c$ So that H-types are offered a package which size equals*

their first-best quantity.

2. $\theta_L v'(q_L^{ia}) = \frac{c}{\left[1-\left(\frac{1-\beta}{\beta}\right)\left(\frac{\theta_H-\theta_L}{\theta_L}\right)\right]} > c$ *So that L-types are offered a package which size is less than their first-best quantity.*

3. $p_H^{ia} = \theta_H v(q_H^{ia}) - (\theta_H - \theta_L)v(q_L^{ia})$ *So that the price of the H-type package is discounted by an information rent.*

4. $p_L^{ia} = \theta_L v(q_L^{ia})$

5. *The seller's value optimized profit is expressed in equation (13).*

6. *The H-type consumer's value function* $U_H^{ia} = (\theta_H - \theta_L)v(q_L^{ia})$.

7. *The L-type consumer's value function is* $U_L^{ia} = 0$.

4 Incorporating Taxation into the Model

In this section, we analyze the potential effects that the introduction of *ad valorem* and specific taxes can have on the retailers' screening strategy. In order to pursue this sections' goal, we modify the original retailer's problem and incorporate tax regimes.

Definition 1. *A **tax regime** (t_s, t_v) is one member of the set of ordered pairs* $T \equiv \{\{[0,1] \times [0,1)\} \setminus (0,0)\}$, *i.e.*
$(t_s, t_v) \in T \equiv \{\{(t_s, t_v) \mid t_s \in [0,1], \ t_v \in [0,1)\} \setminus (0,0)\}$
where t_s denotes a specific (or per-unit) tax and t_v an ad valorem tax.

Note that we exclude combinations of taxes where $t_v = 1$; this is in order to avoid divisions by zero later on. We do not include the singleton

$(t_s, t_v) = (0, 0)$ since it represents the event of no taxation discussed in the previous section. When a tax regime (t_s, t_v) is in effect, the expected profit maximizer retailer faces the following constrained optimization problem:

$$\underset{\langle (p_L, q_L), (p_H, q_H) \rangle}{\text{maximize}} \mathbb{E}[\pi] = (\beta)[(1 - t_v)p_L - t_s q_L - c q_L] + (1 - \beta)[(1 - t_v)p_H - t_s q_H - c q_H]$$

subject to:

$$ICH : \theta_H v(q_H) - p_H \geq \theta_H v(q_L) - p_L$$

$$PCL : \theta_L v(q_L) - p_L \geq \bar{v}_L$$

$$(15)$$

Where, without loss of generality we set $\bar{v}_L = 0$. Naturally, we can rewrite (15) as an unrestricted maximization problem:

$$\underset{q_L, q_H}{\max} \mathbb{E}[\pi] = (1 - t_v)\Big\{ (\beta)[\theta_L v(q_L) - \Psi_L] + (1 - \beta)[\theta_H v(q_H) - (\theta_H - \theta_L)v(q_L) - \Psi_H] \Big\}$$

$$(16)$$

where $\Psi_i \equiv (t_s q_i + c q_i) \div (1 - t_v)$ is the seller's effective cost function. Similarly, let $\psi \equiv \frac{d\Psi_i}{dq_i} = (t_s + c) \div (1 - t_v)$ denote the retailer's effective marginal cost.

In the objective function (16), the expressions within square brackets are virtual surpluses generated by transactions between the seller and an i-type consumer when a tax regime (t_s, t_v) has been implemented. Ad valorem and specific taxes are modeled as modifications to the principal's cost function. Because profit decreases with cost, these virtual surpluses are smaller than those in problem (5). This claim is easy to verify (all proofs can be found in

13

the appendix).

Claim 1. *When a tax regime (t_s, t_v) is in effect and the seller's objective function is (16), the virtual surplus generated by a transaction between a retailer and a buyer (either H or L type) is strictly smaller than the corresponding virtual surplus when there is not a tax regime in effect and the retailer's objective function is (5).*

The first order Kuhn-Tucker conditions of problem (16) are:

$$\text{FOC}[q_H] : \frac{\partial \mathbb{E}[\pi]}{\partial q_H} = (1 - t_v)(1 - \beta)[\theta_H v'(q_H) - \psi] \leq 0$$

where (17)

$$q_H \geq 0 \quad \text{and} \quad \frac{\partial \mathbb{E}[\pi]}{\partial q_H} \cdot q_H = 0$$

$$\text{FOC}[q_L] : \frac{\partial \mathbb{E}[\pi]}{\partial q_L} = (1 - t_v)\{\beta(\theta_L v'(q_L) - \psi) + (1 - \beta)[-(\theta_H - \theta_L)v'(q_L)]\} \leq 0$$

where

$$q_L \geq 0 \quad \text{and} \quad \frac{\partial \mathbb{E}[\pi]}{\partial q_L} \cdot q_L = 0$$

(18)

Now, we analyze three possible strategies that the seller could follow under taxation.

