



Original Contribution

A Dynamic Panel Model of the Associations of Sweetened Beverage Purchases With Dietary Quality and Food-Purchasing Patterns

Carmen Piernas, Shu Wen Ng, Michelle A. Mendez, Penny Gordon-Larsen, and Barry M. Popkin*

* Correspondence to Dr. Barry M. Popkin, Carolina Population Center, University of North Carolina, 123 W. Franklin Street, Chapel Hill, NC 27516 (e-mail: popkin@unc.edu).

Initially submitted November 29, 2013; accepted for publication October 14, 2014.

Investigating the association between consumption of sweetened beverages and dietary quality is challenging because issues such as reverse causality and unmeasured confounding might result in biased and inconsistent estimates. Using a dynamic panel model with instrumental variables to address those issues, we examined the independent associations of beverages sweetened with caloric and low-calorie sweeteners with dietary quality and food-purchasing patterns. We analyzed purchase data from the Homescan survey, an ongoing, longitudinal, nationally representative US survey, from 2000 to 2010 ($n = 34,294$). Our model included lagged measures of dietary quality and beverage purchases (servings/day in the previous year) as exposures to predict the outcomes (macronutrient (kilocalories per capita per day; %), total energy, and food purchases) in the next year after adjustment for other sociodemographic covariates. Despite secular declines in purchases (kilocalories per capita per day) from all sources, each 1-serving/day increase in consumption of either beverage type resulted in higher purchases of total daily kilocalories and kilocalories from food, carbohydrates, total sugar, and total fat. Each 1-serving/day increase in consumption of either beverage was associated with more purchases of caloric-sweetened desserts or sweeteners, which accounted for a substantial proportion of the increase in total kilocalories. We concluded that consumers of both beverages sweetened with low-calorie sweeteners and beverages sweetened with caloric sweeteners had poorer dietary quality, exhibited higher energy from all purchases, sugar, and fat, and purchased more caloric-sweetened desserts/caloric sweeteners compared with nonconsumers.

beverages; caloric sweeteners; dietary quality; low-calorie sweeteners; purchases

Abbreviations: CS, caloric sweetener; IV, instrumental variable; LCS, low-calorie sweetener.

Consumption beverages flavored with caloric sweeteners (CS) (hereafter referred to as CS beverages) has been associated with higher caloric intake and adverse health outcomes (1, 2). However, whether the same associations exist for beverages flavored with low-calorie sweeteners (LCS) (hereafter referred to as LCS beverages) remains unclear (3–7). Overall, the associations with these beverage types and dietary quality and food-purchasing patterns warrant further examination.

LCS are defined as food additives that provide fewer than 3.8 kilocalories per gram and/or are used in quantities low enough that the amount of calories they provide is negligible. All other sweeteners that provide 3.8 kilocalories per gram or more are considered CS because this cutpoint reflects the caloric value of a gram of carbohydrate.

Investigating the relationships of consumption of CS and LCS beverages with dietary quality is challenging because it is difficult to ascertain whether a particular dietary pattern is linked to a specific beverage-consumption pattern or vice versa. Also, unobserved common factors (i.e., obesity, diabetes, personal preferences) could affect both beverage-consumption and dietary patterns. Such effects, which were traditionally known as reverse causality or unmeasured confounding, could create endogeneity issues (8, 9), which could result in biased and inconsistent estimates if the aforementioned issues are not adequately addressed (10). In the context of observational data and a lack of randomly assigned exposures, an approach to investigate the causal association between CS and LCS beverages and dietary quality in the presence of endogeneity is

instrumental variable (IV) analysis (11, 12). IVs are used to correct for bias caused by endogenous exposures and include variables that are directly associated with the endogenous explanatory variables but are not directly associated with the outcomes of interest (8). Furthermore, IVs must be exogenous (uncorrelated with the unobserved error terms) in the model.

In the present study, we implemented a dynamic panel model using data on yearly purchases by individuals included in Homescan from 2000 to 2010 (Nielsen Consumer Panel and Retail Measurement, The Nielsen Co.; <http://www.nielsen.com>) to investigate the relationships of consumption of CS and LCS beverages with dietary quality. To better establish temporality and causal relationships, this approach allowed the outcome (contemporaneous dietary quality) to depend on lagged dietary quality and beverage consumption in the previous year, and we included several market-level IVs that were selected on the basis of theoretical validity and specification tests.

METHODS

Study population

We analyzed purchasing data from the 2000 to 2010 Homescan Consumer Panel (13), an ongoing, longitudinal, nationally representative US survey that captures household purchases of more than 600,000 barcoded store products (14). To better characterize individuals' diets, we selected data that represented single-person adult households ($n = 136,011$ observations from 34,294 individuals) (Table 1). Overall, kilocalories captured in Homescan represented approximately two-thirds of the total caloric intake (15).

Food groups

Information from nutrition-fact panel on total kilocalories and kilocalories from macronutrients (carbohydrates, total sugar, total fat, protein, and saturated fat) and from ingredient lists were linked to each barcoded product (14, 15). Using keyword searches for CS and LCS, we used information from ingredient lists to categorize all foods and beverages with sweeteners. All foods and beverages purchased by Homescan participants were grouped into 1 of 9 beverage and 14 food groups (Web Table 1).

Statistical analysis

Descriptive statistics. Detailed information about the study population, food groups, demographic characteristics,

and modeling approach is provided in the Web Appendix 1 (available at <http://aje.oxfordjournals.org/>). Households included in Homescan reported sociodemographic characteristics, including sex, age, income (percentage of the poverty level), educational level, and race/ethnicity (Hispanic, non-Hispanic white, non-Hispanic black, and other).

Outcome specification. For primary outcomes, we used continuous measures (kilocalories per capita per day) of dietary quality: total energy from all purchases, total energy excluding LCS and CS beverages, total energy from beverages, total energy from foods, and total energy and percentage of energy from macronutrients. For secondary outcomes, we used continuous measures (kilocalories per capita per day) of purchases of other foods and beverages groups based on purchase data. We used yearly measures of purchases to estimate daily purchases.

Exposure specification. Continuous measures of LCS and CS beverage purchases (servings/day) were included as main exposures. We obtained estimates by dividing the total volume (in milliliters) of beverages purchased per day by the standard serving size of a can (355 mL).

Dynamic panel model. Endogeneity and IVs. Endogenous explanatory variables were correlated with the error term in the model, a problem that might have been due to reverse causality or unmeasured confounding and that will lead to biased and inconsistent measures of associations (8–10). Consumption of LCS and CS beverages, as well as dietary patterns, might be potentially endogenous variables because these are “choice variables” to the consumer (e.g., a household might choose to consume a specific type of beverage and diet). If so, our model would have been subject to endogeneity bias if a particular dietary pattern was linked to a particular beverage-consumption pattern or vice versa or if there were unobserved common factors (e.g., obesity, diabetes, or individual preferences) that affected both beverage-consumption and dietary patterns.

