The Effectiveness of Obesity Taxation *

Wei Xiao†

October 2011

Abstract

Obesity costs U.S. businesses billions of dollars each year in medical expenses and lost productivity. Many policy solutions are proposed and under debate to combat America's obesity epidemic, including the obesity taxation, which is a tax on high calory but low nutritious foods. Soft drink is one of the target food categories under debate and this paper examines the obesity taxation policy in this market. Different from the market level demand models, this paper studies the soft drink purchase timing, the choice between regular and diet products, and the quantity decision of individual households. From the scientific findings of the link between soft drink consumption and weight change, I simulate the effectiveness of various tax schemes on weight reduction. I find that the tax proportional to the calory content is around 6 per cent more effective than the tax on the food category and that an obesity tax on the soft drink can cause weight loss, though the effect is small.

Keywords: Obesity Tax, Carbonated Soft Drink, Proportional Hazard Model, Logit Model, Fixed Effects Model

*This paper is adapted from my Ph.D. dissertation at Johns Hopkins University. I am grateful to my advisors, Joe Harrington and Matt Shum, for their continuous encouragement and guidance. All remaining errors are my own.

†Competition Commission, Victoria House, Southampton Row, London WC1B 4AD, United Kingdom. Email: wei.xiao@cc.gsi.gov.uk

1 Introduction

Obesity and being overweight are the principal public health problems in the United States. It is likely to increase the risk of chronic disease and could cause low productivity. Studies even find that the prevalence of obesity and being overweight lead to higher consumption of gasoline, raising the concerns for energy security and the environment.[1] The epidemic has increased sharply among both adults and children just within the past few decades. For example, the obesity rate of children in the USA increased by 100% between 1980 and 1994. Recent national estimates indicate that 24% and 11% of children are above the 85th and 95th reference percentiles of body mass index (BMI), respectively.[2] The rapid spread of obesity and being overweight comes with huge economic cost, it is estimated that obesity costs U.S. businesses about $45 billion a year in medical expenses and lost productivity.[3]

Generally speaking, the cause for obesity and being overweight is that the energy, or calories, consumed exceed the energy expended in daily activities, and to lose weight, people can either consume fewer calories or expend more. Levels of physical activity do not seem to have declined during the past few decades, which means that the obesity epidemic must come from the increased caloric intake. Various diet factors could contribute to the excessive calorie intake, one such factor is the consumption of carbonated soft drinks. According to the data from the US Department of Agriculture (USDA), in the United States, per capita soft drink consumption has increased by 500% over the last 50 years and represents the largest single food source of calories in the US diet. From 1989-91 to 1994-95, average soft drink intake per day rose from 195 to 275 ml in the general population, and from 345 to 570 ml among adolescent boys. These figures approach or exceed the daily limits for total added sugar consumption recommended by the USDA. Among children of school age, total energy intake is positively associated with regular soft drink consumption, the average adjusted energy intake is 10% higher for children who regularly consume carbonated drinks compared to non-consumers.[4] Furthermore, the consumption of soft drinks often displaces other more nutritious drinks with less calories such as milk and fruit juice.

[1] See Li, Liu and Zhang (2009).
[2] See Ludwig et al (2001).
[3] See Arnst (2009).
[4] See James (2005) and Ludwig et al (2001).

Many policy solutions are proposed and under debate to combat America's obesity epidemic, such as the restriction on advertising, education programs, pulling out high calory "junk food" from schools, and other instruments comparable to the smoking reduction campaign. The taxation on high-fat and low nutrition foods (so-called "obesity taxes") is among these proposed solutions. In particular, given the current budget-busting recession, the obesity tax has become even more appealing to legislators. Twenty seven states have already imposed small tariffs of 7% to 8% on vending machine snacks such as candy, soda and baked foods. More states and municipalities are likely to try to impose taxes on soda and sweets. In late 2008, New York Governor David A. Paterson proposed an 18% sales tax on non-diet soda and sugary juice drinks for the fiscal year starting in April 2009. Such a tax, he said, would raise $404 million in 2009 and $539 million in 2010, and can be used for fat-fighting public health programs. Paterson has run into stiff opposition from the soft-drink industry, but several other states are mulling over such taxes. [5]

How effective these proposed "obesity tax" policies will be at reducing obesity? How can different taxation schemes achieve different outcomes? And is one policy better than another, in particular, comparing a universal tax rate on a food category to the one applying different rates depending on the content of calories? This paper is to shed light on these questions in the current policy debate, taking the case of the carbonated soft drink category.

Existing evaluations of obesity tax often rely on market level demand models.[6] However, aggregate demand changes are not appropriate to address the question of how a tax on soft drink might reduce obesity.[7] This paper studies the soft drink purchase decision of individual households from supermarket scanner data in a metropolitan area, and then, based on the scientific finding on the link between soft drink consumption and weight change, to simulate the effectiveness of various soft drink taxation policies on weight reduction.

The second section reviews the background and related literature. The third section presents the framework and the estimation strategy. The fourth section gives the details on the data and

[5]See Arnst (2009).

[6]For example, Rudd Report 2009.

[7]For example, if we find that a tax on soft drink would decrease sales by 20%, it is not clear whether it is because 20% consumers leave the market and all else consumes the same amount, or it is because every consumer reduces the consumption by 20%, or because of other changes. These scenarios would have quite different implications on obesity reduction. In other words, the distribution of the soft drink intake over households matters more than the total market sales.

shows the estimation results. The fifth section simulates and evaluates the effectiveness of different tax policies. And the last section concludes the paper.

2 Background and Related Literature

Although many factors are likely to contribute to the obesity epidemic, various evidences show a link between soft drink consumption and obesity. As a visual and astonishing fact that are unfamiliar to many, Kluger et al (2006) provide a good illustration of the linkage:

"drinking just two 12-oz. cans of soda a day - the average for kids 15 to 19 - and at the end of a week you have poured down enough sugar to fill a 1.5-lb. bag".[8]

Through a project in schools in four communities in Boston, Ludwig et al (2001) finds that the probability of being obese among children increased 1.6 times for each additional can or glass of sugar-sweetened drink in their daily consumption. In contrast, additional diet soda consumption is negatively associated with the obesity incidence.