4.1 Taxed Case II-A: The seller Screens the Market and Serves Both Types of Consumers

The seller offers a menu of two contracts in which both quantities q_H and q_L are strictly positive. Both FOC[q_H] and FOC[q_L] (in 17 and 18 respectively)

14

hold with strict equality, implying:

$$\theta_H v'(q_H) = \psi \tag{19}$$

$$\theta_L v'(q_L) = \psi + \left(\frac{1-\beta}{\beta}\right)(\theta_H - \theta_L)v'(q_L) \tag{20}$$

The retailer designs a menu of contracts in such a way that the effective marginal cost of producing a container intended for an H-type buyer equals this consumer's marginal benefit. The marginal benefit received by an L-type buyer when he purchases his targeted bundle does not equal the effective production marginal unit cost. The seller distorts the size of the package downwards and equates the effective marginal cost of producing the small package to the virtual marginal benefit that the L-type consumer obtains after the purchase:

$$v'(q_L)\left[\theta_L - \left(\frac{1-\beta}{\beta}\right)(\theta_H - \theta_L)\right] = \psi \tag{21}$$

Let $q_L^{iia} > 0$ and $q_H^{iia} > 0$ be the quantities that solve (19) and (20). If it exists, the unique interior solution to program (15) is characterized by the following equations, which imply $q_H^{iia} > q_L^{iia}$.

$$\theta_H v'(q_H^{iia}) = \psi \tag{22}$$

$$\theta_L v'(q_L^{iia}) = \frac{\psi}{[1 - (\frac{1-\beta}{\beta})(\frac{\theta_H - \theta_L}{\theta_L})]} > \psi \tag{23}$$

Since $\psi > c \,\forall (t_s, t_v) \in T$, both types of consumers get less than their first-

15

best optimal quantities. The retailer's expected profit is smaller, compared to the untaxed case because the virtual surplus available to trade is smaller.

Equation (22) suggests that under a tax regime, there is distortion at the top (because $\psi > c$). L-types also receive a smaller quantity compared to the untaxed case I-A, which is already smaller than their first best. The decrease in size of the package tailored for L-type buyers arises from two sources: i) since $\psi > c$, the seller offers smaller quantities across both screening contracts $[p_i, q_i]$, this drives up the cost of serving all consumers, and ii) the marginal cost of serving L-types is also driven up by information rents transferred to H-type clients.

From the participation and incentive compatibility constraints in program (15), we can derive the optimal pricing rules $p_L^{iia} = \theta_L v(q_L^{iia})$ and $p_H^{iia} = \theta_H v(q_H^{iia}) - (\theta_H - \theta_L)v(q_L^{iia})$. Because under a tax regime (t_s, t_v) the package offered to each type of client is smaller, the prices per package are lower, compared to the situation without tax regulation. This is $q_L^{iia} < q_L^{ia}$ and $q_H^{iia} < q_H^{ia}$. However, the price per unit of product are higher when there is a tax regime in effect, i.e. $\frac{p_L^{iia}}{q_L^{iia}} > \frac{p_L^{ia}}{q_L^{ia}}$ and $\frac{p_H^{iia}}{q_H^{iia}} > \frac{p_H^{ia}}{q_H^{ia}}$. Regarding participants' value functions, the seller's expected profit is in equation (24) and consumers' surpluses are in (25). We list these outcomes in proposition (2).

$$\pi^{iia} = (1-t_v)\Big\{(\beta)[\theta_L v(q_L^{iia}) - \psi q_L^{iia}] + (1-\beta)\{[\theta_H v(q_H^{iia}) - (\theta_H - \theta_L)v(q_L^{iia})] - \psi q_H^{iia}\}\Big\} \tag{24}$$

$$U_L^{iia} = 0$$

$$U_H^{iia} = (\theta_H - \theta_L)v(q_L^{iia}) \tag{25}$$

Proposition 2. *Suppose that the retailer decides to screen the market and to serve both types of clients offering two tailored price-size combinations. Assume there is a tax regime (t_s, t_v) in effect with at least one type of tax strictly positive. Then:*

1. *$\theta_H v'(q_H^{iia}) = \psi$. There is distortion at the top.*

2. *$\theta_L v'(q_L) = \psi + (\frac{1-\beta}{\beta})(\theta_H - \theta_L)v'(q_L^{iia})$. L-type consumers get a quantity lower than their first best.*

3. *$q_L^{iia} < q_L^{ia}$ and $q_H^{iia} < q_H^{ia}$. Both packages are smaller compared to the pre-tax case I-A.*

4. *$p_H^{iia} = \theta_H v(q_H^{iia}) - (\theta_H - \theta_L)v(q_L^{iia})$. The price that H-type buyers pay is discounted by the information rent.*

5. *$p_L^{iia} = \theta_L v(q_L^{iia})$. The price paid by L-types is not discounted.*

6. *$p_L^{iia} < p_L^{ia}$ and $p_H^{iia} < p_H^{ia}$. Prices of both packages are lower compared to the pre-tax case I-A.*

7. *The retailer receives profit expressed by value function (24).*

8. *H-type client value function is $U_H^{iia} = (\theta_H - \theta_L)v(q_L^{iia})$.*

9. *L-type buyer value function is $U_L = 0$.*

17

4.2 Taxed Case II-B: The Seller Serves only H-type consumers

Here, we study to a setting where $q_H > 0$ and $q_L = 0$. Let the superscript *iib* denote variables that maximize the retailer's benefit under this strategy. FOC[q_L] in (18) does not bind with equality. Utilizing FOC[q_H] in (17), the pricing rule (4), and the normalizing assumption $v(0) = 0$, we obtain:

$$\theta_H v'(q_H^{iib}) = \psi \tag{26}$$

$$p_H^{iib} = \theta_H v(q_H^{iib}) \tag{27}$$

The retailer decides not to serve L-types and concentrates her marketing efforts on attending H-types exclusively. As a consequence, only one size package is offered and is intended to appeal H-type clients only. The price of the container in this situation is higher than the derived in the taxed case II-A, since the principal no longer needs to elicit truthful revelation of information.