IVs can be used to control for endogeneity. Reliable IVs should be theoretically justified and statistically associated with the endogenous explanatory variables in the model, conditional on the other covariates, but have no direct association with the outcome (other than through the endogenous variables). In addition, IVs should be exogenous (uncorrelated with the error term in the main equation) (8). Several potential market-level IVs (including food/beverage prices and purchases; percentage of market sales that LCS and CS beverages comprise; and average number of shopping trips per year) were considered suitable a priori because theoretically, these variables would be associated with household consumption of LCS and CS beverages under the assumption

Table 1. Sample Sizes of Single-Person Adult Households That Comprised the Study Population in Each Year, Homescan Survey, 2000–2010

Study Sample	Sample Size by Study Year											Total
	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	
Households in ≥ 2 consecutive waves	N/A	6,595	7,073	7,817	8,004	8,847	10,518	11,727	12,122	12,032	11,632	96,367
Households in 1 study wave	8,508	2,335	3,051	2,502	2,379	4,129	3,983	3,277	3,050	2,838	3,592	39,644
Total sample	8,508	8,930	10,124	10,319	10,383	12,976	14,501	15,004	15,172	14,870	15,224	136,011

that household's environment affects behavior. In addition, these market-level variables are outside the control of the individual, so they would have an indirect effect on dietary behavior that is mediated through its association with LCS and CS beverage consumption in the model. Because of this, these market-level IVs could be assumed to be exogenous and not correlated with the unobserved error terms (10). To further test the statistical strength of IVs, we analyzed the association between our endogenous explanatory variables, outcomes, and proposed IVs.

Empirical model. To better establish temporality and causality, we created a dynamic panel model that related diet in the current year to its own lagged value in the previous year, along with lagged measured LCS and CS beverage consumption:

$$D_{it} = \alpha D_{i,t-1} + \beta B_{i,t-1} + \gamma X_i + \pi Z_{it} + \mu_i + \varepsilon_{it}, \quad (1)$$

where D_{it} is diet in the current wave, $D_{i,t-1}$ is diet in the prior wave, $B_{i,t-1}$ represents continuous lagged values of beverage consumption (servings/day of LCS and CS beverages), X_i is the vector of time invariant covariates (i.e., sex, race), Z_{it} represents other time-varying control variables (i.e., age, educational level, and income), i equals 1, . . . , N individuals, t equals 1, . . . , T years, and α , β , γ , π are vectors of coefficients for the explanatory variables. The error terms were μ_i (unobserved time-invariant individual characteristics) and ε_{it} (time-varying error term).

There were potential challenges to account for: 1) endogeneity (i.e., correlation between explanatory variables and μ_i); 2) double endogeneity (i.e., correlation between explanatory variables and ε_{it}); and 3) autocorrelation (i.e., serially correlated ε_{it} and μ_i for the same individual due to the time invariant unobserved heterogeneity, which will result in incorrect standard errors). At minimum, we expected to find that lagged diet was correlated with μ_i , so we used IVs to account for it. One option was to calculate a first difference equation so that μ_i and other time invariant covariates were dropped:

$$\Delta D_{it} = \alpha [\Delta D_{i,t-1}] + \beta [\Delta B_{i,t-1}] + \pi Z_{it} + \Delta \varepsilon_{it} \quad (2)$$

The estimation method used was the generalized method of moments estimator developed by Blundell and Bond (16, 17). To our knowledge, this generalized method of moments system is more efficient than other approaches because it estimates equations 1 and 2 simultaneously. The system uses a large set of moment conditions and simultaneously includes 2 transformations of the main model, the regression in differences (equation 2) and the regression in levels (equation 1). In equation 2, the μ_i and other time-invariant observed variables were dropped. However, additional IVs should be used to account for the potential correlation between the explanatory variables and ε_{it} . We used lagged second and third differences of the endogenous explanatory variables (i.e., $\Delta B_{i,t-2}$ and $\Delta B_{i,t-3}$) as IVs in equation 2 under the assumption that there is no second-order autocorrelation (i.e., $\Delta B_{i,t-2}$ is correlated with $\Delta B_{i,t-1}$ but not correlated to $\Delta \varepsilon_{it}$), so that the lagged differences serve as valid IVs. Because these explanatory variables are time-varying, each successive wave added additional instruments. As was previously discussed, it appeared highly

likely that the endogenous explanatory variables would be correlated with μ_i in equation 1 and with ε_{it} in equation 2. Then, market-level IVs (prices, number of shopping trips, and percentage of market sales) that were found to be valid IVs were additionally used as lagged IVs for the explanatory endogenous variables in both equations 1 and 2.

Specification tests. We used the Sargan-Hansen J test of overidentifying restrictions to test the assumption that, given that there was more than 1 IV available and at least 1 of them was theoretically exogenous and helped predict the endogenous explanatory variable, our IVs were exogenous and as a result were not correlated with the error terms (18). Our main limitation was that the test could not identify whether all IVs used in the main models were truly exogenous; the reliability of the test is contingent upon at least 1 valid IV having been adequately identified and justified a priori. Failure to reject the null hypothesis of overidentification ($P > 0.05$) indicated that the assumption made about exogeneity of the IVs was valid. Secondly, we performed the Arellano-Bond test to investigate whether there was a second-order autocorrelation in equation 2, which would invalidate the use of lagged values of the endogenous variables as IVs in equation 2 (18). Although first-order autocorrelation was expected because of the first differencing, failure to reject the null hypothesis of no second-order autocorrelation ($P > 0.05$) would validate the use of the second and third lags of our explanatory variables as IVs.

Final model. We performed all analyses using Stata, version 12 (StataCorp LP, College Station, Texas). Survey commands accounted for design and weighting to generate nationally representative descriptive results. Our 2-step dynamic panel model with generalized method of moments estimator included lagged measures of the dependent variable (such as total energy), lagged exposures (LCS and CS beverage purchases), IVs (lagged market-level variables; second- and third-year lags of the main exposures), and confounders (age, sex, educational level, race/ethnicity, income, and year). We controlled for these because they were found to be differentially associated with consumption of LCS and CS beverages over this period of time (19). Estimates were presented as β coefficients and means with their respective standard errors. We obtained robust standard errors from dynamic panel models. The β coefficients for the main exposures in the dynamic panel models can be interpreted as the increase in the outcome variable for every additional serving per day of LCS or CS beverages in the previous year. We used model parameters to predict the mean energy from each type of beverage purchased. For example, we specified an increase of 1 serving/day of LCS beverages but 0 servings/day of CS beverages for consumers of LCS beverages and vice-versa for consumers of CS beverages. For nonconsumers, we specified 0 servings/day of LCS and CS beverages. Finally, we provided a comparison model that assumed exogeneity of the main exposures; it illustrates how failure to correct endogeneity can affect the findings. Linear trends were tested using Wald tests. A 2-sided P value of < 0.05 was set to denote statistical significance.