Small changes of daily calories intake, such as by means of soft drink consumption, could have a big impact on the obesity risk.[9] For example, Hill et al (2003) suggests that altering the energy gap by 100 calories each day, which is the energy in an 8 fl oz. glass of regular soft drink, would prevent excessive weight gain in most adult Americans. Furthermore, they suggest that the daily consumption of one can of a regular carbonated drink (120 calories) over a ten year period in a constant environment theoretically will have the potential to result in a 50 kilograms weight gain.

Nestle (2003) describes the trend in the increasing amount of food sold and consumed at any one time since the mid-1980s. For example, Continuing Surveys of Food Intakes by Individuals (CSFII), conducted by the USDA, collected 24-hour recall lists from more than 11,000 individuals on three separate days. Comparing the 1989-1991 and 1994-1996 results, respondents reported consuming larger portions of nearly one third of 107 foods, and soft drink was among the food categories with an increased portion size.[10]

Larger portions have more calories and people tend to eat more when confronted with large

[8] Page 24.
[9] See Ebbeling and Pawlak (2002).
[10] See Smiciklas-Wright et al (2003).

amounts of food. Nestle (2003) further estimates that CSFII respondents increase the portion sizes of soft drink by 2 oz. (25 calories) on average, and an increase of 25 calories per day from soft drinks alone comes to more than 9,000 calories per year, which is an amount approaching a 3-pound weight gain. However, this estimate is descriptive and more about an illustration of a hypothetical scenario that is not based on actual consumption and rigorous quantitative analysis.

Given that around two thirds of Americans are overweight and the linkage between high-calory food consumption and obesity, many policies analogous to those on tobacco and alcohol are proposed to combat the obesity epidemic, such as the advertising ban and obesity tax. For example, a ban on sweetened-cereal advertising aimed at kids and a tax on high-fat and low-nutrition food are called in the U.S..[11]

Chou et al (2008) investigates the relationship between fast food restaurant advertising and children and adolescents overweight in the United States. They employ a pooled cross sectional micro level national data set for children (aged 3-11) and adolescents (aged 12-18) with their height and weight measures, combined with a fast food restaurant television advertising database at the designated market area (DMA) level, from 1997 to 1999.[12] The data is fitted in a linear model to examine how fast food restaurant advertising exposure affects children and adolescents' probabilities of being overweight.

They find a strong positive effect of exposure to fast food restaurant advertising on the probability of a child being overweight. In addition, they estimate the effect of a fast food restaurant television advertising ban on children and adolescents' obesity. A complete advertising ban on television would reduce the number of overweight children in a fixed population by 18 per cent and for adolescents, a decline of 14 per cent.

While the obesity tax on soft drinks was proposed and under debate,[13] to my knowledge, no study has examined the effect on weight reduction from a taxation on soft drinks. This paper is the first attempt to quantify the effectiveness of an obesity tax on the soft drink.

[11] See Huckabee (2006).

[12] The advertising data consists of the dollar expenditure and the annual number of seconds of messages aired on television for a wide range of fast food restaurant chains.

[13] For example, see Rudd Report in Fall 2009: Soft Drink Taxes: A Policy Brief.

3 The Model and Estimation

Consumer's decision on soft drink purchase can be modeled in three stages. Firstly she decides whether to make a purchase at some time point. Secondly she choose between regular and diet drinks. And thirdly she decides the purchase quantity. This section describe the model and estimation strategy for each of the three stages.

3.1 First Stage: The Rate of Purchase

Consider a random variable T that represents the time until an event occurs (e.g. the purchase of a soft drink) and let $h(t)$ denote the hazard function of T. The hazard function specifies the instantaneous probability of purchase given that no purchase has been made up to time t, and is defined as

$$h(t) = \lim_{\delta t \to 0} \frac{P[t \leq T \leq t + \delta t | T \geq t]}{\delta t} = \frac{f(t)}{1 - F(t)} = \frac{f(t)}{S(t)},$$

where $f(t)$, $F(t)$ and $S(t)$ are the probability density function (pdf), the cumulative distribution function and the survivor function of T, respectively. It is clear that the hazard function is finite and non-negative. It can be shown that the three specifications $f(t)$, $h(t)$ and $S(t)$ for the probability distribution of T are interrelated as follows (Kalbfleisch and Prentice 1980):

$$S(t) = exp[-\int_0^t h(u)du] \tag{3.1}$$

$$f(t) = h(t)exp[-\int_0^t h(u)du]. \tag{3.2}$$

It is clear that there is a one-to-one relationship between the pdf of T and its hazard function. Hence, the hazard function uniquely determines the pdf of T, and we can restrict ourselves to $h(t)$ in studying the purchase timing decisions of households.

Let $h(t|X_1, \theta)$ represent the hazard function of T conditional on a vector of covariates (X_1) and an unobserved heterogeneity component (θ). We use for $h(t|X_1, \theta)$ the following representation:[14]

$$h(t|X_1, \theta) = h_0(t) \cdot \psi(X_1) \cdot \phi(\theta). \tag{3.3}$$

In equation (3.3), $h_0(t)$ denotes the baseline hazard, $\psi(X_1)$ a function of the covariates, and $\phi(\theta)$ the specification for the unobserved heterogeneity. We assume that $h_0(t)$, $\psi(X_1)$ and $\phi(\theta)$ are all non-negative. If the effects of the covariates and the unobserved heterogeneity are ignored, then the purchase timing decision is characterized only by the baseline hazard function and hence, the corresponding probability density function. The effects of covariates and the unobserved heterogeneity are to shift the hazard from its baseline.

(3.1) is a general formulation. An important issue is the specification of the $h_0(t)$, $\psi(X_1)$ and $\phi(\theta)$ functions. Let us first consider $h_0(t)$. There are two approaches available. In the Cox model, $h_0(t)$ is simply left unparameterized, and through conditioning on failure times, estimates of coefficient are obtained. In the parametric approach, a functional form for $h_0(t)$ is specified. For example, if we assume $h_0(t) = exp(a)$ for some a, then we have the exponential model. The baseline hazard is assumed constant over time and there is an extra parameter, a, to estimate. If we assume $h_0(t) = exp(a)exp(\gamma t)$, then we obtain the Gompertz model, with two ancillary parameters a and γ to estimate.