Notice that the size of the package designed to appeal L-types (q_L) does not appear in the first expression in $\theta_H v'(q_H) = \psi$, this suggests that the amount of product offered to H-types under this marketing strategy is the same as the quantity offered in case II-A, where the retailer continues to offer two differentiated packages. Since, compared to case II-A, the price of the container offered to H-types is larger while its size remains the same, the per unit price is higher. The H-type value function is zero, as well as the L-types' because they do not get to purchase any product. Expected profit

is in equation (28). We list these outcomes in proposition (3).

$$\pi^{iib} = (1 - t_v)(1 - \beta)[\theta_H v(q_H^{iib}) - \psi q_H^{iib}] \qquad (28)$$

Proposition 3. *Suppose that the retailer decides not to screen the market and offers one size-price option designed to serve H-type buyers solely. Assume there is a tax regime (t_s, t_v) in effect with at least one type of tax strictly positive. Then:*

1. *$\theta_H v'(q_H^{iib}) = \psi > c$ There is distortion at the top.*

2. *$q^{iib} = q^{iia} < q_H^{ia}$ so that the size of the H-type package offered in case II-B is the same with the offered in case II-A and smaller than the sizes offered in cases I-A without taxes.*

3. *$p_H^{iib} = \theta_H v(q_H^{iib})$ so that information rents are not being transferred to the consumers.*

4. *$p_H^{iia} < p_H^{iib}$ so that the price of the H-type package is higher in case II-B compared to case II-A.*

5. *Per unit price is higher compared to per unit price in case II-A.*

6. *The seller's value functions is expressed by equation (28).*

7. *The H-type consumer's value function is $U_H = 0$.*

4.3   Taxed Case II-C: One-Size-Fits-All

Here we study a strategy where the retailer, ignoring potential willingness to pay discrepancies across clients, pools the market and stops customizing

bundles. This implies that only one size package $q_L > 0$ is offered by the seller and it is intended to cover all of the demand. The retailer does not need to motivate truthful revelation of private information. She only needs to assure participation of L-type clients and this automatically guarantees that H-types would also participate in trades. Her optimization problem can be written as follows:

$$\max_{p,q} \ \mathbb{E}\left[\pi\right] = p_L - \psi q_L$$

$$\text{subject to:} \tag{29}$$

$$PCL : \theta_L v(q) - p_L \geq \bar{v}_L$$

Without loss of generality we set $\bar{v}_L = 0$. Since the Spence-Mirrlees single crossing condition is satisfied, if the L-type consumers engage in transactions, the H-types will also participate. Profit maximization in this case implies:

$$\theta_L v'(q_L^{iic}) = \psi \tag{30}$$

$$p_L^{iic} = \theta_L v(q_L^{iic}) \tag{31}$$

Thus, the L-type consumers do not get their first best quantity. Let p_L^{iic} and q_L^{iic} be the optimal price and quantities. The retailer's expected benefits are in (32) and the clients' value functions in (33). We summarize the aforementioned results in proposition (4).

$$\pi^{iic} = (1 - t_v)[\theta_L v(q_L^{iic}) - \psi \cdot q_L^{iic}] \tag{32}$$

20

$$U_L^{iic} = 0$$
$$U_H^{iic} = (\theta_H - \theta_L)v(q_L^{iic}) \tag{33}$$

Proposition 4. *Suppose that the retailer decides not to screen the market and offers one size-price contract designed to serve both types of buyers. Assume there is a tax regime* (t_s, t_v) *in effect with at least one class of tax strictly positive. Then:*

1. $\theta_L v'(q_L^{iic}) = \psi$ *so that L-types are provided with a quantity smaller than their first best.*

2. *The price of the only size of package offered is* $p_L^{iic} = \theta_L v(q_L^{iic})$.

3. *The seller's value function is expressed by equation (32).*

4. *The L-type consumer value function is zero.*

5. *The H-type consumer value function is* $U_H^{iic} = (\theta_H - \theta_L)v(q_L^{iic})$.

4.4 How does taxation affect retailers' choice of scheme?

Throughout this paper, we assume that in absence of taxation the seller screens the market and offers a menu of two price-size combinations (as in our benchmark case I-A). We also assume that when a tax regime comes into play, she would prefer to keep offering differentiated price-size combos (as in taxed case II-A). In this section, we derive the conditions that would provoke the principal to shift her marketing scheme from a separating approach (II-A) to either strategy II-B (Serve H-type only) or scheme II-C (one-size-fits-all).

We first analyze the case where the seller moves from screening the market to start serving H-type buyers exclusively. In order for this shift to occur, the expected profit in case II-B should be larger than in case II-A. We make the following tie-breaking assumption: if $\pi^{iia} = \pi^{iib}$, then the retailer decides to keep offering two differentiated screening contracts.

Proposition 5. *Suppose that a tax regime $(t_s, t_v) \in T$, comes into effect. Then, the seller will stop offering two price-size consumption options and will exclusively target H-type consumers iff the following inequality holds true:*

$$[\theta_L v(q_L^{iia}) - \psi q_L^{iia}] < \left(\frac{1 - \beta}{\beta}\right)(\theta_H - \theta_L)v(q_L^{iia}) \tag{34}$$

Inequality (34) can be interpreted as follows: the retailer will switch her strategy from serving both types of consumers to target H-types solely, if and only if the virtual surplus generated by transactions with L-types is smaller than the total burden of information rent transfers.

Now we analyze the case where the retailer changes her scheme from screening to pooling (II-C) the market. We first derive a necessary condition under which the implementation of a tax regime would provoke the retailer to move from a screening strategy to a pricing scheme serving all consumers with a single price-size combination. In order for this to occur, the profit earned by the seller in case II-C should be larger than her earnings in case II-A. We make the tie-breaking assumption that if $\pi^{iia} = \pi^{iic}$, then the retailer decides to keep offering two differentiated screening contracts.