RESULTS

From 2000 to 2010, single-person households from Homescan comprised predominantly non-Hispanic white

Table 2. Beverage Consumption and Demographic Characteristics of Single-Person Adult Households That Comprised the Study Population, Homescan Survey, 2000–2010

Characteristic	Percentage of the Study Population by Year											P Value ^a
	2000 (n = 8,508)	2001 (n = 8,930)	2002 (n = 10,124)	2003 (n = 10,319)	2004 (n = 10,383)	2005 (n = 12,976)	2006 (n = 14,501)	2007 (n = 15,004)	2008 (n = 15,172)	2009 (n = 14,870)	2010 (n = 15,224)	
Consumption, servings/week ^b												
LCS beverages	2.4 (0.1)	2.4 (0.1)	2.6 (0.1)	2.8 (0.1)	2.9 (0.1)	3.0 (0.1)	2.9 (0.1)	2.7 (0.1)	2.7 (0.1)	2.6 (0.1)	2.7 (0.1)	0.034
CS beverages	3.1 (0.1)	3.1 (0.1)	3.0 (0.1)	3.0 (0.1)	2.7 (0.1)	2.7 (0.1)	2.8 (0.1)	2.6 (0.1)	2.7 (0.1)	2.6 (0.1)	2.5 (0.1)	<0.001
Age ^b	57.5 (0.3)	57.4 (0.2)	57.5 (0.2)	58.4 (0.2)	58.6 (0.2)	57.2 (0.2)	57.3 (0.2)	57.5 (0.2)	57.1 (0.2)	56.8 (0.2)	56.5 (0.2)	<0.001
Male sex	49.8	50.5	51.4	49.0	49.6	47.7	49.2	49.8	50.2	50.3	51.4	0.461
Race/ethnicity												
Non-Hispanic white	86.1	84.1	84.9	84.9	83.3	82.5	81.6	81.6	80.1	79.9	78.2	<0.001
Non-Hispanic black	10.6	11.1	10.5	10.5	11.3	11.1	11.7	11.5	12.3	11.9	12.5	<0.001
Hispanic	2.1	2.9	2.6	2.7	3.2	3.8	4.1	4.3	4.6	4.9	5.3	<0.001
Other	1.1	1.8	2.0	1.9	2.3	2.7	2.7	2.7	3.0	3.4	4.1	<0.001
Educational level												
High school diploma or less	32.6	32.8	29.7	33.0	32.6	36.9	35.4	35.1	34.9	34.0	34.5	<0.001
At least some college	67.4	67.2	70.3	67.0	67.4	63.1	64.6	64.9	65.1	66.0	65.5	<0.001
Income ^c												
<185%	23.8	22.7	22.7	22.3	22.3	33.3	32.9	32.7	31.5	30.0	29.1	<0.001
185%–400%	43.9	43.3	42.4	37.2	37.8	31.4	31.3	30.0	29.7	34.1	34.6	<0.001
>400%	32.2	34.0	34.9	40.5	39.9	35.3	35.8	37.3	38.8	35.9	36.3	<0.001

Abbreviations: CS beverages, beverages flavored with caloric sweeteners; LCS beverages, beverages flavored with low-calorie sweeteners.

^a P for linear trend; Wald test $P < 0.05$.

^b Values are expressed as mean (standard deviation), with sample weights used to account for selection probability and sampling design.

^c Percentage of the poverty level.

Table 3. Associations Between Lagged Market-Level Instrumental Variables and Lagged Outcomes and Exposures in Single-Person Adult Households That Comprised the Study Population, Homescan Survey, 2000–2010^a

Lagged Instrumental Variable ^b	Lagged Endogenous Exposures ^b in the Dynamic Panel Models, servings/day		Lagged Outcome Variables ^b in the Dynamic Panel Models, kcal/day			
	LCS Beverages	CS Beverages	Total Energy	Total Energy Excluding LCS/CS Beverages	Total Beverages Excluding LCS/CS Beverages	Total Energy From Food
Average prices, \$						
Food price index per 100 g	−0.00090 (0.00087)	−0.00080 (0.00077)	−1.46138 (0.99770)	−1.37329 (0.96991)	−0.75704 (0.22817) ^c	−0.61481 (0.89496)
LCS beverage price per 100 mL	0.52942 (0.34428)	0.14600 (0.30687)	333.84320 (395.64580)	219.28940 (384.62350)	109.74910 (90.48177)	110.63400 (354.90260)
CS beverage price per 100 mL	−0.29223 (0.96174)	0.61266 (0.85719)	1506.79700 (1105.16800)	1167.58100 (1074.37500)	657.44210 (252.73670) ^c	504.60710 (991.35250)
<i>P</i> value ^{d,e}	0.150	0.777	0.504	0.569	0.011	0.924
Average household purchases/year, \$						
Total food	−0.00004 (0.00003)	−0.00010 (0.00003) ^c	0.08474 (0.03846) ^c	0.09587 (0.03738) ^c	−0.01517 (0.00879)	0.11114 (0.03449) ^c
Total beverages	−0.00022 (0.00011) ^c	0.00012 (0.00010)	−0.22801 (0.12904)	−0.21344 (0.12542)	0.11441 (0.02947) ^c	−0.32580 (0.11572) ^c
Total LCS/CS beverages	0.00117 (0.00025) ^c	0.00047 (0.00022) ^c	0.14505 (0.28738)	0.12068 (0.27935)	−0.18146 (0.06566) ^c	0.30224 (0.25774)
<i>P</i> value ^{d,e}	0.001	0.006	0.086	0.034	0.004	0.001
Average no. of household grocery trips/year	0.00126 (0.00067)	−0.00019 (0.00060)	3.46010 (0.76735) ^c	3.44305 (0.74592) ^c	1.03024 (0.17536) ^c	
<i>P</i> value ^{d,f}	0.059	0.755	<0.001	<0.001	<0.001	0.005
% of market sales						
LCS beverage purchases	0.00565 (0.00108) ^c	−0.00064 (0.00096)	0.53795 (1.24340)	0.69859 (1.20873)	−0.17449 (0.28431)	0.87125 (1.11531)
CS beverage purchases	0.00071 (0.00110)	0.00261 (0.00098) ^c	0.79737 (1.26644)	0.51711 (1.23114)	0.44346 (0.28957)	0.07739 (1.13599)
<i>P</i> value ^{d,g}	<0.001	<0.001	0.820	0.846	0.015	0.575
<i>P</i> value ^{h,i}	<0.001	0.003	0.002	<0.001	<0.001	0.002

Abbreviations: CS beverages, beverages flavored with caloric sweeteners; LCS beverages, beverages flavored with low-calorie sweeteners.

^a Values are expressed as β (standard error) and were obtained from longitudinal random-effects models that were adjusted for year, market, sex, age, race/ethnicity, educational level, and income.

^b The first lagged value was used.

^c Significant at $P < 0.05$.

^d *P* value for the χ^2 joint test of significance for each group of instrumental variables.