The next issue is the specification of the $\psi(X_1)$ function that measures the effects of covariates. We require $\psi(X_1)$ to be a nonnegative function. For reasons of simplicity, we use

$$\psi(X_1) = exp[\sum_{j=1}^{J} X_{1j}(\tau)\beta_j],$$

where $X_{1j}(\tau)$ is the value of the jth covariate, $j = 1, 2, \ldots J$, at calendar time τ when purchase is made, β_j is the associated coefficient.[15]

$\phi(\theta)$ captures unobserved heterogeneity. The effect, $\phi(\theta)$, is known as a frailty and represents

[14]See Cox (1972), Lancaster (1979), Heckman and Singer (1984) and Flinn and Heckman (1982, 1983).

[15]The covariates are allowed to vary with time but cannot be explicit functions of time.

that individuals in the population are heterogeneous because of factors that remain unobserved. The frailties are positive quantities not estimated from the data but instead assumed to have mean one (for purpose of identifiability) and variance θ, and θ is estimated from the data.[16] Since the purchase observations are correlated within households because of some overall group characteristic (a frailty) that is not being measured, we want to model the correlation using a shared frailty model. Shared frailty models are the survival-data analogous to random effects models. To illustrate, the hazard function can be written as[17]

$$h(t_{ij}|X_{ij1}, \theta) = h_0(t_{ij}) \cdot \psi(X_{ij1}) \cdot \phi_i(\theta).$$

Here the index i denotes the household (i = 1,. . . , n), and j denotes the observation within household ($j = 1, \ldots, n_i$). The frailties, $\phi_i(\theta)$, are shared within each group and are assumed to follow either a gamma or inverse-Gaussian distribution. The frailty variance, θ, is estimated from the data and measures the variability of the frailty across groups.

Since the hazard rate is an observed quantity, to estimate the parameters of this model using observed household purchase data, we use the maximum likelihood method. The parameters to be estimated are the distribution parameters of $h_0(\cdot)$, β_j and θ. Denoting these parameters as ν, the likelihood function of ν for the ith household conditional on θ is

$$L_i(\nu|\theta) = \sum_{r_i=1}^{R_i} [f(t|\theta)]^{\delta_{r_i}} [S(t|\theta)]^{1-\delta_{r_i}}, \tag{3.4}$$

where R_i is the total number of purchase of the ith household, and

$$\delta_{r_i} = \begin{cases} 1 & \text{if the } r_i\text{th spell ends in a purchase} \\ 0 & \text{otherwise} \end{cases}$$

The terms $[S(t|\theta)]^{1-\delta_{r_i}}$ accounts for right censoring, as the last purchase generally does not

[16]For example, $\phi(\theta) = exp(c\theta)$, where θ follows standard normal distribution with mean zero and variance one, then $\phi(\theta)$ is distributed with mean one and variance $exp(c)$.

[17]In other places, the indexes i and j are suppressed for simplicity and clarity.

coincide with the end of the data collection period. In the estimation, we do not account for left censoring of the data which occurs when the first purchase does not coincide with the beginning of the observation period. Substituting for $S(t|\theta)$ and $f(t|\theta)$ from (3.1) and (3.2) in equation (3.4), and assuming the covariates remain constant during spells, we get

$$L_i(\nu|\theta) = \sum_{r_i=1}^{R_i} [h(t|X_1, \theta) exp[-\int_0^t h(u|X_1, \theta) du]^{\delta_{r_i}} \times [exp[-\int_0^t h(u|X_1, \theta) du]]^{1-\delta_{r_i}}.$$

The unconditional likelihood function $L_i(\nu)$ is obtained by integrating over the distribution of θ, i.e.

$$L_i(\nu) = \int_\theta L_i(\nu|\theta) dG(\theta),$$

where $G(\cdot)$ is the distribution function of θ. The parameter estimates are obtained by maximizing the likelihood function across all N households in the sample:

$$L(\nu) = \sum_{i=1}^{N} L_i(\nu).$$

3.2 Second Stage: Choice Between Regular and Diet

Conditional on purchasing the soft drink, the consumer chooses either a regular soft drink ($Regular = 1$) or a diet one ($Regular = 0$).[18] We believe that a set of factors, such as price and family size, gathered in a vector X_2 explain the decision, so that

$$Prob(Regular = 1|X_2) = F(X_2, \beta_2)$$

$$Prob(Regular = 0|X_2) = 1 - F(X_2, \beta_2).$$

The set of parameters β_2 reflects the impact of changes in X_2 on the probability. Assume that $F(\cdot)$ follows the logistic distribution, then we have the familiar logit model and the probability can

[18] It is possible that some households purchase both regular and diet products on one shopping trip. In this case, the choice of regular drink and the choice of diet drink are treated as separate observations, as if the purchases were made by two identical households.

be written as

$$Prob(Regular = 1|X_2) = \frac{e^{X_2'\beta_2}}{1 + e^{X_2'\beta_2}}.$$

Estimation of the regular and diet choice model is based on the method of maximum likelihood. Each observation is treated as a single draw from the logistic distribution and the joint probability, or likelihood function is

$$Prob(R_1 = r_1, R_2 = r_2, \dots, R_n = r_n|X_2) = \sum_{r_i=0}[1 - F(X_{2i}\beta_2)] \sum_{r_i=1} F(X_{2i}\beta_2),$$

where X_2 denotes $[x_{2i}]_{i=1,\dots,n}$. The likelihood function for a sample of n observations can be conveniently written as

$$L(\beta_2) = \sum_{i=1}^{n}[F(X_{2i}\beta_2)]^{r_i}[1 - F(X_{2i}\beta_2)]^{1-r_i},$$

here $r_i = 1$ if the regular drink is chosen and $r_i = 0$ if diet drink is chosen.

3.3 Third Stage: Purchase Quantity

Conditional on consumer i making a purchase and after she decides on a regular or diet soft drink, she will then consider the purchase quantity. The purchase quantity of consumer i at time t can be modeled in the regression:

$$Y_{it} = X_{3it}'\beta_3 + Z_i'\alpha + \varepsilon_{it}$$

where X_{3it} is a vector of variables that may affect the consumer's quantity demand, such as the price and family size; and Z_i is the individual heterogeneity.

The estimation for the panel demand quantity is standard. For the fixed effects model, the classic least square regression is sufficient, while for the random effects model, the generalized least squares (GLS) is fitted.