Proposition 6. *Suppose that a tax regime $(t_s, t_v) \in T$, comes into effect. Then, the seller will stop offering two price-size combinations and will start*

serving both types of consumers with a one-size-fits-all strategy if and only if the following inequality holds true:

$$(\beta)[\theta_L v(q_H^{iia}) - \psi q_L^{iia}] + (1-\beta)[\theta_H v(q_H^{iia}) - (\theta_H - \theta_L)v(q_L^{iia}) - \psi q_H^{iia}] < \theta_L v(q_L^{iic}) - \psi q_L^{iic}$$

$$(35)$$

Proposition (6) indicates that the retailer stops differentiating bundles and adopts a on-size-fits-all marketing approach if and only if the convex combination of the virtual surpluses obtained form the screening scheme is strictly smaller than the total virtual surplus gained under the pooling strategy.

Next, we present a sufficient condition for the retailer to switch her strategy to the scheme delineated in case II-C. If this condition holds, it is in the retailers' best interest to change her pricing tactic.

Proposition 7. *Suppose that a tax regime (t_s, t_v), with at least one class of tax strictly positive, comes into effect. Let assumption 1 to hold true:*

Assumption 1. *The virtual surplus of serving an L-type is larger than or equal to the virtual surplus of trading with an H-type in case II-A:*

$$[\theta_L v(q_L^{iia}) - \psi q_L^{iia}] \geq [\theta_H v(q_H^{iia}) - (\theta_H - \theta_L)v(q_L^{iia}) - \psi q_H^{iia}]$$

Then, the seller will stop offering two price-size combinations and will start serving both types of consumers consumer only if the following inequality holds true:

$$\theta_L v(q_H^{iia}) - \psi q_H^{iia} < \theta_L v(q_L^{iic}) - \psi q_L^{iic} \tag{36}$$

which indicates that the virtual surplus obtained when serving an L-type in scenario II-A is strictly smaller than the surplus gained when serving an L-type in case II-C.

The intuition behind this result is the following: Serving H-type buyers is costly because the retailer needs to provide incentives so that they truthfully reveal their information. As the proportion of H-types increases (β decreases), the retailer's benefits decrease and her costs (in form of information rent transfers) increases. If assumption 1 does not hold and if we do not add more stringent parametric assumptions, it is not clear when a tax regime causes the seller to stop segmenting the market.

In sum, the event of the retailer switching marketing tactic towards a one-size-fits-all approach is unlikely because two stringent conditions need to hold: i) there must be a very low proportion of L-type purchasers in the population of buyers and ii) serving L-types should be more profitable than attending H-types, in other words, information rents should be a prohibitively high burden for the retailer.

5 Policy Implications: Taxation Influence on Package Size and Surplus Distribution

5.1 Effects on Quantities

The primal goal of taxing certain foods, especially calorie-dense and sugary products, is to reduce consumption. Here we compare the size of the packages offered in the pre-tax scenario I-A to those designed in the taxed cases II-A, II-B and II-C.

Regardless of the pricing strategy adopted by the retailer, after a tax regime is implemented, the size of the package designed to appeal H-type customers is smaller when compared to case I-A.

Proposition 8. *Suppose that a tax regime $(t_s, t_v) \in T$ is implemented. Then:*

- $q_H^{ia} > q_H^{iia}$

- $q_H^{iib} = q_H^{iia} \implies q_H^{ia} > q_H^{iib}$

Under taxation, the size of the package designed for L-type buyers in taxed case II-A is smaller when compared to pre-tax scenario I-A. The amount offered to L-type consumers in taxed pooling case II-C is lower compared to taxed screening case I-A if a certain condition holds true.

Proposition 9. *Suppose that a tax regime $(t_s, t_v) \in T$, comes into effect. Then:*

- if $\psi > \dfrac{c}{\left[1 - \left(\frac{1-\beta}{\beta}\right)\left(\frac{\theta_H - \theta_L}{\theta_L}\right)\right]}$

1. $q_L^{ia} > q_L^{iic} > q_L^{iia}$

- if $\psi < \frac{c}{\left[1-\left(\frac{1-\beta}{\beta}\right)\left(\frac{\theta_H-\theta_L}{\theta_L}\right)\right]}$

  1. $q_L^{iic} > q_L^{ia} > q_L^{iia}$

- if $\psi = \frac{c}{\left[1-\left(\frac{1-\beta}{\beta}\right)\left(\frac{\theta_H-\theta_L}{\theta_L}\right)\right]}$

  1. $q_L^{ia} = q_L^{iic} > q_L^{iia}$

In the pooling equilibrium strategy II-C when the retailer implements a one-size-fits-all strategy, only if the tax is high enough, the size of the package offered in case II-C is smaller than the size of the package tailored for L-types in screening cases I-A (pre-tax) and II-A (taxed).

5.2 Effect of Taxation on Expected Profit and Consumers' Welfare

Let us first analyze the retailers' expected profit. As we already mentioned in claim (1), compared to the pre-tax case, the virtual surplus available to trade between buyer and seller is smaller, thus expected profit in the screening-with-tax case (II-A) is smaller than the expected benefit without tax in case I-A.

If the retailer chooses to execute a non-screening strategy, such as those described in taxed cases II-B and II-C, then she stops gaining extraordinary rents and her expected profit drops to zero. This is an unsurprising consequence, since the root of the rents captured by a monopolist is price discrimination.