^e Three degrees of freedom.

^f One degree of freedom.

^g Two degrees of freedom.

^h *P* value for the χ^2 joint test of significance for all instrumental variables.

ⁱ Nine degrees of freedom.

Table 4. Dynamic Panel Modeling of the Associations of 1-Serving per Day Increases in the Consumption of Beverages Sweetened With Caloric Sweeteners and Low-Calorie Sweeteners With Dietary Quality and Macronutrients in Single-Person Adult Households That Comprised the Study Population, Homescan Survey, 2000–2010^a

Outcome at Time <i>t</i>	Lagged Endogenous Explanatory Variables ^b			Sargan-Hansen Test ^c <i>P</i> Value	Arellano-Bond Test of Autocorrelation ^d	
	Outcome	LCS Beverages	CS Beverages		First Order <i>P</i> Value	Second Order <i>P</i> Value
Energy, kilocalories per capita per day						
Total energy	0.39 (0.18) ^e	86.01 (29.61) ^e	112.95 (55.31) ^e	0.513	<0.001	0.161
Total energy excluding LCS/CS beverages	0.31 (0.18)	92.51 (29.24) ^e	73.03 (37.23) ^e	0.383	<0.001	0.460
Total energy from food	0.23 (0.15)	99.41 (27.96) ^e	84.59 (32.68) ^e	0.468	<0.001	0.981
Total energy from all beverages	0.53 (0.22) ^e	−3.54 (7.20)	23.58 (32.14)	0.142	<0.001	0.184
Total energy from beverages excluding LCS/CS	0.74 (0.11) ^e	−2.17 (4.77)	−3.24 (5.21)	0.590	<0.001	0.379
Total daily macronutrients from all purchases, kilocalories per capita per day						
Carbohydrates	0.34 (0.17) ^e	42.29 (15.91) ^e	85.94 (38.29) ^e	0.465	<0.001	0.263
Sugar	0.26 (0.20)	19.41 (9.65) ^e	80.38 (35.88) ^e	0.786	<0.001	0.561
Protein	0.37 (0.17) ^e	10.46 (5.15) ^e	8.88 (5.06)	0.866	<0.001	0.179
Total fat	0.25 (0.16)	45.41 (14.01) ^e	38.54 (17.31) ^e	0.592	<0.001	0.718
Saturated fat	0.37 (0.18) ^e	14.10 (5.57) ^e	11.01 (6.51)	0.650	<0.001	0.330
Total daily macronutrients, %						
Carbohydrates	0.42 (0.16) ^e	−0.39 (0.28)	0.41 (0.44)	0.069	0.036	0.610
Sugar	0.54 (0.18) ^e	−0.24 (0.29)	0.32 (0.59)	0.065	<0.001	0.067
Protein	0.19 (0.13)	0.34 (0.23)	0.18 (0.21)	0.631	0.060	0.857
Total fat	0.69 (0.09) ^e	0.24 (0.22)	0.13 (0.23)	0.150	<0.001	<0.001
Saturated fat	0.62 (0.12) ^e	0.15 (0.11)	0.15 (0.10)	0.531	<0.001	0.033

Abbreviations: CS beverages, beverages flavored with caloric sweeteners; LCS beverages, beverages flavored with low-calorie sweeteners.

^a Values are expressed as β (standard error) and were obtained from a generalized method of moments 2-step system dynamic panel model with average no. of household grocery trips per year and percentage market sales made up of LCS beverages and CS beverages (specified for the level equation and differenced equation) as the lagged instrumental variables in the first step and the second and third lags of LCS and CS beverage purchases (specified for the differenced equation) as the lagged instrumental variables in the second step. There were 41 instrumental variables in total. The models were adjusted for age, sex, educational level, race/ethnicity, income, and year.

^b The first lagged value was used.

^c Sargan-Hansen test of overidentifying restrictions. If $P > 0.05$, the null hypothesis of overidentification indicated that the assumptions made about exogeneity of the instrumental variables were valid.

^d Arellano-Bond test of autocorrelation of the time-varying error term in the differenced equation. If $P > 0.05$, the null hypothesis of no second-order autocorrelation indicates that the second and third lags of our endogenous explanatory variables are valid instrumental variables for the differenced equation.

^e $P < 0.05$.

middle-aged adults were more highly educated and in the middle and upper categories of income (Table 2). The population of single-person households included in the present analysis was slightly older and included fewer Hispanics compared with the population of multiperson households originally included in Homescan (data not shown).

Among consumers of LCS and CS beverages, most persons purchased more than 0 but less than 1 serving/day of either beverage (Web Table 2). Approximately 12% of consumers purchased 1 or more servings/day of either type of beverage from 2000 to 2010.

Over the same period, we observed significant secular decreasing trends in purchases of total daily kilocalories, kilocalories excluding LCS and CS beverages, and kilocalories from food and beverages, as well as decreases in total daily

kilocalories from all macronutrients (Web Table 3). Changes over time in food groups are shown in Web Tables 4 and 5.

Overall prices of foods, prices of LCS and CS beverages, and yearly dollar expenditures at the market level increased significantly from 2000 to 2010 (Web Table 6). The average number of grocery trips per household per year decreased significantly over time.

Specification of IVs

The IVs that were used in the model (percentage of market sales and number of grocery trips per year) helped predict our endogenous exposures (Table 3) and were found to have empirical validity as measured by the specification tests. The percentages of market sales that LCS and CS beverages

Table 5. Dynamic Panel Modeling of the Associations of 1-Serving per Day Increases in the Consumption of Beverages Sweetened With Caloric Sweeteners and Low-Calorie Sweeteners With Dietary Purchasing Patterns, Homescan Survey, 2000–2010^a

