Are you are what you eat?
Overweight Status and Soft Drink Choices

Rui Huang
Rigoberto A. Lopez
University of Connecticut

Selected Paper prepared for presentation at the Agricultural & Applied Economics Association's 2009 AAEA & ACCI Joint Annual Meeting, Milwaukee, WI, July 26-28, 2009.

Are you are what you eat? Overweight Status and Soft Drink Choices

Introduction

The wide-spreading obesity epidemic has become a severe public health crisis. Heavy soft drink consumption is among the major culprits. The principle of energy balance unmistakably states that the higher the calorie intake is, the higher risk of being overweight, other things being equal. As the average Americans become bigger, perception on what the normal body weight is changes. A shirt of size small today is larger than a shirt of same size sold ten years ago. Hospitals are ordering larger medical equipment and beds to accommodate heavier patients. An interesting question thus arises: does a social norm that casts a heavier weight in more favorable view influence consumption of caloric food? This is what this study attempts to partially answer. We empirically investigate whether social norms of body weight have any effect on soft drink consumption choices, and if so, what the size of the effects may be.

The question is important because it would shed light on how the obesity epidemic spreads in a society. The existence of such an effect means that the spread of obesity is accelerated by a "social multiplier effect". If the social norm factor is important in shaping individual choices, then policies aiming to change the social perception of being overweight would be effective.

Besides looking at whether social norm affects total soft drink consumption, we also try to examine whether social norms play a role on brand choices of soft drinks. Identifying the effects of social norm on brand choices is of interest to marketers of the soft drink industry, as the information will help them to decide what product attributes are valued more by consumers given the current social norm. Moreover, the industry is also able to influence social norms through advertising so their products become more attractive to consumers.

Literature

Spurred by the enormous concern about the accelerating obesity problem around the world, there has been a rapidly growing literature on the causes and consequences of the obesity epidemic. Most economic analyses so far have been focused on the fall of food prices and exercise. Philipson and Posner (2008) surveyed this literature and they pointed out that social aspects of obesity may have a multiplier effect on the growth of obesity. Burke and Heiland (2006) calibrated a theoretical dynamic model to explain the rising obesity with falling food prices, endogenous social body weight norms, and heterogeneous human metabolism. Elite (2007) utilized French survey data to examine how social norms, defined as self-reported ideal body weight, affects food attitudes.

We intend to contribute to this growing literature by investigating whether food consumption choices are affected by social norms. Our paper is closest in spirit to Elite (2007), but instead of self-reported food attitudes, we are focusing on whether social norm of body weight affects revealed food choices. To our knowledge, no previous studies have linked actual consumption data with social norm of body weight.