Proposition 10. *Suppose that a tax regime $(t_s, t_v) \in T$. Regardless of the pricing strategy implemented by the retailer, her expected profit is smaller compared to the pre-tax case I-A. Moreover, if the retailer stops screening the market, her expected profit collapses to zero.*

In the context of a government mandated tax on junk food, proposition (10) suggests that food retailers would result negatively affected, unambiguously, as a result of this measure.

Now, we concentrate on assessing the distributional consequences of the tax regime. According to our model, L-type purchasers are hold at their reservation value (which we normalized to zero) across all of the marketing strategies we analyzed. Thus, L-type buyers do not see their surplus affected by the imposition of a tax.

On the other hand, H-type consumers' surplus is likely to decrease after the intervention. However, the surplus they get to obtain could remain unaffected or even increase (if the retailer applies a one-size-fits-all strategy). The implementation of a pricing scheme resembling case II-C is unlikely, as we already mentioned; however, if the seller adopts this marketing approach, and if the tax regime is sufficiently mild, then H-type consumers' welfare actually increases. In order for this to occur, the effective marginal cost after taxation should be smaller than the marginal benefit received by L-types in the pre-tax scenario, i.e. if $\psi < \frac{c}{\left[1 - \left(\frac{1-\beta}{\beta}\right)\left(\frac{\theta_H - \theta_L}{\theta_L}\right)\right]}$. Proposition 11 summarizes these results.

Proposition 11. *Suppose that a tax regime $(t_s, t_v) \in T$, comes into effect. Then, compared to pre-tax screening case I-A:*

- *L-type consumers' welfare remains unaffected, regardless of the pricing*

strategy implemented by the seller.

- *If $\psi > \frac{c}{\left[1-\left(\frac{1-\beta}{\beta}\right)\left(\frac{\theta_H-\theta_L}{\theta_L}\right)\right]}$.*

    1. *H-type consumers' welfare decreases, regardless of the pricing strategy implemented by the seller.*

- *If $\psi < \frac{c}{\left[1-\left(\frac{1-\beta}{\beta}\right)\left(\frac{\theta_H-\theta_L}{\theta_L}\right)\right]}$.*

    1. *In cases II-A and II-B, H-type consumers' decreases.*

    2. *In case II-C, H-type consumers' welfare increases.*

- *If $\psi = \frac{c}{\left[1-\left(\frac{1-\beta}{\beta}\right)\left(\frac{\theta_H-\theta_L}{\theta_L}\right)\right]}$.*

    1. *In cases II-A and II-B, H-type consumers' decreases.*

    2. *In case II-C, H-type consumers' remains unaffected.*

In a food retailing environment, this would imply that one group of consumer do get negatively impacted by a tax; namely, those who prefer to purchase large quantities of the product in question. The other buyers remain unaffected, because even without a tax they were already giving the maximum amount they would be willing to pay for the food.

6 *Ad Valorem* and Per Unit Tax Incidence

Thus far, we have not differentiated the effects of an *ad valorum* versus a per unit tax. This is because the results derived above hold true whether the government applies an *ad valorum*, a per unit or a combination of both types of taxes. In this section, we present the incidences that each class of tax has

on prices, according to our model. The expressions below can be interpreted as the predictions our model makes regarding changes in prices as response to in increment in taxes.

We begin by analyzing the incidence that each tax has on prices when the retailer screens the market (case II-A). Integrating out the equations (19) and (21), and using the expressions for prices in (3) and (4), we can rewrite the prices of each package in terms of quantities and the model's parameters:

$$p_L = q_L \left[\frac{\psi}{1 - \left(\frac{1-\beta}{\beta}\right)\left(\frac{\theta_H - \theta_L}{\theta_L}\right)} \right] \tag{37}$$

$$p_H = \psi \left[q_H - \frac{(\theta_H - \theta_L)q_L}{\theta_L - \left(\frac{1-\beta}{\beta}\right)(\theta_H - \theta_L)} \right] \tag{38}$$

We encounter familiar outcomes. The price of the small package equals the marginal utility of L-types at consumption level q_L. The price of the large package equals the marginal benefit of H-types at level of consumption q_H minus a discount factor that depends on q_L and adjusted by the marginal elective cost.

Proposition 12. *Suppose the retailer is implementing the strategy delineated in case II-A. The effects that each tax has on package prices are:*

- $\frac{dp_L}{dt_s} = \left(\frac{1}{1-t_v}\right)\left[\frac{q_L}{1 - \left(\frac{1-\beta}{\beta}\right)\left(\frac{\theta_H - \theta_L}{\theta_L}\right)}\right]$

- $\frac{dp_L}{dt_v} = \left(\frac{t_s + c}{1 - t_v}\right) \cdot \frac{dp_L}{dt_s}$

- $\frac{dp_H}{dt_s} = \left(\frac{1}{1-t_v}\right)\left[q_H - q_L \cdot \frac{(\theta_H - \theta_L)}{\theta_L - \left(\frac{1-\beta}{\beta}\right)(\theta_H - \theta_L)}\right]$

29

- $\frac{\mathrm{d}p_H}{\mathrm{d}t_v} = \left(\frac{t_s+c}{1-t_v}\right) \cdot \frac{\mathrm{d}p_H}{\mathrm{d}t_s}$

Now we turn to the scenario where the seller implements the pricing strategy depicted in case II-B, i.e. she decides to serve exclusively H-type buyers. Integrating out equation (26) and using the expression for price in (27), we can express the price of the only package in the following form:

$$p_H = \left(\frac{t_s + c}{1 - t_v}\right) q_H \tag{39}$$

Proposition 13. *Suppose the retailer is implementing the strategy delineated in case II-B. The effects that each tax has on the price are:*