3.4 Overall Effect and Tax Policy Evaluation

What we are interested in is how does tax policy affect an individual's consumption of soft drinks. Denote the joint distribution of the above three stages as

$$D(Hazard, Regular, Quantity) \equiv D(H = h, R = r, Q = q)$$

Here H, R and Q stand for purchase frequency, choice of regular or diet drink, and the purchase quantity respectively; h, r and q are consumer's choices respectively. On a specific day, the consumer's consumption of soft drink is the overall effect of the three stages:

$$C(h, r, q) = h \cdot r \cdot q$$

The economic meaning of the above equation is that at a time point, the individual's consumption of soft drink depends on the probability that she would make a purchase times her probability of choosing a regular or diet drink and then times the quantity she would buy. Consequently, the expectation of the soft drink consumption can be written as

$$E[C(h, r, q)] = \int_h \int_r \int_q h \cdot r \cdot qD(h, r, q)dhdrdq.$$

Assume that the distribution of purchase frequency, choice of regular or diet drink, and the quantity decision are independent,[19] then the soft drink consumption becomes

$$E[C(h, r, q)] = \int_h \int_r \int_q h \cdot r \cdot qD(h)D(r)D(q)dhdrdq$$

$$= \int_h hD(h)dh \cdot \int_r rD(r)dr \cdot \int_q qD(q)dq = E(h) \cdot E(r) \cdot E(q).$$

[19] This is also the implicit assumption for the estimation of three-stage purchase decisions. If the choice of regular or diet drink and the quantity decision are dependent, the estimates of the quantity equation are biased and inconsistent, and logit estimates of the second stage are inefficient. The two-step maximum likelihood estimate (2SML) suggested by Lee and Trost (1978) or the consistent two stage procedure suggested by Heckman (1979) can be used for joint estimation. For more discussion on estimating dependent models, see Dubin and McFadden (1984) and Krishnamurthi and Raj (1988). Krishnamurthi and Raj (1988) jointly estimate households' brand choice and purchase quantity for two categories of frequently purchased products, and they find that the price coefficients of quantity equation in the 2SML are generally depressed in magnitude. In this case, the price effect on the purchase quantity is overestimated in my specification and so as the effect of obesity tax.

Given a taxation policy, denoted as T, its impact on individual's consumption will be

$$\frac{dE[C(h,r,q)]}{dT} = \frac{dE(h)}{dT_h} \cdot \frac{dE(r)}{dT_r} \cdot \frac{dE(q)}{dT_q}.$$

Here T is the overall tax policy. T_h, T_r and T_q are the taxation effects on purchase hazard, choice on drink types and purchase quantity respectively, and they may be different under the same taxation scheme. In terms of the elasticity, which is the percentage change in the soft drink consumption in response to a one percent price change caused by the tax policy, it can be written as

$$\frac{dE[C(h,r,q)]/E(C(h,r,q)}{dT/T} = \frac{dE(h)/E(h)}{dT_h/T_h} \cdot \frac{dE(r)/E(r)}{dT_r/T_r} \cdot \frac{dE(q)/E(q)}{dT_q/T_q} \equiv e_h \cdot e_r \cdot e_q, \tag{3.5}$$

where e_h, e_r and e_q are the price elasticities in the consumer's three decision stages from the same tax policy.

4 Data and Results

4.1 Data

The data used for demand estimation is a scanner data set collected by Information Resource Inc.(IRI), containing household-level soft drink purchase data and store-level sales data in 8 store chains in a large metropolitan area. The sample period is 104 weeks from June 1991 to June 1993. The household-level data tracks 1024 household's soft drink shopping history with the information on shopping date, units and prices of products purchased, store locations and household demographics. The store-level data contains the weekly prices and the promotional activity each week. The products' calory level are obtained from CocaCola and PepsiCo's websites or supermarket packages, assuming that the contents of a brand do not change over time.

In order to obtain a static sample of continuously active households, I eliminate those panelists who made less than eight purchases over the observed period of 104 weeks. This is done in order to avoid mistaking unrecorded purchases for "observed" lengthy inter-purchase times. This results in 912 households with 35755 spells. The mean inter-purchase times for soft drink is 17 days.

Table 1 presents the summary statistics of the covariates used in the three stages of the model, including the relevant prices, market mixes and demographics. The reported statistics are for the sample taken to estimation. The prices used in the three stages are in different measures to reflect the different impacts of price in each stage of the customers' decision process. In the first stage of purchase frequency, it is the overall price of the soft drink category which affects the customer's hazard rate to participate in the market. In the second stage, when a household makes the choice between regular and diet soft drinks, the aggregate prices of regular drink and diet drink take effect. And in the third stage, after a customer has chosen a certain product or products and starts to consider the quantity to buy, the price of the chosen products is expected to affect the customer's decision.

The bottom section of Table 1 provides the statistic summary of household demographics in the sample. The family size is the number of household members, which is capped at six and if the number of family members is more than six, it takes the value of six. The income is the family income, which is categorized into eleven windows. The values one to eleven represents the income from low to high. Male and female educations are category variables with values zero to seven, ordered from low to high educational levels.[20] Given that the income and education measures are discrete, I only interpret their coefficients qualitatively but not quantitatively in the estimation results.

4.2 First Stage: The Rate of Purchase

The set of covariates comprises the average per pound price of all the soft drink products in the week that the household made the purchase,[21] and the family size of the household. I also include a promotion dummy, which equals one if the purchased products have a feature, a display or a coupon when the purchase was made. Since seasonal effect is important in the soft drink product

[20] Value zero of male (female) education means that there is no male (female) in the household. Value one is some grade school and value seven is post graduate study.

[21] The average price here is the simple average of all products, not the weighted average by the products' market share. This is because when a consumer makes a purchase decision, she does not have knowledge of the market shares of soft drink products, therefore the market shares are not likely to influence her purchase decision.

market, for example, the demand peaks in the summer and the pricing strategy by manufacturers or retailers may also be different over the seasons, not controlling for the seasonal effect may result in a biased estimation. Thus I also include the dummy variable of weekdays, which has value one if the purchase was made from Monday to Friday, and the dummy variable of different seasons.[22] The estimation results are presented in Table 2.