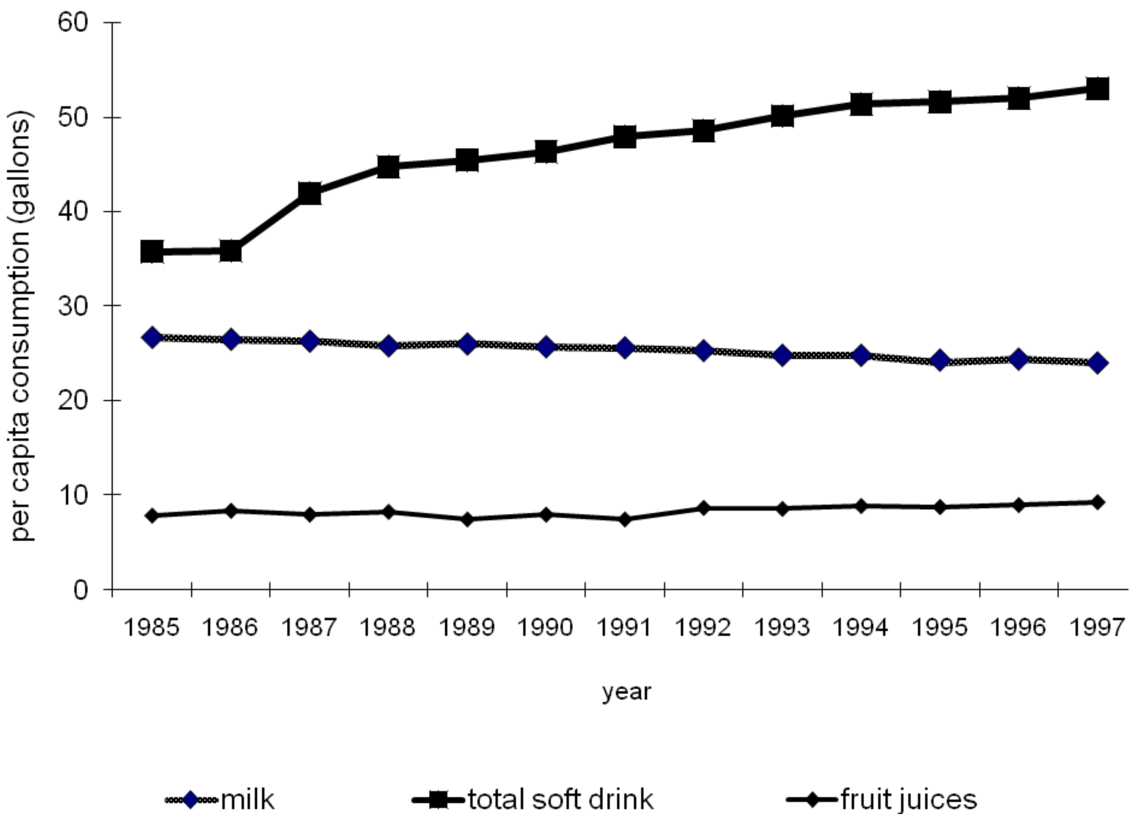
Data Sources

We have two primary data sources. One is an IRI scanner data set of the soft drink category. The IRI scanner data consist of quarterly information on volume sales, dollar sales, unit sales, product description, average price, percentage of volume sales with mechanizing, as well as market level basic demographic information such as population, median household income, and median household size. The data span from the first quarter of 1989 to the last quarter of 1992, a period which is in the midst of the rapid increase in obesity. The other data source is Center for Disease Control's Behavioral Risk Factor Surveillance System (BRFSS). The BRFSS is a large random cross-sectional sample of the resident population 18 years and older based on telephone survey conducted yearly in participating states. BRFSS contains data on self-reported weight and height, as well as other demographic characteristics of the surveyed individuals. We use the annual BRFSS data from 1989 to 1992 in our analysis. In addition, we collect product calorie information for the major brands in our IRI data from manufacturer's websites, and from nutrition information websites such as www.**nutrition**data.com .

Trends on Reference BMI and Soft Drink Consumption

Americans have been drinking more and more soft drinks during the past several decades. Figure 1 showed the per capita consumption of soft drink from 1985 to 1997, a period that includes our data period. Per capita consumption of regular soft drink far exceeds other healthier alternatives such as tea, milk, bottled water and fruit juice, more over, it has been increasing at a

