Neighbourhood variation in the price of soda relative to milk and its association with neighbourhood socio-economic status and race

David M Kern, Amy H Auchincloss* Lance S Ballester and Lucy F Robinson
Department of Epidemiology and Biostatistics, School of Public Health, Drexel University, Nesbitt Hall, 3215 Market Street, Philadelphia, PA 19104, USA

Abstract

Objective: Soda consumption is high in the USA, especially among minorities and individuals of lower socio-economic status (SES); this may be due to its affordable price in relation to healthier alternatives. The objective of the present study was to examine geospatial variation in price of milk and soda, and the price of milk relative to soda, by neighbourhood SES and proportion of Hispanic and black individuals.

Design: Retailer soda and milk prices (*n* 2987; Information Resources, Inc. Academic Data Set 2004–2011) were linked to census block group sociodemographic characteristics (American Community Survey 2005–2009). Linear hierarchical regression models were used to adjust for confounders.

Setting: Large chain supermarkets and superstores (*n* 1743) in forty-one states and 1694 block groups (USA).

Results: For equivalent fluid ounces, price of soda on average was 62 % lower than milk ($US 0·23 *v.* $US 0·63 per serving) and there was high dispersion in milk price across geographic areas. After adjustment for confounding, neighbourhoods with a higher concentration of black and Hispanic individuals tended to have lower soda prices and higher milk prices (−$US 0·001 and +$US 0·007 in price per serving, respectively, for a one quintile increase in black/Hispanic population), while soda and milk both became less expensive as SES decreased (−$US 0·002 and −$US 0·015 in serving price per one SD decrease in SES index, respectively).

Conclusions: Neighbourhoods with a higher concentration of blacks and Hispanics may be at greater risk of higher soda consumption due to more affordable prices, in absolute terms and relative to the price of milk.

Soft drink sales per capita are higher in the USA than anywhere else in the world, other than Argentina[1]. Approximately one-half of the US population consumes a sugar drink on any given day[2], with blacks, Hispanics and individuals of low socio-economic status (SES) most likely to do so[3]. Sugar-sweetened beverages, such as soda, are nutrient deficient and most or all of the energy they provide comes from refined added sugars, typically high-fructose corn syrup. The US Department of Agriculture recommends no more than 10 % of energy be obtained from added sugars[4], yet a single can of cola includes 586 kJ (140 kcal) of added sugar[5], 7 % of energy in a 8368 kJ/d (2000 kcal/d) diet. Consumption of soda and other sugar-sweetened beverages has been linked to poor health outcomes such as diabetes[6–9], obesity/weight gain[10–12] and CVD[13–15]. Furthermore, these chronic diseases tend to be more prevalent in blacks, Hispanics and those of low SES[16–19].

Conversely, consumption of fluid milk has dropped by nearly 18 % in the last 10 years[20] and current consumption is lower than levels recommended by the US Department of Agriculture[4]. Additionally, blacks consume less milk per capita than other races and adolescents from lower-income households consume less than those from higher incomes[21]. While milk's health benefits have been questioned[22–24], in general, epidemiological evidence supports a positive association between dairy consumption and health, specifically a lower risk of chronic diseases. For example, studies have found that increased consumption of low-fat dairy is associated with a lower risk of developing diabetes in men[25] and women[26], as well as lower odds of developing insulin resistance syndrome and obesity[27], and consumption of dairy in general is associated with a lower risk of CVD, all-cause death, stroke and diabetes[28,29]. These benefits are not limited to low-fat dairy; there is evidence that even

Corresponding author: Email aha27@drexel.edu

full-fat dairy may decrease the risk of being overweight or obese compared with low-fat dairy intake[30]. Some possible reasons for the benefit conferred by milk are appetite satiety and replacing less healthful drinks with milk[31,32].

High consumption of sugar-sweetened beverages may be driven in part by their affordability compared with more healthful alternatives[33]. Differential availability across sociodemographic factors may also contribute to disparities in intake. It is possible that individuals of certain neighbourhoods are at an increased risk for consuming soda due to its relative affordability. This is especially likely for individuals of lower SES who are more sensitive to food prices[34] – due to a higher proportion of income spent on food[35] – and those who are at an increased risk of obesity and diabetes, such as blacks and Hispanics. While the prices of unhealthy and healthy foods have been studied previously[36–40], there is little information regarding how the costs of those foods are related to the demographics of the neighbourhood in which they are sold.

The present study compiled an analytic data set of soda and milk prices across the USA in order to examine differences between the two beverages and whether that difference is associated with neighbourhood characteristics, including race and SES. Furthermore, the study aimed to understand variation in prices by the degree of urbanization and geographic locale and spatial scale. To our knowledge, no previous study has examined the relative price of healthy v. unhealthy foods at the store level and the association with neighbourhood-level demographics across a wide-ranging geographic area. Understanding how prices of foods differ by neighbourhood may help explain differences in diet and health for residents of those same neighbourhoods.

Methods

Price data

The current retrospective analysis utilized product pricing data from the Information Resources, Inc. (IRI) Academic Data Set. IRI is a market research group focused on consumer packaged goods sold in large chain supermarkets and superstores (Albertson's, A&P, Food Emporium, Pathmark, Super Fresh, Kmart, BJ's, Sam's Club, Walmart, etc.) across the USA[41,42]. Data were collected weekly at the Universal Product Code (UPC) level within each store and contained the store identifier, item UPC, number of units sold and the price of the item, among other variables. In total, data used in the present study span January 2004 through December 2011 and came from 1743 stores across forty-one states (including Washington, DC) and 1694 block groups. The wide temporal range of data was used to maximize available UPC across stores in order to increase sample size and the stability of prices, and was combined to create a cross-sectional data sample.