As expected, the overall soft drink price has a significant negative effect on the hazard rate in making a purchase and the coefficients of all three models are at a similar level (-0.73 to -0.85). The product promotion increases a consumer's likelihood to make a purchase, as reflected by the significant positive coefficients. Also, the larger the family size, the probability to buy the soft drink is higher. The seasonal effects of soft drink demand are confirmed in the estimation results as well. In the summer season (June to August), the consumers buy soft drinks most frequently, followed by the fall (September to November) and winter (December to February). The price elasticities on the purchase duration are reported in the bottom of Table 2, the elasticities from the three models do not have big differences, ranging from 0.62 to 0.72. However, the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) tests slightly prefer the exponential model,[23] so the estimates from the exponential model will be used in the tax policy effects simulation.

I also test the effect of the chosen product price on household's purchase hazard rate. It turns out that the single product price does not affect the purchase frequency. I want to emphasize that it is the average price of the soft drink market affect the purchase frequency, not the single product price. This fact would have different policy implications, as will be illustrated in the policy simulation section.

4.3 Second Stage: Choice Between Regular and Diet

Consumers' choice on regular or diet soft drinks is fitted in a logit model. The dependent variable is the binary outcome of the consumer's choice, which is defined as one if the regular drink is picked up and as zero if the diet drink is chosen. The key explanatory variable is the prices of regular and diet soft drinks, since both the regular drink price and diet drink price will impact the consumer's

[22] More specifically, the variable"season1" is the dummy for months June to August, "season2" is the dummy for months September to November, and "season3" is for months December to February of the following year.

[23] Smaller values of AIC or BIC correspond to better fitting models.

decision on whether to buy the regular drinks, and the estimated price coefficients can be used to evaluate different tax policies. For example, comparing the universal tax on all soft drinks to taxing only the regular soft drinks with high calories is of interest. There are several choices of price variables. I tried the average and median of the product-level prices per pound.[24] The estimation results are reported in Table 3.

The first section of Table 3 presents the coefficient estimates of two specifications. All the estimates are significant with expected signs. The price of regular drinks has a negative effect on the probability of consumers choosing the regular drink, while the higher the price of diet drinks the higher the probability that the regular product will be bought. The education levels of both male and female in the household have negative sign, meaning that higher educated people are more likely to drink diet products. A relative finding is that high income households prefer diet drinks. Considering that the education level and income are positively correlated, the results are consistent in this sense. The bottom section of Table 3 shows the price elasticity of regular and diet products, which measures the changes of probability to purchase regular drinks in response to a price change in percentage terms.[25] The elasticities of the regular drink's price ($0.60 - 0.65$) and diet drink's price (0.35) are at a similar level for both specifications. The middle section of Table 3 gives the log likelihood of the two models. The model using the mean price per pound has a higher log likelihood, so it will be used in the policy effects simulation. It is interesting to note that the significance of regular drinks' price elasticity is higher than the diet drinks', suggesting that applying the same tax rate on both types of drinks, the negative effect on a household's choice of a regular drink from a change in the prices of regular drink will outweigh the positive effect from changes in diet drink prices.

4.4 Third Stage: Purchase Quantity

For the quantity regression, I tried both the fixed effects and random effects models. Conditional on that consumers have made the purchase and decided on the type, the price of the product of

[24] For the same reason as in the first stage, the prices here are unweighted.
[25] These elasticities are calculated at the means of independent variables.

her choice is supposed to affect her quantity decision. I also control for the coupon dummies,[26] household characteristics such as the education level of male and female, the family size and the income. Table 4 reports the regression results.

The price coefficient is significantly negative at a similar level in both models (-0.316 and -0.318). Since the dependent variable is the log of the quantity and the price is also in log form, the price coefficient is the elasticity of quantity to price. The coefficient of the coupon dummy is significantly positive as expected in both models (0.30 and 0.31), because when there is a feature, display or coupon discount, the consumers are likely to buy more of the products. The Hausman specification test favors the fixed effects model[27] , so I will take the price coefficient from the fixed effect model to simulate the effects of taxation policies.

5 Taxation Policy Evaluation

Based on the above estimation results, we can now simulate the effects of various tax policies based on the scientific finding in Nestle (2003) that an additional 3000 calories per year would result in a 1 pound weight gain. After straightforward unit conversion, this turns into the fact that the intake of 6.614 calories leads to a one gram weight gain.[28] We know that one ounce of regular soft drink contains around 12.5 calories, so if we know how much an obesity tax policy will change the amount the household consumes regular soft drinks per year, we can calculate the weight change of the household in that year.

Table 5 presents the results for this section, and I will review them in detail in the following sections. The first section provides an overview of the consumption of regular soft drink and the consumed calories in the United States. The first row presents the distribution consumption of regular drinks per year in ounces from the data set. And the second row is the conversion of individual's calory intake from the consumption of regular soft drinks per year.[29]

[26]IRI categorizes the coupons into seven types depending on their source and format, but I do not distinguish them and group them into one. "Coupon dummy" has the value of one if a coupon is available on the shopping day and zero if not.

[27]The F ratio for testing the joint significance of the individual effects is $F(2, 36380) = 711.69$, which is strongly in favor of a household specific effect in the data.

[28]One pound is equivalent to 453.59 grams.

[29]The conversion is an approximation that one ounce of regular soft drink is equivalent to 12.5 calories. The actual

5.1 Universal Tax v.s. Tax on Ingredients

From (3.5) the overall elasticity of regular soft drink consumption with respect to the taxation rate is the product of the price elasticity of each of the three decision stages. Our purpose is to compare the effectiveness of the universal tax rate on the whole soft drink category to the taxation depending on the the calory level.

Two effects come into play in this comparison. On the one hand, universal taxation has a greater impact than the targeted tax of shortening the purchase duration, so it would lead consumers to lose more weight. On the other hand, universal taxation has less impact than the targeted tax on consumers purchasing regular soft drinks, which would lead consumers to lose less weight. These two effects counteract each other, and which effect weigh more is an empirical question.