Figure 1. Per Capita Beverage Consumption

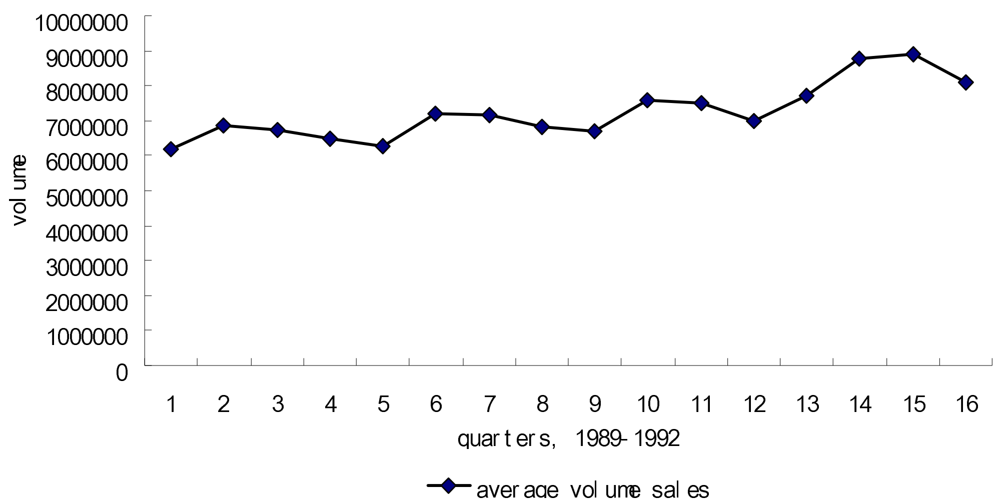


Data source: USDA ERS Report No. 965. Food Consumption, Prices, and Expenditures, 1970-97

much faster rate than any of the other alternatives, including diet soft drink.

Figure 2 shows the average quarterly volume sales of total soft drink across all the metropolitan areas in the IRI data. A similar upward trend, with cyclic seasonal fluctuations showing a peak of consumption in the summer.

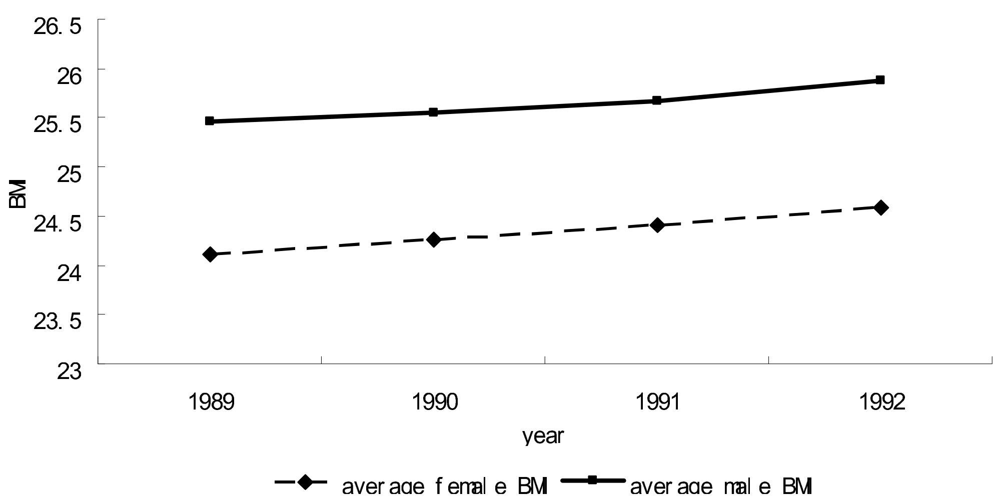
Figure 2. Average Soft Drink Sales in IRI Data



Source: our computation from IRI data

Americans are also getting heavier during the data period. Figure 3 shows the trends in average male and female BMI computed from the BRFSS data for the period of 1989 to 1992.

Figure 3. Average BMI across States



Source: our computation from BRFSS data.

The parallel rising trends in the per capita or total consumption of soft drink and in the average BMI suggest some quite strong relationship between soft drink consumption and average weights in a society. The most obvious relationship between the two is that an increase in soft drink consumption on the average leads to an increase of weight on the average. We argue that the relationship could potentially flow in the opposite direction. That is, when a person observes an increase in what the society she lives in regards as normal or acceptable in body weight, she becomes more likely to drink an additional can of soft drink because an increase in her own body weight has become more acceptable now. If there is indeed such a relationship between the social norm of body weight and soft drink consumption, then both the average body weight and average soft drink consumption will accelerate because an increase in one fuels the other.