Outcome at Time <i>t</i> , kilocalories per capita per day	Lagged Endogenous Explanatory Variables ^b			Sargan-Hansen Test ^c P Value	Arellano-Bond Test of Autocorrelation ^d	
	Outcome	LCS Beverages	CS Beverages		First Order P Value	Second Order P Value
Beverage group						
Juice, sweetened	0.73 (0.23) ^e	−2.28 (2.07)	−1.52 (2.20)	0.891	<0.001	0.509
Milk and milk drinks, sweetened	−0.07 (0.18)	−0.48 (0.78)	1.24 (0.94)	0.500	0.064	<0.001
Milk, plain unsweetened	0.36 (0.17) ^e	1.82 (2.33)	2.22 (2.57)	0.438	<0.001	0.516
Coffee/tea, sweetened	−0.44 (0.31)	0.69 (0.71)	−1.08 (0.98)	0.755	0.924	0.006
Coffee/tea, unsweetened	0.76 (0.17) ^e	−0.73 (0.65)	0.40 (0.81)	0.966	<0.001	0.017
Water and other beverages, unsweetened	−0.24 (0.49)	0.00 (0.05)	0.05 (0.05)	0.970	0.282	0.292
Alcohol	0.88 (0.10) ^e	−1.80 (2.21)	−2.83 (1.98)	0.453	<0.001	0.731
Food group						
Dairy, sweetened	0.39 (0.20)	1.76 (1.55)	0.98 (1.43)	0.483	<0.001	0.593
Dairy, plain and unsweetened	0.76 (0.14) ^e	0.92 (0.58)	0.82 (0.53)	0.045	<0.001	0.944
Fruit, processed and sweetened	−0.21 (0.21)	−0.36 (0.57)	0.41 (0.56)	0.874	0.122	0.003
Plain fruits and vegetables	0.28 (0.21)	0.85 (1.53)	0.27 (1.50)	0.522	0.001	0.323
Ready-to-eat cereal, sweetened	0.05 (0.15)	8.13 (3.39) ^e	2.14 (2.66)	0.565	0.003	0.006
Grains and breads	0.81 (0.09) ^e	−0.40 (3.55)	−1.40 (3.52)	0.299	<0.001	<0.001
Desserts and sweeteners, low-calorie	0.39 (0.13) ^e	1.34 (1.77)	−1.29 (1.23)	0.085	<0.001	0.581
Desserts and sweeteners, caloric	0.24 (0.19)	40.18 (14.04) ^e	36.00 (17.30) ^e	0.213	<0.001	0.415
Salty snacks	0.70 (0.27) ^e	1.66 (2.74)	0.04 (2.57)	0.904	<0.001	0.068
Cheese	0.45 (0.20) ^e	5.21 (2.58) ^e	3.92 (2.85)	0.203	<0.001	0.914
Cooking fats and dressings	0.89 (0.22) ^e	−2.22 (7.00)	−7.29 (7.96)	0.311	<0.001	<0.001
Nuts and seeds	0.53 (0.23) ^e	3.10 (3.48)	2.62 (2.80)	0.819	<0.001	0.656
Meat, fish, poultry, and eggs	0.80 (0.08) ^e	−1.71 (3.15)	−1.55 (2.95)	0.878	<0.001	0.004
Ready-to-eat mixed, frozen/fast food meals	0.69 (0.17) ^e	6.37 (3.93)	5.78 (4.78)	0.724	<0.001	0.042

Abbreviations: CS beverages, beverages flavored with caloric sweeteners; LCS beverages, beverages flavored with low-calorie sweeteners.

^a Values are expressed as β (standard error) and were obtained from a generalized method of moments 2-step system dynamic panel model with average no. of household grocery trips per year and percentage market sales made up of LCS beverages and CS beverages (specified for the level equation and differenced equation) as the lagged instrumental variables in the first step and the second and third lags of LCS and CS beverage purchases (specified for the differenced equation) as the lagged instrumental variables in the second step. There were 41 instrumental variables in total. The models were adjusted for age, sex, educational level, race/ethnicity, income, and year.

^b The first lagged value was used.

^c Sargan-Hansen test of overidentifying restrictions. If $P > 0.05$, the null hypothesis of overidentification indicated that the assumptions made about exogeneity of the instrumental variables were valid.

^d Arellano-Bond test of autocorrelation of the time-varying error term in the differenced equation. If $P > 0.05$, the null hypothesis of no second-order autocorrelation indicates that the second and third lags of our endogenous explanatory variables are valid instrumental variables for the differenced equation.

^e $P < 0.05$.

comprised were found to be significant predictors for the explanatory endogenous variables (LCS and CS beverages) but were not associated with main outcomes. We found the number of grocery trips per year to be a significant predictor for the outcomes in the model (e.g., total energy), but the number of grocery trips per year was not associated with the number of LCS and CS beverage purchases. The specification tests showed that, for most of our models, the null hypotheses of overidentification and no second-order autocorrelation cannot be rejected, thus indicating that the majority of our models with IVs were correctly specified (Tables 4 and 5, Web Appendix 1, and

Web Tables 7 and 8). Compared with these results, associations obtained from models that assumed exogeneity of the main exposures (without IVs) were generally weaker, but they were in the same direction (Web Tables 9 and 10).

Consumption of LCS and CS beverages and dietary quality

Each 1-serving/day increase in purchases of either LCS or CS beverages resulted in higher total daily energy (86 kcal and 113 kcal, respectively), total energy excluding energy

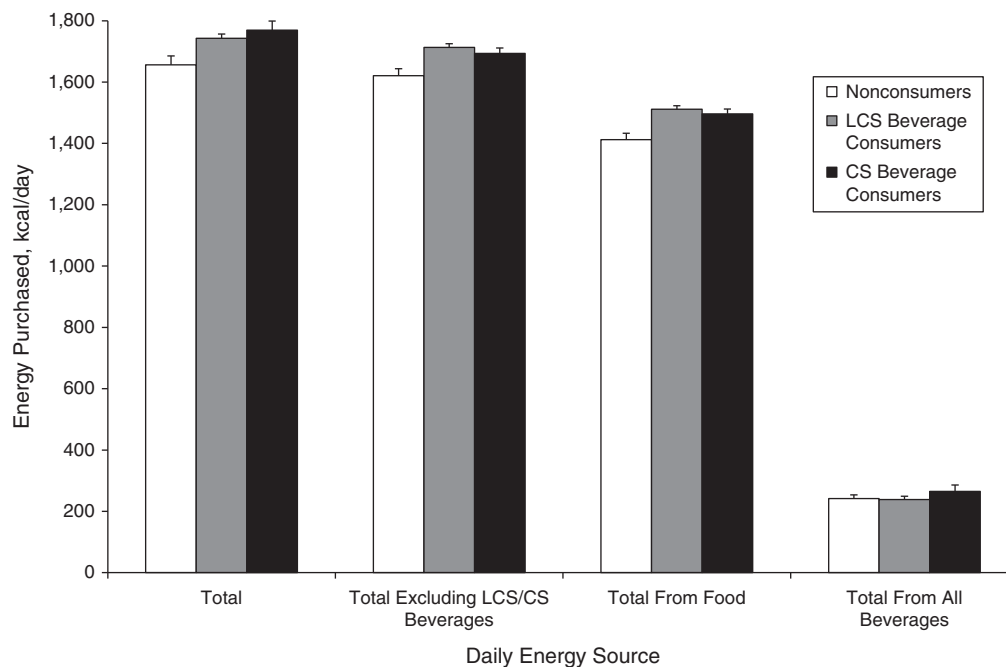


Figure 1. Mean energy (kcal/day) purchased by each beverage-consumer profile, Homescan survey, 2000–2010. Results are presented as means with standard error bars obtained from a generalized method of moments 2-step system dynamic panel model. In the first step, average number of household grocery trips per year and percentages of market sales made up of beverages sweetened with low-calorie sweeteners (LCS beverages) and beverages sweetened with caloric sweeteners (CS) (specified for the level equation and differenced equation) as the lagged instrumental variables. In the second step, we used the second and third lags of LCS and CS beverage purchases (specified for the differenced equation). There were 41 instrumental variables in total. The models were adjusted for age, sex, educational level, race/ethnicity, income, and year of data collection. Using margins commands in the fully adjusted models, we predicted the mean of the outcome for consumers of LCS beverages, consumers of CS beverages, and persons who did not consume sweetened beverages (nonconsumers). LCS beverage consumers were persons who purchased 1 serving/day of LCS beverages but 0 servings/day of CS beverages; the opposite was true for CS beverage consumers. Purchases of total daily energy, energy excluding that from LCS/CS beverages, and energy from food were significantly different between nonconsumers and LCS and CS beverage consumers ($P < 0.05$).

from LCS and CS beverages, and total energy from purchased food (Table 4, Figure 1), as well as more total kilocalories from carbohydrates, sugar, and total fat (Table 4, Web Figure 1). Every additional serving per day of LCS beverages resulted in higher daily kilocalorie purchases from protein and saturated fat. The proportions of each macronutrient did not change with each additional serving of LCS or CS beverages.