To illustrate the effect of obesity taxation on weight reduction, suppose that the government imposes a 1% tax on all soft drink products. And assume that the consumers bear all the tax, in the form of a price increase of 1%.[30] We want to find out how this tax policy affects the households' regular soft drink consumption over a year. The first stage estimation shows that the price elasticity on the inter-purchase duration is 0.68, which means that the purchase frequency in one year will reduce by 0.68% from the 1% price increase, or a 1% tax rate, and consumption will become $(100 - 0.68) = 99.32$ percent of the original amount. Since this is a universal tax, both the regular and diet drinks' price will rise by 1%, so the probability that the household will choose regular drink becomes $(100 - 0.65 + 0.35) = 99.70$ percent of the probability before the tax. Conditional on making the purchase and choosing the product type, the purchased quantity will reduce by 0.32%, from the elasticity presented in Table 4. The overall regular drink consumption after a 1% tax as a percentage of the pre-taxation level will be a product of the above three shares, i.e. $99.32\% \times 99.70\% \times 99.68\% = 98.71\%$. Since one ounce of regular soft drink corresponds to a certain amount of calory intake, the reduction of calory intake is reduced by $100 - 98.71 = 1.29$ percent as a result of the 1% universal taxation.

The same principal can be used to calculate the effect on calory intake from a 1% tax on regular

calories contained in eight ounce of regular drink range from 91 to 120.

[30] Under the assumption that the entire tax is passed through to the final price, the estimated tax effect on weight reduction is the upper bound.

soft drinks only, with the exception of two differences. Firstly, since the overall soft drink market price affects the consumers' inter-purchase duration, we need to find out how the price change in the regular product affects the market price level. Regular drinks account for around 60% of all the sales in our data set, so I assume that the 1% price increase in regular products turns into a $60\% \times 1\% = 0.6\%$ market price increase. Secondly, in the second stage decision to choose between a regular and diet drink, when the diet product price does not change, the effect on household's probability to buy regular drinks depends solely on the price elasticity of the regular product, i.e. -0.65%. After these two adjustments, I derive that the reduction of calory intake is reduced by 1.37%.

Now we can see that targeted taxation of 1% on regular drinks with high calories is more effective than a universal tax scheme on weight reduction, and more precisely, the targeted taxation is $100 \times (1.37/1.29 - 1) = 5.96$ percent more effective than the universal taxation. Similarly, for the 5% and 10% tax rates, the targeted tax scheme is 5.79% and 5.57% more effective, respectively.

5.2 Measuring the Effectiveness of an Obesity Tax

The individual regular soft drink consumption is defined as the annual regular drinks purchased by the household divided by the number of family members, assuming that each family member consumes the same amount of soft drinks purchased. The idea to evaluate various obesity tax policies is to derive the reduction of regular drinks consumption across the consumption percentiles due to those policies, as it can be converted into the amount of calory intakes, which is linked to the weight reduction from the scientific finding that 6.614 calories can be transferred into one gram of weight gain.

The first section of Table 5 gives the consumption of regular soft drinks in a year in percentiles and the corresponding calory intake. From the elasticity analysis in the last section, I simulate the reduction in calory intake across the percentiles of annual regular soft drink consumption from different taxation schemes, and the results are presented in the bottom three sections in Table 5. Here is an example on how to read the table. Given a universal tax rate of 10%, the person at the 90th percentile of consuming regular soft drinks would reduce consumption of 3584.69 calories per year. Since 6.614 calories a year transforms into one gram, this suggests that the person at the 90th percentile will lose weight by $3584.69/6.614 = 540$ gram in a year.

A more interesting question is to evaluate the obesity tax policy's impact on weight reduction of the overweight or obese group. One constraint here is that I do not have the weight or height information of the panelists in the data set. Being overweight and obese are determined by both a person's weight and height as the "body mass index" (BMI), which is defined as the body weight in kilograms divided by the square of height in meters (kg/m^2). Being overweight is classified as BMI in the range from 25.0 to 29.9, obesity is for BMI greater than or equal to 30.0, and when BMI exceeds 40.0 it is extreme obesity. According to the latest release of the National Health and Nutrition Examination Survey (NHANES), among adults aged 20 or older in 2007-2008, approximately 34.2% is overweight, 33.8% is obese and 5.7% is extremely obese in the U.S..[31]

Based on the finding of Ludwig et al (2001) that regular soft drink consumption and being overweight are positively associated, I assume that the amount of regular soft drink consumption is (perfectly) positively correlated to the BMI distribution, i.e. consumption and the BMI can be sorted in the same order and the percentiles can be matched. Under this assumption, the simulated tax effect on weight reduction of the obese or overweight population is the upper bound.[32] Given that the obesity rate is 33.8%, the corresponding regular soft drink consumption percentile is 66.2% ($(100 - 33.8)\%$) and above. Similarly, the regular soft drink consumption percentile for the overweight group is between 33% ($(100 - 34.2 - 33.8)\%$) and 66.2%, and for the extreme obese group the percentile is 94.3% ($(100 - 5.7)\%$) and higher.

Table 6 shows the weight reduction effect on the marginal percentiles of three groups of the overweight, obese and extreme obese under different taxation policies. The calculation is based on the same principle as in Table 5. Let us first look at the most impacted group in the table. The targeted 10% tax on the regular drinks would reduce the weight of the marginal extreme obese (94.3th percentile) person by 725.64 grams (1.6 pounds) per year. In ten years, a 10% taxation on regular drinks could accumulate to reducing the weight by 16 pounds on the marginal extreme

[31]NHANES is maintained by the Centers for Disease Control and Prevention (CDCP). See National Center for Health Statistics Heath (NCHSH) E-Stat for more detailed summarized statistics.

[32]This is because the heavily consumers are more affected by the taxation and thus would reduce more weight. When the regular drink consumption and the weight are not perfectly correlated, more weight reduction would go to the group that is not overweight or obese, so the effect of taxation on overweight reduction would be less significant.

obese person.

The obesity taxation's effect on individuals depends on their position in the BMI and soft drink consumption distribution. However, the effect of the obesity tax is positively monotonic on the consumption percentiles under the assumption that the BMI and regular soft drink consumption distributions are perfectly positively correlated, therefore I can estimate the range of weight reduction for the overweight groups. Specifically, in a year, a 10% target taxation on regular drink could reduce the weight of overweight (but not obese) persons by a range of 69.19 grams (0.15 pounds) to 210.7 grams (0.46 pounds); could reduce the obese (but non-extreme obese) persons' weight by 210.7 grams (0.46 pounds) to 725.64 grams (1.6 pounds); and could reduce the extreme obese group by more than 725.64 grams (1.6 pounds). In comparison, the annual weight reduction from a 5% tax on regular soft drink for overweight, obese and extreme obese groups range from 0.08 to 0.24 pounds, from 0.24 to 0.81 pounds and over 0.81 pounds respectively.