In the following we tried to test whether a change in the social norm of body weight affects consumption of soft drink. We conduct two complementary analyses. The first analysis relies on reduced form regressions to see whether social norm of body weight influences per capita consumption of soft drink and that of regular soft drink. The second analysis turns to whether individual brand choices are affected by social norm of body weight.

Reduced Form Analysis

Our reduced form analysis investigates whether social norm of body weight influences soft drink consumption. We utilizes market level data compiled from our two data sets, IRI market level scanner data and BRFSS repeated cross-section survey data. A market is defined as a quarter-locality combination in the IRI data.

We specify the following regression equation:

$$y_m = X_m\beta + \gamma \cdot refBMI_m + \varepsilon_m \qquad (1)$$

Where y_m is an outcome variable in a market, X_m is a matrix of market specific control variables, such as price of soft drink on the average, the median household income and median household size in the market. Our variable of interest is $refBMI_m$, the reference or social norm of BMI in the market. We use the BMI averaged across all individuals in the BRFSS sample for a state in a given year as the reference BMI. Therefore, this measure is in essence an average of a relatively representative sample of adults of 18 years and older in a state for a given year. Finally, ε_m is an idiosyncratic error term that captures unobserved shocks to market level demand. We look at two outcome variables, one is per capita soft drink volume sales, and the other is per capita calorie contained in volume sales of soft drink in the market.

We obtain market level volume sales of total soft drink or regular soft drink by aggregating product level volume sales in the IRI scanner data. The soft drink category includes all carbonated beverages, such as cola, ginger ale and root beer. We simply include all products categorized as soft drink in the IRI data, when computing soft drink sales. Then per capita soft drink volume sales are defined as total volume sales across all products in a market divided by population in the market. For the other outcome variable, per capita calorie consumption, we only use a subset of all the products from IRI data because we do not have calorie information for all the products. We do not have calorie content for each of the product in our IRI data for three reasons. First, for some obscure or even discontinued brands, the information is not available. Second, some national brand products in the original IRI data, for instance, “all other Coca-cola products”, are already an aggregated product for which we cannot identify individual brand. Last but not the least, the IRI data do not distinguish private label products carried by different grocery chains and therefore we cannot identify any of the individual private label products either. Therefore, we decide to keep only the top 23 national brand products that collectively account for 74.8% of market share in the data when computing the average per capita consumption of calorie from soft drink. The average prices of regular soft drink are also based on this subset of products for which we can have calorie content information.

There are two issues with the market level reduced form analysis. First, the price could be endogenous in explaining per capita consumption. The average price of soft drink could be related to factors observed by the retailers in different markets but not observed in the data. For example, retailers could potentially use the soft drink category as a loss leader, that is, retailers intentionally lower the prices of soft drink products to attract more foot traffic in anticipated period of high demand. Moreover, there could some market level specific factors that are related with the prices of soft drink in that area. For example, if consumers in the South have a tradition of drinking more soft drink than consumers in other areas in the United States, manufacturers and retailers could follow certain zone pricing strategy and price their products differently in the South. We address the potential price endogeneity issue in two ways, namely, by using instrumental variables for price, and by adding market fixed effects that control for market specific unobserved demand shocks. We instrument the prices with input prices which are definitely correlated with prices, but which should not be directly correlated with unobserved demand shocks.