In terms of food groups, each additional serving per day of either LCS or CS beverages was significantly associated with higher total daily kilocalories from CS or desserts sweetened with CS (40 kcal and 36 kcal, respectively) (Table 5, Figure 2B). Among consumers of LCS beverages, every additional serving per day was associated with higher total daily kilocalorie purchases from CS cereals and cheese (Table 5, Figure 2A).

DISCUSSION

Using a dynamic panel model with IVs, we investigated the associations of CS and LCS beverage consumption with dietary quality and food-purchasing patterns over the past decade in the United States. Despite overall declines in purchases of kilocalories from all sources (kilocalories per capita per day), we found that every additional serving per day of

either CS or LCS beverages was associated with higher total energy from all purchases and kilocalories from most macronutrients. However, the dietary distribution of macronutrients did not change, which indicates that the absolute changes were driven by an absolute increase in the total number of kilocalories consumed. We also found that every additional serving per day of either beverage was associated with a significant increase in the number of purchases of caloric desserts/sweeteners, which accounted for an important proportion of the increase in overall kilocalories (excluding those from CS and LCS beverages).

Consumption of CS beverages has been associated with higher caloric intakes and poorer dietary patterns (7, 20–22). In a recent study on the long-term associations between different profiles of beverage consumers and dietary patterns, Piernas et al. (23) found that households in which either LCS or CS beverages were consumed were significantly less likely to follow healthier dietary patterns compared with nonconsumers. However, LCS beverage consumers also had a higher probability of following a prudent dietary pattern, which is characterized by eating fruits/vegetables augmented with snacks and LCS desserts. In another study, Duffey et al. (7) reported that consumers of LCS beverages had a lower cardiometabolic

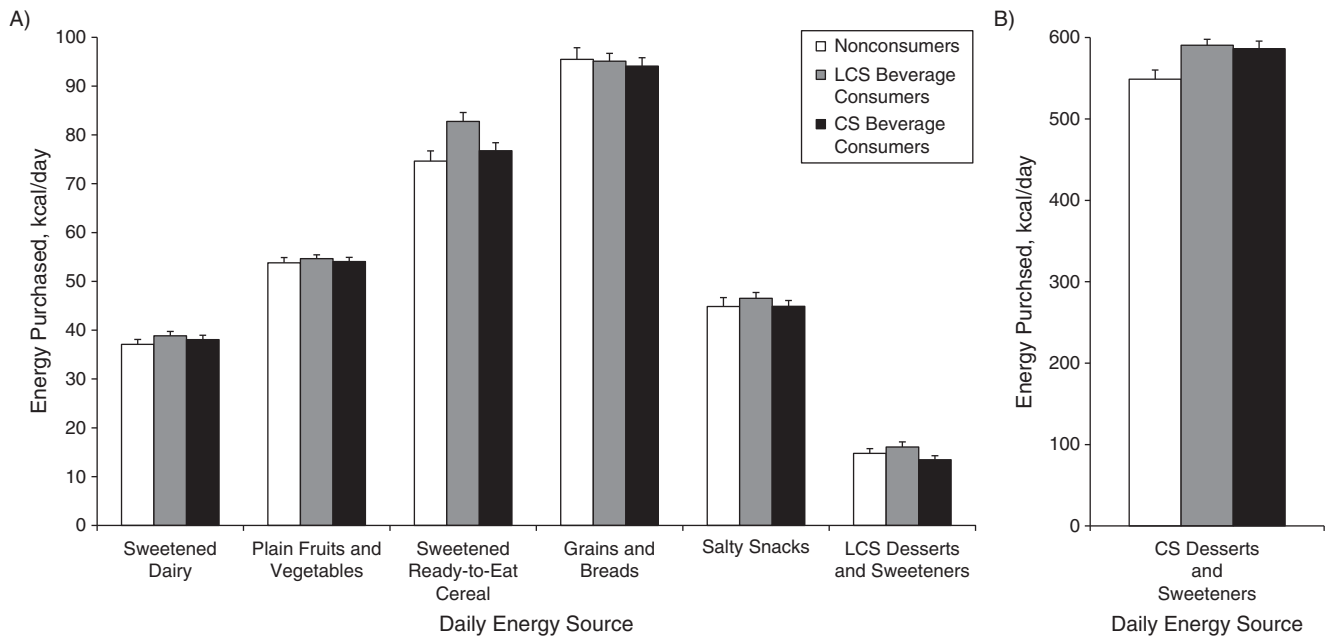


Figure 2. Mean energy (kcal/day) from selected food groups (A) and caloric-sweetened desserts and sweeteners (B) by each beverage-consumer profile, Homescan survey, 2000–2010. Results are presented as means with standard error bars obtained from a generalized method of moments 2-step system dynamic panel model. In the first step, average number of household grocery trips per year and percentages of market sales made up of beverages sweetened with low-calorie sweeteners (LCS beverages) and beverages sweetened with caloric sweeteners (CS) (specified for the level equation and differenced equation) as the lagged instrumental variables. In the second step, we used the second and third lags of LCS and CS beverage purchases (specified for the differenced equation). There were 41 instrumental variables in total. The models were adjusted for age, sex, educational level, race/ethnicity, income, and year of data collection. Using margins commands in the fully adjusted models, we predicted the mean of the outcome for consumers of LCS beverages, consumers of CS beverages, and persons who did not consume sweetened beverages (nonconsumers). LCS beverage consumers were persons who purchased 1 serving/day of LCS beverages but 0 servings/day of CS beverages; the opposite was true for CS beverage consumers. LCS beverage consumers had total daily purchases of energy from sweetened ready-to-eat cereals that were significantly different from those of nonconsumers and CS beverage consumers ($P < 0.05$). Nonconsumers had total daily energy purchases from CS desserts and sweeteners that were significantly different from those of LCS and CS beverage consumers ($P < 0.05$).

risk when following a prudent diet rather than a Western diet. An earlier cross-sectional study using purchases recorded by Homescan in 1999 compared the food patterns of CS and LCS beverage consumers and concluded that LCS beverage consumers chose foods that had lower energy contents or that were less energy dense than those chosen by CS beverage consumers (24). However, none of these studies could determine the directionality of these relationships. In the present study, we demonstrated that, after accounting for endogenous decisions about food choices and other unmeasured confounding factors, consumers of either type of sweetened beverage had higher caloric purchases from all sources and from most macronutrients than did nonconsumers.