6 Conclusion

This paper examines the the effectiveness of an obesity tax on soft drinks under the current policy debate to combat the epidemic of obesity and being overweight in the United States. It sheds light on two questions: the effectiveness of different obesity taxation schemes in reducing obesity and being overweight, in particular, comparing the tax applied to the category of foods to a tax proportional to the content of calories; and the effectiveness of a tax on soft drinks in reducing the weight of obese and overweight people.

Consumer's decision on soft drink purchases are modeled in three stages. Firstly she decides whether to make a purchase at some time point through a proportional hazard model. Secondly she chooses between regular and diet drinks through the logit model. And thirdly she decides the purchase quantity through the fixed effects model.

On the inter-purchase frequency, seasonal effects exist in soft drink demand. The consumers buy soft drinks most frequently in the summer, followed by the fall and winter. And the chosen product price does not affect a household's purchase hazard rate, instead, the overall average price of the soft drink market would affect the consumers' purchase frequency. On the household's choice between regular and diet drinks, both regular and diet prices have effects. However, the significance of a regular drink's price elasticity is higher than the diet drinks, suggesting that applying the same tax rate on both types of drinks would discourage the consumers buying regular drinks. And for the quantity decision, when there is a feature, display or coupon discount, the consumers are likely to buy more of the products.

Finally, I find that the tax proportional to the calory level is around 6% more effective than the universal tax on weight reduction. The obesity tax on soft drinks is effective in reducing the weight of overweight groups. Specifically, within a year, a 10% target tax on regular drinks as a maximum could reduce the weight of overweight (but not obese) persons in the range of 0.15 to 0.46 pounds; could reduce the obese persons' weight by 0.46 to 1.6 pounds; and could reduce the extreme obese person's weight by over 1.6 pounds.

References

[1] Arnst, C, “Taxing the Rich - Foods, That Is ”, *Business Week*, (February 12, 2009)

[2] Chou, Shin-Yi, Inas Rashad and Michael Grossman, “Fast-Food Restaurant Advertising on Television and Its Influence on Childhood Obesity ”, *Journal of Law and Economics*, (2008), Vol.51, pp. 599-618

[3] Cox, D.R., “Regression Models and Life-Tables ”, *Journal of Royal Statistical Society*, (1972), Vol.34, pp. 187-200

[4] Dubin, J.A. and D.L. McFadden, “An Econometric Analysis of Residual Electric Appliance Holdings and Consumption ”, *Econometrica*, (1984), Vol.52, pp. 345-62

[5] Ebbeling, C. and D.B. Pawlak, “Childhood Obesity: Public Health Crisis, Common Sense Cure”, *Lancet*, (2002), Vol.360, pp. 473-82

[6] Flinn, C. and J. Heckman, “Models for the Analysis of Labor Force Dynamics”, *Advances in Econometrics*, (1982), Vol.1, pp. 35-95

[7] Flinn, C. and J. Heckman, “The Likelihood Function for the Multistate-Multiepisode Model in ‘Models for the Analysis of Labor Force Dynamics‘”, *Advances in Econometrics*, (1983), Vol.2, pp. 225-31

[8] Heckman, J., “Sample Selection Bias as a Specification Error ”, *Econometrica*, (1979), Vol.47, pp. 153-61

[9] Heckman, J. and B. Singer, “A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data ”, *Econometrica*, (1984), Vol.52, pp. 271-320

[10] Hill, J., H.R. Wyatt, G.W. Reed and J.C. Peters, “Obesity and Environment: Where Do We Go From Here”, *Science*, (2003), Vol.299, pp. 852-57

[11] Huckabee, M “Battling America’s Obesity Epidemic”, *Time*, May 8, 2006, Vol.167, Issue 19, pp. 80

[12] Jain, Dipak C. and Naufel J. Vilcassim, "Investigating Household Purchase Timing Decision: A Conditional Hazard Function Approach ", *Marketing Science*, (1991), Vol.10, No. 1, pp. 1-23

[13] James, J and D. Kerr, "Prevention of childhood obesity by reducing soft drinks ", *International Journal of Obesity*, (2005), Vol. 29, S54-S57

[14] Kalbfleisch, J.D. and R.L. Prentice, *The Statistical Analysis of Failure Time Data*, New York: John Wiley and Sons (1980)

[15] Kluger, J., J. Caplan, E. Kauffman, J. Ressner, B. Rubiner and K. Tumulty "How Bill Put the Fizz in the Fight Against Fat", *Time*, May 15, 2006, Vol.167, Issue 20, pp. 22-25

[16] Krishnamurthi, L. and S.P. Raj, "A Model of Brand Choice and Purchase Quantity Price Sensitivities ", *Marketing Science*, (1988), Vol.7, No. 1, pp. 1-20

[17] Lancaster, T., "Econometric Methods for the Duration of Unemployment", *Econometrica*, (1979), Vol.47, pp. 939-56

[18] Lee, L.F. and R.P. Trost, "Estimation of Some Limited Dependent Variable Models with Application to Housing Demand", *Journal of Econometrics*, (1978), Vol.8, pp. 357-82

[19] Li, S, Y. Liu and J. Zhang, "Lose Some, Save Some: Obesity, Automobile Demand, and Gasoline Consumption in the U.S. ", *RFF Discussion Paper 09-34*, (2009)

[20] Ludwig, D, K. Peterson and S. Gortmaker, "Relation between consumption of sugar-sweetened drinks and childhood obesity:a prospective, observational analysis ", *The Lancet*, (2001)

[21] Nestle, M, "Increasing portion sizes in American diets: More calories, more obesity ", *Journal of the American Dietetic Association*, (2003)

[22] Philipson, T and R. Posner, "Is the Obesity Epidemic A Public Health Problem? A Decade of Research on The Economics of Obesity ", *NBER Working Paper 14010*, (2008)

[23] Rudd Report, "Soft Drink Taxes: A Policy Brief ", *Rudd Center for Food Policy and Obesity*, (Fall 2009), Yale University