The other econometric issue is the endogeneity or reverse causality of the reference BMI in the market. Although the mechanism of BMI determination varies for each individual due to inter-person differences in physiological and genetic features, BMI is in general a function of the net energy intake, or energy intake from food subtracted with energy expenditure through metabolism and physical exercises. Therefore, more consumption of soft drink will undoubtedly contribute to weight gain, everything else being equal. In other words, an unobserved demand shock that affects consumer’s consumption of soft drink will also be correlated with the reference BMI, giving rise to the endogeneity issue. In principal, we can also use instrumental variables to address the endogeneity of the reference BMI. The instrumental variables should be correlated with how the reference BMI is determined in a society, but not with demand shocks that affect average soft drink consumption in this society. For example, genetic make-ups that have been scientifically proven to have an effect on the average BMI of a society, but which does not interfere with consumers’ tastes for soft drink can serve as a valid instrument. Lacking such information, we resort to two measures to address this issue. First, again, we try to add market specific fixed effects to control for market level demand shocks that also affect social norm of body weight in the market. For example, if a market consists of a majority of health conscientious consumers, we expect to see that the consumers in the specific market to pay more attention on their diet and on their fitness. Therefore, this market would exibit a lower average BMI, as well as a lower per capita consumption. The use of market fixed effects helps to isolate this market level unobserved healthy orientation out and partly deals with the endogeneity issue. Second, instead of using the current average BMI as the reference BMI, we use the average BMI in the last year as the reference BMI. The idea is that an individual consumer adapts her perceived social norm of BMI at the beginning of a year, given what she observes in the past year. The lagged average BMI is regarded as predetermined. Hence the lagged average BMI should not be correlated with the unobserved demand shocks that influence current soft drink demand. We acknowledge that the use of lagged BMI is not a perfect solution to the endogeneity problem however we are limited by data availability.

We report the summary statistics of the variables for the reduced form analyses in Table 1, and the results from the two reduced-form analyses in Table 2 and Table 3. For the two different outcome variables, per capita soft drink volume and per capita calories from caloric soft drink, we report results from four specifications, namely, OLS and instrumental variable regressions, with and without locality fixed effects. Our variable of interest is the lagged reference BMI, a proxy for social norm. We also include other market level control variables including quantity weighted average prices of all soft drink or of caloric soft drink, median household income and median household size.

For per capita soft drink sales, lagged reference BMI is positive and significant at 0.001 levels for all four specifications. Particularly, in the fixed effect instrumental variable regression, if reference BMI goes up by 1, then the per capita consumption increases by 0.275 units. The average per capita consumption in the data is just 2.87 unit. Therefore, the impact of reference BMI is relatively large. For per capita calorie consumption, lagged reference BMI is positive and significant at 0.01 level for all but one specification, suggesting that an increase in reference BMI will lead to increased caloric intake from soft drink. In sum, these results support that an increase in BMI has considerable and positive impact on total soft drink consumption, as well as total caloric soft drink consumption.

Brand-level Analysis

The reduced form analyses provide supportive evidence that a heavier socially normal body weight, as captured by an increased reference BMI, leads to more consumption of soft drink overall and regular soft drink. In a second part of the empirical analysis, we further investigate whether a change in reference BMI has any impacts on brand choices of soft drink. That is, does a higher reference BMI lead to a higher evaluation of calorie content of differentiated soft drink products?

An ideal data set suitable for the brand level analysis should contain information on consumer level soft drink purchases as well as consumer level perception of social norm. For identification purposes, the social norm should also be exogenous to soft drink demand and brand choices. Unfortunately, we do not have such data. Instead, we have an aggregated market level scanner data with quantities and prices for each brand in a market. We also have another data set with consumer demographical information. We resort to simulation techniques to relate the consumer level information in the latter data set to the market level brand shares provided by the former data set. This kind of data-augmenting techniques was first promoted by Berry (1994) and Berry, Levinsohn and Pakes (1995) where consumer level demographic information from Current Population Survey are used in estimating brand level demand for differentiated products from aggregated data.

We model consumer brand choices with a random-coefficient logit model, following Berry, Levinsohn and Pakes (1995) and Nevo (2001). In this model, a consumer decides to choose one of the products available in the market. She can also choose to purchase none of the available brands, in this case, she is said to choose an outside alternative. The utility a consumer i derives from product j in market m is given by

$$u_{ijm} = -\alpha_i p_{jm} + X_{jm}\beta_j + \varepsilon_{ijm} \qquad (2)$$

Where p_{jm} is the price of product j in market m, X_{jm} is a matrix of product characteristics of the product, such as calorie content and brands, and ε_{ijm} is an idiosyncratic error term. The coefficients vary across individuals according to observed and unobserved individual characteristics,

$$\begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} + \pi D_i + \Sigma v_i \qquad (3)$$

Where α, β are average values of the coefficients, D_i is a matrix of observed individual characteristics, such as age and gender, and v_i is an unobserved consumer level taste shock.