- $\frac{\mathrm{d}p_H}{\mathrm{d}t_s} = \frac{q_H}{1-t_v}$

- $\frac{\mathrm{d}p_H}{\mathrm{d}t_v} = \left(\frac{t_s+c}{1-t_v}\right)\frac{\mathrm{d}p_H}{\mathrm{d}t_s}$

Lastly we present the scenario where the seller implements the pricing strategy depicted in case II-C, where she decides to pool the market and serves both type of buyers with a single undifferentiated price-size combo. Integrating out equation equation (30) and using the expression for price in (31), we can express the price of the package in terms of quantities and parameters:

$$p_L = \left(\frac{t_s + c}{1 - t_v}\right) q_L \tag{40}$$

Proposition 14. *Suppose the retailer is implementing the strategy delineated in case II-C. The effects that each tax has on the price are:*

- $\frac{\mathrm{d}p_L}{\mathrm{d}t_s} = \frac{q_L}{1-t_v}$

30

- $\frac{\mathrm{d}p_L}{\mathrm{d}t_v} = \left(\frac{t_s+c}{1-t_v}\right)\frac{\mathrm{d}p_L}{\mathrm{d}t_s}$

The results above make clear that the price of all packages increase following an increment in either t_s or t_v. Whether the tax increment is overshifted or undershifted will depend on the parametrization we choose for consumers' demand. The general results we have presented here are aligned with what the Public Finance literature predicts regarding tax incidence under nonlinear pricing[3].

7   Conclusion

In this paper we present an analysis of the economic consequences of taxation on products offered in the food retailing industry. We also study how these policies are likely to influence marketing decisions. We do not aim to advocate for or against any food regulation. We do not intend to design a model that mimics all of the details of any "fat tax" in the United States or abroad, either implemented or proposed. We offer a general study of the more likely consequences that these interventions would have on package size and welfare distribution.

The key findings in this paper are that after the introduction of a tax regime and comparing to an untaxed counterfactual state of the world: i) the size of the package offered to H-type decreases, except for an unlikely case where the retailer aims to cover her entire demand using a one-size-fits-all scheme and the tax is mild enough; ii) the L-type consumers' welfare remains unaffected regardless of the marketing strategy adopted by the retailer; iii)

---

[3]See, for example Jensen and Schjelderup (2011)

H-type consumers' welfare decreases under separating and H-exclusive marketing strategies, however, if the retailer stops screening and offers a small package to all of her customers, the effect on H-type buyers's welfare is ambiguous and depends on the severity of the regulation, and iv) The retailer sees her expected profit unambiguously reduced.

Not every consumer suffers welfare loss under a restricted nonlinear pricing scheme, when the constraint takes the form of a tax; only those customers with high willingness to pay for the food see their welfare diminished. Consumers with lower appetite for the product always pay the maximum they are willing to give for the good, regardless of the strategy employed by the seller. Since per-unit prices rise, the seller adjusts downwards the size of the small package so she does not loose "low-type" clients and continues extracting surplus from them.

On the other hand, in the more plausible settings, customers that enjoy large quantities of the product see their welfare diminished. In the case where the retailer screens the market in spite of the tax, the decline in welfare of "high-type" clients comes from the size reduction of the *small* container, this is because this quantity directly influences the magnitude of the quantity discounts that these customers enjoy (in other words, the amount of information rents transferred to motivate revelation of information). If the retailer decides to serve "high-type" customers solely, then these clients are hold to their reservation value, i.e. they pay the maximum amount they are willing to give for the product (this resembles the marketing strategies adopted by "exclusive" or "high end" brands in other industries). When the seller, aiming to cover all of her demand, offers a small package to all customers,

"high-type" buyers lose welfare because the size of this package is smaller compared to the small package offered when the retailer screens the market, this is akin to a lower quantity discount.

The retailer's expected profit decreases under taxation compared to an untaxed setting. If she decides to keep offering tailored packages in order to appeal both type of consumers, her expected benefit is lower because the introduction of a tax regime is akin to an increment in the marginal cost of production. If she chooses not to screen the market, the reduction in expected profit has an additional cause, she stops receiving the extraordinary rents associated with second-degree price discrimination.

Future research projects should include the effects of changes in the model's parameters. Examples could include modification in taste hetero-geneity (i.e. the difference in willigness to pay between high and low types) via advertising campaigns or others, and changes in the proportion of con-sumers with high willingness to pay for these foods.

References

Brian Bourquard and Steve Y Wu. An Economic Analysis of Beverage Size Restrictions. *Purdue University, mimeo*, 2016.

Luc Cohen. Philadelphia mayor rewrites playbook for soda-tax advocates, June 2016. URL `www.reuters.com/article/us-beverages-philadelphia-sodatax-mayor-idUSKCN0Z20BI`. [Online; posted 16-June-2016].

M Arantxa Colchero, Barry M Popkin, Juan A Rivera, and Shu Wen Ng. Beverage purchases from stores in mexico under the excise tax on sugar sweetened beverages: observational study. *bmj*, 352:h6704, 2016.

Jason M Fletcher, David E Frisvold, and Nathan Tefft. The effects of soft drink taxes on child and adolescent consumption and weight outcomes. *Journal of Public Economics*, 94(11):967–974, 2010.

Jeffrey Grogger. Soda taxes and the prices of sodas and other drinks: evidence from mexico. Technical report, National Bureau of Economic Research, 2015.

Sissel Jensen and Guttorm Schjelderup. Indirect taxation and tax incidence under nonlinear pricing. *International Tax and Public Finance*, 18(5):519–532, 2011.