We showed that kilocalories from carbohydrates, sugars, and CS desserts/sweeteners significantly increased with every additional serving per day of either CS or LCS beverages. In an earlier study in which Homescan data from 1999 were used, Binkley et al. (24) reported that consumers of CS and LCS beverages purchased significantly candy than did nonconsumers, whereas consumers of LCS beverages purchased significantly more cookies and low-fat ice cream than did CS beverage consumers or nonconsumers (24). However, in a recent randomized controlled trial in which participants were randomized into groups that substituted CS beverages with either

LCS beverages or water, the investigators did not find a differential effect in intakes of kilocalories, macronutrients, or desserts in the LCS beverage compared with the water group (18). This conflicting finding could be explained by the fact that the patients enrolled in the randomized controlled trial were overweight participants who were highly motivated to lose weight. In the present study, we observed participants' behavior in free-living conditions, and both LCS and CS beverage consumers increased their daily intakes of kilocalories from CS desserts/sweeteners.

We determined that there are many similarities between LCS beverage consumers and CS beverage consumers in terms of food choices. Some researchers have hypothesized that consumers of sweeteners and sweetened beverages have a higher preference for sweetness (25, 26). Also, repeated exposure to LCS uncoupled with energy was hypothesized to modify the natural relationship between sweet taste and energy, which could affect appetite and energy intake by disrupting hormonal and neurobehavioral pathways that control hunger and satiety (27–31). However, in a recent randomized control trial, researchers showed that children randomized to LCS or CS beverages had no differences in satiety (32). On the other hand, dietary intake is also influenced by the important mechanisms and behaviors involved in food selection. From the behavioral point of view,

persons might use consumption of LCS as a rationale to consume an unhealthy diet or even larger portion sizes because of the common belief that “diet” products (which typically contain LCS) have fewer calories. A preference for sweet taste is considered to be a universal trait, and it involves biological mechanisms related to food reward (i.e., consumption of certain foods reinforces eating behavior) and other nutritional properties of sugars (25, 33). Although there are large variations in the preferred level of sweetness that modulate patterns of consumption of sweeteners and sweet-tasting products, highly processed and intensely sweet products are becoming increasingly popular and marketed in the United States (19).

The longitudinal nature of Homescan allowed us to study long-term dynamics in purchasing patterns for a large sample of individuals while controlling for unmeasured individual determinants that affect food selection. The main advantage of our conceptual dynamic panel model and our statistical approach was that confounding due to endogeneity and unobserved heterogeneity could be addressed, which allowed us to obtain potentially valid estimates of the true relation between dietary quality and sweetened-beverage consumption. However, although the conditions of theoretical exogeneity of IVs (which cannot be empirically tested) and relevancy of IVs (in predicting our endogenous exposures) had to be met in order for us to have confidence in our IV approach, there is still the possibility that our IVs were weak, given that many of these theoretical and empirical requirements are difficult to test and are sometimes unmet. In this adverse scenario, the consequences of using an IV approach could be worse than those of using ordinary least-squares models for our estimates of association. On the other hand, our dynamic panel model allowed use of lagged values of beverage consumption and IVs, which helped to set up an adequate temporality such that the main exposure could occur before the outcome. The use of purchasing data from Homescan constitutes an alternative method for characterizing population eating patterns (34). Food-purchasing and expenditure surveys have been previously used to measure household food availability. While these data sets do not capture individuals’ actual dietary intakes, they are useful for characterizing the wide variability in food-consumption patterns (35–37). Although the process of scanning and recording the purchases might be time-consuming and prone to recording errors, Homescan has been validated using retailer’s transaction and diary survey data, and its overall accuracy is in line with many other commonly used surveys of this type (38, 39). One important advantage of using Homescan is the availability of ingredient information for each product in the US marketplace. Our approach also addressed issues related to measurement of sweeteners in the food supply that no other databases have achieved. All products that contained sweeteners were objectively identified and classified, which let us avoid the potential misclassification error that likely affects data self-reported by individuals who are potentially unaware of the type of sweetener in their products.

The main limitation of using Homescan data is the unavailability of information on foods and beverages consumed outside of the home, although this consumption accounts for less than one-third of CS and LCS beverage consumption (9). In addition, we could be underestimating the associations with other foods that are usually consumed outside the home (e.g.,

ice cream). Furthermore, the generalizability of the study might be limited because it included only single-person households comprising individuals who were mostly middle-aged and non-Hispanic white and had higher educational levels and incomes. Adults who live alone might not be representative of an average person; they might have different dietary patterns and away-from-home eating patterns. People who live in single-person households might purchase some food to share with others and might waste food more, especially perishable products. Another source of measurement error might come from missing information about purchases of nonbarcoded, random-weight products that are not prebagged (e.g., loose fruits, nuts).

In conclusion, because consumers appear to be turning to LCS as their sweetener of choice more frequently, our study opens up new pathways that relate consumption of both LCS and CS beverages to purchases containing more overall energy, carbohydrates, sugar, and caloric desserts/sweeteners. While the current state of research on this topic is incomplete and unclear, our results have significant public health implications, especially with regard to consumption of products that contain LCS. It is essential to understand whether sweetener consumption translates into a better or worse dietary quality before continuing with more complex studies that relate sweetener intake to health outcomes. Our research combined an advanced approach and sweetener classification to contribute new evidence to the understanding of the mechanisms potentially implicated in the relationship between sweetener consumption and lower nutritional quality. Our findings, which suggest that consumption of any type of sweetened beverage might negatively affect diet, can potentially inform future intervention strategies and nutrition policy recommendations aimed at improving diet and nutrition in the United States.

ACKNOWLEDGMENTS

Author affiliations: Department of Nutrition, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina (Carmen Piernas, Shu Wen Ng, Michelle A. Mendez, Penny Gordon-Larsen, Barry M. Popkin).

This work was supported by the Robert Wood Johnson Foundation (grant 70017) and the Carolina Population Center 5 R24 HD050924.

We thank the entire University of North Carolina Food Research Program for their assistance. We are especially grateful to Dr. Donna Miles for exceptional assistance with the data management and programming, Frances L. Dancy for administrative assistance, and Dr. David Guilkey (Department of Economics, University of North Carolina) for his contributions of reviewing and improving this work.

Conflict of interest: none declared.

REFERENCES

1. Malik VS, Hu FB. Sweeteners and risk of obesity and type 2 diabetes: the role of sugar-sweetened beverages. *Curr Diab Rep.* 2012;12(2):195–203.