Each consumer purchases whichever product among all the available products that gives her the highest utility. If the highest utility is less than the utility arisen from choosing the outside alternative, which is normalized to zero in our model, the consumer will not make any purchases. The market share of product j in a market is therefore the aggregated consumer choices in that market.

$$s_{jm} = \int_{A_{jm}} dF(D)dG(v) \quad (4)$$

Where A_{jm} is the set of consumers who chooses product j over all other alternatives in the market m, and F(.), G(.) are the distribution of observed and unobserved consumer characteristics in the population in this market.

The demand model is then estimated using Generalized Method of Moments. In our model, calorie content is one of the product characteristics that enter the utility a consumer enjoys from the products. The individual taste of a specific consumer for calorie content, we assume, is a function of the mean evaluation for calorie content among consumers, her own weight status, and the social norm of body weight she perceives. Implicitly, we are looking at a more general utility-maximizing problem for the consumer. The problem she faces is to decide energy intake to maximize her utility, which depends on three components. The first of the components is the net pleasure she derives from the caloric food, the second component is health benefit of her own weight status, and the third is a social benefit of how closely she conforms to the perceived social norm. Energy intake positively affects one's own body weight, which in turn affects the last two components. Without functional form assumptions for the utility function, the health benefit and the social benefit function, it is not clear a prior what the relationship between the utility-maximizing caloric intake and one's actual and reference body weight is in such a general analytical model. Thus, the goal of our empirical exercise is to examine whether the observed choices reveals something about the possible relationship.

The reference BMI in the brand-level analysis is defined on a finer scale than in the market level analysis because we can exploit the individual level data here. We define the reference BMI as the median BMI of the reference group a consumer is in. We assume the reference group which a consumer identifies with has the some of the same social-demographic attributes she possesses. For example, we let the reference group to be the group of individuals in the BRFSS data who are in the same state/race/age category/income category. Then for each year, we compute the median BMI for each of the reference group. The approach of defining reference BMI can be problematic for two reasons. First, an individual's reference group is not necessarily the one who have identical social-demographic attributes as she does. For example, a middle-class teenager girl might perceive the body shape and body weight of models as the social norm. It is also possible that a lot of individuals perceive the group of higher income to represent the social norm. A second, more fundamental problem with this definition of reference BMI is that it might reflect the "context effect" rather than the "social interaction effect". The context effect arises when all the people in a same group are exposed to a same environmental factor that contributes to demand for caloric intake. For example, if consumers of lower income tend to live in urban areas where grocery stores do not carry low calorie content sodas, then they tend to buy more caloric sodas than other consumers, and as a result, they become heavier and the reference BMI of this group rises. Lacking better ways to address the problem, we experiment with different specifications of reference BMI to see whether the results are sensitive to the specifications.

We report the results from the random-coefficient model in Table 4. All the coefficients of the mean evaluations are statistically significant at 0.01, but none of the random parts of the coefficients are statistically significant, suggesting we might have some identification issues for the random coefficients. The calorie coefficient is negative, suggesting that after controlling for prices and brands, there is a general dislike of calorie content in the population.

Conclusion

Understanding how obesity spread in a society is important in devising effective policies to combat the obesity epidemic. We investigated empirically whether social norm of body weight affects food consumption choices using data from soft drink category. We found strong evidence that overall consumption of soft drink and calorie intake from soft drink increases with a rising social norm of obesity. The finding suggests of social multiplier effect in the spread of obesity around the world. We also examined whether social norm of body weight affects consumers evaluation for calorie content in a brand-level demand model. We failed to find any statistically significant effects from our preliminary results, which might reflect not a lack of relationship but problematic definition of social norm in the implementation. In the next step, we will try to develop a unified analytical framework of calorie intake decision and brand choices, and refine our measures for social norm to tackle the identification issues. Any comments and suggestions are greatly welcomed.