Joanna P MacEwan, Julian M Alston, and Abigail M Okrent. The consequences of obesity for the external costs of public health insurance in the

united states. *Applied Economic Perspectives and Policy*, 36(4):696–716, 2014.

Zhen Miao, John C Beghin, and Helen H Jensen. Taxing sweets: sweetener input tax or final consumption tax? *Contemporary Economic Policy*, 30 (3):344–361, 2012.

Roger B Myerson. Incentive compatibility and the bargaining problem. *Econometrica: journal of the Econometric Society*, pages 61–73, 1979.

Marion Nestle. Why the 'consumer choice' argument is an automatic win for big soda, July 2012. URL www.theatlantic.com/health/archive/2012/07/why-the-consumer-choice-argument-is-a [Online; posted July-2012].

Cynthia L Ogden, Margaret D Carroll, and Katherine M Flegal. Prevalence of obesity in the united states. *Jama*, 312(2):189–190, 2014.

JA Rivera-Dommarco, M Hernández-Ávila, CA Aguilar-Salinas, F Vadillo-Ortega, and C Murayama-Rendón. Obesidad en México: recomendaciones para una política de estado. *México: Universidad Nacional Autónoma de México (UNAM)*, 2013.

Roland Sturm, Lisa M Powell, Jamie F Chriqui, and Frank J Chaloupka. Soda taxes, soft drink consumption, and childrens body mass index. *Health Affairs*, pages 10–1377, 2010.

Appendix-For On line Publication

Proof of Claim 1

Proof. Assume that the virtual surplus under tax regime (t_s, t_v) is larger than or equal to the virtual surplus generated when no tax regime is in effect. This would imply:

$$\theta_H v(q_H) - (\theta_H - \theta_L)v(q_L) - c(q_H) \leq \theta_H v(q_H) - (\theta_H - \theta_L)v(q_L) - \Psi_H$$

$$cq_H \geq \Psi_H$$

$$c \geq \psi = \left(\frac{t_s + c}{1 - t_v}\right)$$

The effective marginal cost per unit ψ takes arguments (t_s, t_v) on the domain $\mathrm{T} = \{[0,1] \times [0,1)\} \setminus (0,0)$. Since $t_s \in [0,1]$, $t_v \in [0,1)$, and $c > 0$, the only case when the inequality above is true is when $(t_s, t_v) = (0,0)$. However, by definition, $(0,0) \notin T$.

Therefore, for all valid points $(t_s, t_v) \in T$, it is true that $\left(\frac{t_s+c}{1-t_v}\right) > c$. This indicates that the virtual surplus under tax regime (t_s, t_v) strictly smaller than the virtual surplus generated when no tax regime is in effect.

Similarly for L-types, this reduces to show that for all valid points $(t_s, t_v) \in T$, it is true that $\left(\frac{t_s+c}{1-t_v}\right) > c$. $\qquad\square$

Proof of Proposition 5

Proof. The proof is straightforward. We only need to find out under which

condition π^{iia} is lower than π^{iib}:

$$\pi^{iia} = (1 - t_v)\Big\{(\beta)[\theta_L v(q_L^{iia}) - \psi q_L^{iia}] + (1 - \beta)\{[\theta_H v(q_H^{iia}) - (\theta_H - \theta_L)v(q_L^{iia})] - \psi q_H^{iia}\}\Big\}$$

$$\pi^{iib} = (1 - t_v)(1 - \beta)[\theta_H v(q_H^{iib}) - \psi q_H^{iib}]$$

$$\pi^{iia} < \pi^{iib} \implies [\theta_L v(q_L^{iia}) - \psi q_L^{iia}] < \left(\frac{1 - \beta}{\beta}\right)[(\theta_H - \theta_L)v(q_L^{iia})]$$

We make use of the previously shown fact that $\theta_H v'(q_H^{iia}) = \theta_H v'(q_H^{iib}) = \psi \implies q_H^{iia} = q_H^{iib}$. $\qquad\square$

Proof of Proposition 6

Proof. The proof is straightforward. We only need to find out under which condition π^{iia} is lower than π^{iic}:

$$\pi^{iia} = (1 - t_v)\Big\{(\beta)[\theta_L v(q_L^{iia}) - \psi q_L^{iia}] + (1 - \beta)\{[\theta_H v(q_H^{iia}) - (\theta_H - \theta_L)v(q_L^{iia})] - \psi q_H^{iia}\}\Big\}$$

$$\pi^{iic} = (1 - t_v)[\theta_L v(q_L^{iic}) - \psi q_L^{iic}]$$

$$\pi^{iia} < \pi^{iic} \implies$$

$$\beta[\theta_L v(q_H^{iia}) - \psi q_L^{iia}] + (1 - \beta)[\theta_H v(q_H^{iia}) - (\theta_H - \theta_L)v(q_L^{iia}) - \psi q_H^{iia}] < \theta_L v(q_L^{iic}) - \psi q_L^{iic}$$

$$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$$

Proof of Proposition 7

Proof. Proposition 7 is the extreme case of 7. This is, when the left-hand-side of inequality (35) is maximized. Here, we demonstrate that the left-hand-side of expression (35) is maximized when $\beta = 1$.