[24] Smiciklas-Wright, H., Mitchell, D.C., Mickle, S.J., Goldman, J.D. and A. Cook, "Foods Commonly Eaten in the United States, 1989-1991 and 1994-1996: Are Portion Sizes Changing? ", *Journal of the American Dietetic Association*, (2003), Vol 103, pp. 41-47

Table 1: Statistic Summary

	Mean	Std. Dev.	Min	Max
First Stage				
No. of Purchase	64.29	44.69	8	303
Weekly Mean Price	0.85	0.015	0.80	0.88
Promotion Dummy	0.51	0.50	0	1
Second Stage				
Regular Drink Choice	0.64	0.48	0	1
Regular Price*	0.81	0.10	0.65	1.20
Diet Price*	0.75	0.09	0.61	1.09
Third Stage				
Quantity**	6.49	5.93	0.22	117
Price per Pound	0.61	0.24	0.08	1.99
Coupon Dummy	0.08	0.27	0	1
Demographics				
Family Size	2.67	1.40	1	6
Income	5.97	3.04	1	11
Male Education	3.42	2.51	0	7
Female Education	4.40	1.68	0	7

Table 2: Proportional Hazard Model Estimates

Variable	Exponential	Gompertz	Weibull
Constant	-2.46***	-2.50***	-2.44***
	(0.33)	(0.33)	(0.33)
Price	-0.80**	-0.73*	-0.85**
	(0.39)	(0.39)	(0.39)
Promotion Dummy	0.03**	0.03**	0.03**
	(0.01)	(0.01)	(0.01)
Family Size	0.05**	0.05**	0.05**
	(0.02)	(0.02)	(0.02)
Weekday Dummy	0.02*	0.02*	0.02*
	(0.01)	(0.01)	(0.01)
Season1	0.20***	0.19***	0.20***
	(0.02)	(0.02)	(0.02)
Season2	0.11***	0.10***	0.11***
	(0.02)	(0.02)	(0.02)
Season3	0.06**	0.06**	0.06**
	(0.02)	(0.02)	(0.02)
Log Likelihood	70791	70791	70791
AIC	-141564	-141562	-141562
BIC	-141487	-141477	-141477
Duration Elasticity			
Price	0.68**	0.62*	0.72**
	(0.33)	(0.33)	(0.33)

Note :
(1) Standard errors are in parentheses.
(2) *** $t-value > 5$; ** $t-value > 2$; * $t-value > 1$.
(3) The estimation are shared frailty model with Gamma distribution.

Table 3: Logit Model Estimates on Regular Drinks: Different Price Variables

	Mean Price per Pound	Median Price per Pound
Constant	1.65***	1.51***
	((0.09)	(0.11)
Regular Price	-2.31***	-2.48***
	(0.15)	(0.17)
Diet Price	1.34***	1.42***
	(0.17)	(0.17)
Male Education	-0.01**	-0.01**
	(0.00)	(0.00)
Female Education	-0.03**	-0.03***
	(0.01)	(0.01)
Family Size	0.31***	0.32***
	(0.01)	(0.01)
Income	-0.14***	-0.14***
	(0.00)	(0.00)
Log Likelihood	-36374	-36399
Elasticity		
Regular Price	-0.65***	0.60***
	(0.04)	(0.04)
Diet Price	0.35***	0.35***
	(0.04)	(0.04)

Note :
(1) Standard errors are in parentheses.
(2) *** $t - value > 5$; ** $t - value > 2$; * $t - value > 1$.

Table 4: Log Quantity Regression: Fixed and Random Models

Variables	Fixed Effects	Random Effects
Constant	1.35***	1.06***
	(0.01)	(0.08)
Log(Price)	-0.316***	-0.318***
	(0.011)	(0.011)
Coupon Dummy	0.30***	0.31***
	(0.01)	(0.01)
Male Education	-	-0.004
		(0.007)
Female Education	-	-0.04**
		(0.01)
Family Size	-	0.09***
		(0.01)
Income	-	0.04***
		(0.01)
Overall R-square	0.06	0.12

Note :
(1) Standard errors are in parentheses.
(2) *** $t - value > 5$; ** $t - value > 2$; * $t - value > 1$.

Table 5: Obesity Taxation Evaluation - Calories Intake

Percentile	5th	10th	15th	25th	50th	75th	85th	90th	95th
$RegularConsumption(Ounce)^1$	17.3	59.3	101.1	197.2	496.4	1167.5	1802.6	2294.2	3073.8
$ConsumedCalories/Year^2$	216.3	741.3	1263.8	2465	6205	14593.8	22532.5	28677.5	38422.5
$Universal1\%Tax^3$	2.81	9.64	16.43	32.00	80.67	189.72	292.92	372.81	499.49
$Regular1\%Tax^4$	3.03	10.38	17.69	34.51	86.87	204.31	315.46	401.49	537.92
$Universal5\%Tax^3$	13.84	47.44	80.88	157.76	397.12	934.00	1442.08	1835.36	2459.04
$Regular5\%Tax^4$	14.49	49.66	84.67	165.16	415.74	977.78	1509.68	1921.39	2574.31
$Universal10\%Tax^3$	27.03	92.66	157.97	308.13	775.63	1824.22	2816.56	3584.69	4802.81
$Regular10\%Tax^4$	28.55	97.85	166.82	325.38	819.06	1926.38	2974.29	3785.43	5071.77

Note :
(1) The individual consumption of regular soft drinks in ounces, defined as the amount purchased by household divided by the family size.
(2) The amount of individual intaking calories from regular soft drink consumption per year, calculated from the regular soft drink consumption.
(3) The changes of calories intaken from the taxation rate on the whole soft drink category.
(4) The changes of calories intaken from the taxation rate on the regular products only.

Table 6: Obesity Taxation Evaluation - Weight Reduction (Gram)

Weight Percentile	33th Overweight	66.2th Obese	94.3th Extremely Obese
Universal 1% Tax	6.81	20.75	71.46
Regular 1% Tax	7.34	22.35	76.96
Universal 5% Tax	33.55	102.16	351.82
Regular 5% Tax	35.12	106.95	368.32
Universal 10% Tax	65.52	199.53	687.16
Regular 10% Tax	69.19	210.70	725.64

Note :
(1) This table reports the weight reduction in grams per year at the regular soft drink consumption percentile corresponding to the percentiles of overweight and obese in the U.S..