2. Mattes R. Fluid calories and energy balance: the good, the bad, and the uncertain. *Physiol Behav.* 2006;89(1):66–70.
3. de Koning L, Malik VS, Kelllogg MD, et al. Sweetened beverage consumption, incident coronary heart disease, and biomarkers of risk in men. *Circulation.* 2012;125(14):1735–1741.
4. de Koning L, Malik VS, Rimm EB, et al. Sugar-sweetened and artificially sweetened beverage consumption and risk of type 2 diabetes in men. *Am J Clin Nutr.* 2011;93(6):1321–1327.
5. Mattes R. Effects of aspartame and sucrose on hunger and energy intake in humans. *Physiol Behav.* 1990;47(6):1037–1044.
6. Mattes RD, Popkin BM. Nonnutritive sweetener consumption in humans: effects on appetite and food intake and their putative mechanisms. *Am J Clin Nutr.* 2009;89(1):1–14.
7. Duffey KJ, Steffen LM, Van Horn L, et al. Dietary patterns matter: diet beverages and cardiometabolic risks in the longitudinal Coronary Artery Risk Development in Young Adults (CARDIA) Study. *Am J Clin Nutr.* 2012;95(4):909–915.
8. Greenland S. An introduction to instrumental variables for epidemiologists. *Int J Epidemiol.* 2000;29(4):722–729.
9. Johnston KM, Gustafson P, Levy AR, et al. Use of instrumental variables in the analysis of generalized linear models in the presence of unmeasured confounding with applications to epidemiological research. *Stat Med.* 2008;27(9):1539–1556.
10. Wooldridge JM. *Econometric Analysis of Cross Section and Panel Data.* Cambridge, MA: MIT Press; 2002.
11. Hernán MA, Robins JM. Estimating causal effects from epidemiological data. *J Epidemiol Community Health.* 2006;60(7):578–586.
12. Hernán MA, Robins JM. Instruments for causal inference: an epidemiologist's dream? *Epidemiology.* 2006;17(4):360–372.
13. The Nielsen Co. Nielsen Consumer Panel and Retail Measurement. <http://www.nielsen.com/ca/en/solutions/measurement/retail-measurement.html>. Accessed November 1, 2012.
14. Ng SW, Popkin BM. Monitoring foods and nutrients sold and consumed in the United States: dynamics and challenges. *J Acad Nutr Diet.* 2012;112(1):41–45.e4.
15. Slining MM, Ng SW, Popkin BM. Food companies' calorie-reduction pledges to improve U.S. diet. *Am J Prev Med.* 2013;44(2):174–184.
16. Blundell R, Bond S. Initial conditions and moment restrictions in dynamic panel data models. *J Econom.* 1998;87(1):115–143.
17. Blundell R, Bond S, Windmeijer F. *Estimation in Dynamic Panel Data Models: Improving on the Performance of the Standard GMM Estimator.* Bingley, United Kingdom: Emerald Group Publishing Limited; 2001.
18. Arellano M, Bond S. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Rev Econ Stud.* 1991;58(2):277–297.
19. Piernas C, Ng SW, Popkin B. Trends in purchases and intake of foods and beverages containing caloric and low-calorie sweeteners over the last decade in the United States. *Pediatr Obes.* 2013;8(4):294–306.
20. Duffey KJ, Popkin BM. Adults with healthier dietary patterns have healthier beverage patterns. *J Nutr.* 2006;136(11):2901–2907.
21. Frary CD, Johnson RK, Wang MQ. Children and adolescents' choices of foods and beverages high in added sugars are associated with intakes of key nutrients and food groups. *J Adolesc Health.* 2004;34(1):56–63.
22. Mathias KC, Slining MM, Popkin BM. Foods and beverages associated with higher intake of sugar-sweetened beverages. *Am J Prev Med.* 2013;44(4):351–357.
23. Piernas C, Mendez MA, Ng SW, et al. Low-calorie- and calorie-sweetened beverages: diet quality, food intake, and purchase patterns of US household consumers. *Am J Clin Nutr.* 2014;99(3):567–577.
24. Binkley J, Golub A. Comparison of grocery purchase patterns of diet soda buyers to those of regular soda buyers. *Appetite.* 2007;49(3):561–571.
25. Drewnowski A. Sweetness, appetite, and energy intake: physiological aspects. *World Rev Nutr Diet.* 1999;85:64–76.
26. Holt SHA, Cobiac L, Beaumont-Smith NE, et al. Dietary habits and the perception and liking of sweetness among Australian and Malaysian students: a cross-cultural study. *Food Qual Prefer.* 2000;11(4):299–312.
27. Blundell JE, Hill AJ. Paradoxical effects of an intense sweetener (aspartame) on appetite. *Lancet.* 1986;1(8489):1092–1093.
28. Tordoff MG. How do non-nutritive sweeteners increase food intake? *Appetite.* 1988;11(Suppl 1):5–11.
29. Davidson TL, Swithers SE. A Pavlovian approach to the problem of obesity. *Int J Obes Relat Metab Disord.* 2004;28(7):933–935.
30. Swithers SE, Davidson TL. A role for sweet taste: calorie predictive relations in energy regulation by rats. *Behav Neurosci.* 2008;122(1):161–173.
31. Brown RJ, Walter M, Rother KI. Ingestion of diet soda before a glucose load augments glucagon-like peptide-1 secretion. *Diabetes Care.* 2009;32(12):2184–2186.
32. de Ruyter JC, Katan MB, Kuijper LD, et al. The effect of sugar-free versus sugar-sweetened beverages on satiety, liking and wanting: an 18 month randomized double-blind trial in children. *PLoS One.* 2013;8(10):e78039.
33. Drewnowski A, Mennella JA, Johnson SL, et al. Sweetness and food preference. *J Nutr.* 2012;142(6):1142S–1148S.
34. Fan JX, Brown BB, Kowaleski-Jones L, et al. Household food expenditure patterns: a cluster analysis. *Mon Labor Rev.* 2007;130:38.
35. Harris JM. Using Nielsen Homescan data and complex design techniques to analyze convenience food expenditures [abstract]. Presented at 2005 Annual meeting, Providence, RI, July 24–27, 2005.
36. Naska A, Paterakis S, Eeckman H, et al. Methodology for rendering household budget and individual nutrition surveys comparable, at the level of the dietary information collected. *Public Health Nutr.* 2001;4(5B):1153–1158.
37. Naska A, Vasdekis VG, Trichopoulou A. A preliminary assessment of the use of household budget survey data for the prediction of individual food consumption. *Public Health Nutr.* 2001;4(5B):1159–1165.
38. Einav L, Leibtag E, Nevo A. On the accuracy of Nielsen Homescan data. *Econ Res Rep.* 2008;69.
39. Einav L, Leibtag E, Nevo A. *Not-so-classical Measurement Errors: a Validation Study of Homescan.* Cambridge, MA: National Bureau of Economic Research; 2008.