References

Berry, S., J. Levinsohn, and A. Pakes, (1995), Automobile Prices in Market Equilibrium, *Econometrica*, Vol. 63, 841-890.

Burke,M., and Heiland, F. (2006), Social Dynamics of Obesity, FRB of Boston Public Policy Discussion Paper No. 06-5.

Etile, F (2007) Social Norms, Ideal Body Weight, and Food Attitudes, *Health Economics*, Vol. 16: 945-966

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

Table 1. Summary statistics for reduced form regressions

Variable	Obs	Mean	Std. Dev.	Min	Max
per capital soft drink volume	864	2.87	0.74	0.97	7.32
average price of soft drink	864	3.50	0.33	2.63	4.49
per capita calorie from regular soft drink	864	19.12	4.71	5.89	41.59
average price of regular soft drink price	864	3.65	0.31	2.82	4.55
volume share of diet soft drink	864	0.07	0.02	0.02	0.14
lagged reference BMI	864	24.79	0.33	23.63	25.51
population	864	2722645	2828788	307181	15700000
median income	864	33330	7302	18086	53429
median household size	864	2.60	0.14	2.30	3.20
no. of localities	63				
no. of quarters	16				
no. of markets	864				

Table 2. Reduced form regressions of per capita soft drink volume

Dependent variable: Per capita volume of all soft drink sales	OLS	IV	OLS	IV
average price of soft drink	-0.903 (0.067)	-2.668 (0.287)	-1.481 (0.085)	-1.674 (0.226)
lagged reference BMI	0.465 (0.070)	0.209 (0.140)	0.319 (0.079)	0.275 (0.093)
median household income	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
median household size	-0.348 (0.146)	-0.082 (0.201)	-0.881 (0.271)	-0.858 (0.273)
constant	-5.164 (1.896)	16.976 (4.272)	1.930 (2.194)	3.705 (2.926)
city fixed effects			Y	Y
Adjusted R-square	0.348	NA	0.3331	NA
F-statistics		60.08		
Wald				56618
first stage F for IV regression		38.43		62.78
no. obs	864	864	864	864

Instrumental variables for price variables: wholesale prices of sugar and corn sweeteners, gasoline prices, average state wages, federal fund rates and commercial paper interests.
Numbers in parenthesis are standard errors.

Table 3. Reduced form regressions of per capita calories from caloric soft drink

Per capita calories from caloric soft drink sales	OLS	IV	OLS	IV
average price of caloric soft drink	-5.785 (0.491)	-13.357 (1.645)	-6.204 (0.495)	-10.860 (1.489)
lagged reference BMI	2.123 (0.479)	0.070 (0.685)	1.133 (0.450)	0.814 (0.484)
median household income	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
median household size	0.935 (1.031)	0.967 (1.166)	-4.862 (1.583)	-4.560 (1.671)
constant	-15.604 (13.069)	62.132 (21.697)	24.934 (12.289)	49.711 (14.929)
city fixed effects			Y	Y
R square	0.1915		0.1678	
F-statistics		30.10		
Wald				65824
first stage F		29.84		45.71
no. obs	864	864	864	864

Instrumental variables for price variables: wholesale prices of sugar and corn sweeteners, gasoline prices, average state wages, federal fund rates and commercial paper interests.

Numbers in parenthesis are standard errors.

Table 4. Random-coefficient brand-level demand results

variables	mean	unobserved	reference BMI	obese dummy
constant	-7.224 (0.144)	1.4732	-0.159	-0.0589
price	-3.421 (0.356)			
calorie	-0.056 (0.021)	0.0927	-0.053	-0.5376
coke	1.427 (0.048)			
pepsi	1.177 (0.037)			
hickha	0.865 (0.032)			
no. obs Obj. function	11448 431.6999			