First, we need to show that left-hand-side of statement (35) (call it φ) increases as the fraction of L-types, gets larger. Differentiating this term

with respect to β:

$$\frac{d\varphi}{d\beta} = \{[\theta_L v(q_L^{iia}) - \psi q_L^{iia}] - [\theta_H v(q_H^{iia}) - (\theta_H - \theta_L)v(q_L^{iia}) - \psi q_H^{iia}]\}$$
$$+ \{ - \beta[\theta_H - \psi]\frac{dq_H^{iia}}{d\beta}\} \tag{41}$$
$$+ \{\beta[\theta_L v'(q_L^{iia}) + (\theta_H - \theta_L)v'(q_L^{iia}) - \psi]\frac{dq_L^{iia}}{d\beta}\}$$

Let us infer the sign of each expression between curly braces in equation (41). By assumption 1, the first term in curly braces is nonnegative. The second expression is zero, because from FOC$[q_H]$ in (17), we know that $\theta_H v'(q_H^{iia}) = \psi$ is independent from β, which implies $\frac{dq_H^{iia}}{d\beta} = 0$. Regarding the third expression, the term in square brackets is positive, since virtual surpluses are concave in quantity. Thus, in order to determine the sign of $\frac{d\varphi}{d\beta}$, we need to know that of $\frac{dq_L^{iia}}{d\beta}$, we can do so by differentiating FOC$[q_L]$ in (18):

$$\frac{dq_L^{iia}}{d\beta} = -\frac{\theta_L v'(q_L^{iia}) + (\theta_h - \theta_L)v'(q_L^{iia}) - \psi}{\beta\theta_L v''(q_L^{iia}) - (1 - \beta)v''(q_L^{iia})} \tag{42}$$

The denominator of equation (42)is negative because the retailer's objective function is concave. From equation (23), we know that $\theta_L v'(q_L^{iia}) > \psi$, thus the numerator is positive. These facts render $\frac{dq_L^{iia}}{d\beta} > 0$ and therefore $\frac{d\varphi}{d\beta} > 0$.

The result $\frac{d\varphi}{d\beta} > 0$ suggests that the left-hand-side of inequality (35) is maximized when $\beta = 1$, this would reduce the expression to $\theta_L v(q_H^{iia}) - \psi q_L^{iia}$. $\qquad\square$

Proof of Proposition 8

Proof.

From equations (11) and (22):

$$\theta_H v'(q_H^{ia}) = c < \theta_H v'(q_H^{iia}) = \psi \implies q_H^{ia} > q_H^{iia}$$

From equations (22) and (26):

$$\theta_H v'(q_H^{iib}) = \theta_H v'(q_H^{iic}) = \psi \implies q_H^{iic} = q_H^{iic}$$

$\square$

Proof of Proposition 9

Proof.

From equations (12) and (23):

$$\theta_H v'(q_L^{ia}) = \frac{c}{[1 - (\frac{1-\beta}{\beta})(\frac{\theta_H - \theta_L}{\theta_L})]} < \theta_L v'(q_L^{iia}) = \frac{\psi}{[1 - (\frac{1-\beta}{\beta})(\frac{\theta_H - \theta_L}{\theta_L})]}$$

which implies: $q_L^{ia} > q_L^{iia}$

From equations (12) and (30), $q_L^{ia} > q_L^{iic}$ iff:

$$\theta_L v'(q_L^{ia}) = c + (\frac{1-\beta}{\beta})(\theta_H - \theta_L)v'(q_L^{ia}) < \theta_L v'(q_L^{iic}) = \psi$$

which holds true is and only if:

$$\psi > c + (\frac{1-\beta}{\beta})(\theta_H - \theta_L)v'(q_L^{ia})$$

From equations (23) and (30), it is easy to deduce that:

$$q_L^{iic} > q_L^{iia}$$

$\square$

Proof of Proposition 11

Proof. The i-type consumers' value function is $U_i = \theta_i v(q) - p_i$. For L-types, we have: $U_L^{ia} = U_L^{iia} = U_L^{iib} = U_L^{iic} = 0$. On the other hand, for H-types: $U_H^{ia} = (\theta_H - \theta_L)v(q_L^{ia})$, $U_H^{iia} = (\theta_H - \theta_L)v(q_L^{iia})$, $U_H^{iib} = 0$, and $U_H^{iic} = (\theta_H - \theta_L)v(q_L^{iic})$.

As long as $\psi > \frac{c}{\left[1 - \left(\frac{1-\beta}{\beta}\right)\left(\frac{\theta_H - \theta_L}{\theta_L}\right)\right]}$, $q_L^{ia} > q_L^{iic} > q_L^{iib}$, thus $U_H^{ia} > U_H^{iia} > U_H^{iic} > U_H^{iib}$.

On the contrary, if $\psi < \frac{c}{\left[1 - \left(\frac{1-\beta}{\beta}\right)\left(\frac{\theta_H - \theta_L}{\theta_L}\right)\right]}$, then $\theta_L v'(q_L^{iic}) < \theta_L v'(q_L^{ia}) \implies q_L^{iic} > q_L^{iia}$. This implies $U_H^{iia} < U_H^{iic}$. $\qquad \square$

Proof of Proposition 10

Proof. Since expected profit decreases with cost, as confirmed by claim 1, we have $\pi^{ia} > \pi^{iia}$.

Because $q_H^{iia} = q_H^{iib}$, then $\pi^{ia} > \pi^{iib}$. Indeed, (from equation 26) $\theta_H \int v'(q_H^{iib}) \mathrm{d}q_H^{iib} = \theta_H v(q_H^{iib}) = \psi q_H^{iib} \implies \pi^{iib} = 0$.

From equation (30), $\theta_L \int v'(q_L^{iic}) \mathrm{d}q_L^{iic} = \theta_L v(q_L^{iic}) = \psi q_L^{iic} \implies \pi^{iic} = 0$. Thus, $\pi^{ia} > \pi^{iic}$. $\qquad \square$