IRI's Academic Data Set represented only a small number of products available in stores (thirty packaged products that included foods like mayonnaise and mustard, and non-foods such as deodorant, facial tissue, etc.). Refrigerated milk and sugar-sweetened soda were selected because of their generally favourable and unfavourable associations with population health respectively, their ability to serve as rough proxies for healthier fresh foods/beverages v. long shelf-life foods/beverages that offer little or no nutritional benefit, and due to both of these products having high sales. Milk is recommended as a healthful beverage in dietary guidelines which call for three cups of dairy per day for adolescents and adults[4]. Diet soda was not considered for the healthier beverage category because it offers no nutritional value and is not widely viewed as a healthful product due to associations with increased body weight and higher incidence of diabetes[12,43,44]. Bottled water was not available.

To increase comparability across beverage categories and across stores, only brand-name items (according to the IRI classification) were retained and analysed. Branded soda represents 95 % of the soda sales observed in the data set (Coke, Pepsi, Sprite, etc.). Branded milk includes national or semi-national (Horizon, Organic Valley, Hood, Deans, etc.) and regional (Alta Dena, Kemps, Tuscan Dairy, etc.) brands and represents a majority (53 %) of milk sales observed in the data set. Whole, low-fat and skimmed dairy milk were included because prices varied only slightly between these types (skimmed was slightly cheaper than whole milk). Refrigerated non-dairy alternatives (e.g. almond milk, soya milk, etc.) were excluded because they had low sales. Only sugar-sweetened sodas were included for analysis. Diet soda products were excluded because they had lower sales than sugar-sweetened soda and had the same price for equivalent brand and size (thus redundant with sugar-sweetened soda). To avoid potential confounding due to different packaging sizes of items, each product was limited to a single size: 2-litre bottles of soda and half-gallon containers of milk, chosen for their similar volumes (2 litres = 0·53 gallons). All products that met the criteria described above were included in the analysis. Stores sold an average of 81 (SD 24·5) unique soda items, accounting for a median of 580 942 units of soda over the entire observation period; and an average of 45 (SD 18·1) unique milk items and median of 213 922 units of milk sold during the same time.

The IRI Academic Data Set includes data for 2001–2011 but because inflation-adjusted prices were more or less stable during 2004–2011, we excluded 2001–2003 and aggregated prices across the 8-year period to create a cross-sectional data set, one observation per store. This approach was taken since we were most interested in analysing spatial variation in prices.

Price variables

Prices reflect the shelf price and included store-level promotions and retailer coupons, but did not include changes in price applied at the cash register, including

taxes and manufacturer coupons. Volume equivalent prices were calculated as the price of an item divided by the number of fluid ounces (fl. oz) the item contained and multiplied by 12 to create a price per 12 fl. oz (equivalent to 'serving size'), and were calculated for every product meeting the inclusion criteria described previously. A standard can of soda is 12 fl. oz (~355 ml), and by selecting 12 fl. oz for soda and milk we were able to compare across the two products. The price of soda and the price of milk for each store were calculated by aggregating over all items sold by the store during the entire observation period (2004–2011). Store-level prices were weighted according to the number of 12 fl. oz servings sold of each item in order to not overweight items that were infrequently purchased, adjusted for inflation according to the US Bureau of Labor and Statistics Consumer Price Index and reported in 2011 US dollars[45].

The average price of brand-name toilet paper was calculated for each store in order to control for baseline costs specific to each store which may be influencing the price of food (e.g. rent, distribution and employee wages) that may not be captured through other variables. Toilet paper is a good proxy for the cost of doing business because it is non-perishable, takes up significant shelf space and is a supermarket consumer staple item[46], thus not likely influenced by differences in demand across stores the way individual food prices may be.

Outcome variables

The outcomes of interest were price of milk alone, price of soda alone and the relative price of soda compared with milk. The relative price was operationalized as the ratio of the average price per 12 fl. oz of soda divided by the average price per 12 fl. oz of milk. A ratio <1·0 indicates that the average price of a single serving of soda is less than the price of a serving of milk, while a ratio >1·0 indicates that a serving of soda is more expensive than a serving of milk on average. The interpretation of the ratio is price of soda as a percentage of the price of milk; for example, a ratio of 0·42 indicates that the price of a serving of soda is 42 % of the price of a serving of milk.

Census variables

Stores were assigned to the population-weighted centroid of their block group (store block group was known but street address was not). In order to characterize the census composition around the stores, block groups within 1 mile (1·6 km) of each store were selected and census data were averaged for those block groups. Block groups intersecting the 1-mile buffer were included as the store's neighbourhood and data from each of the block groups in the neighbourhood were population weighted. The buffer was chosen to expand the supermarket's neighbourhood beyond the block group in which it was located, which may be an industrial area with low population and not representative of the surrounding neighbourhood. Others have referred to a 1-mile buffer as a relevant consumer market area for a supermarket[47]. In order to test how sensitive the results were to the chosen radius, a second analysis was performed using a 2-mile buffer (3·2 km) and results were consistent with the original analysis (data not shown).

Census data were obtained from the 2005–2009 American Community Survey 5-year summary file, approximately the mid-years of the price data file. Neighbourhood SES was operationalized as a continuous measure as described by Diez-Roux et al.[48], using six variables from the American Community Survey representing wealth and income, housing value, education, and managerial or professional occupations (see Table 1 footnote). These variables were combined into a summary score for each block group by creating a Z-score for each variable (i.e. the number of standard deviations away from the overall mean) and summing across all six Z-scores. This measure was then transformed into a Z-score. The second exposure of interest in the present study was race, measured as the proportion of individuals who were black or non-white Hispanic. To control for potential differences in age distribution across block groups the proportion of individuals aged 20 to 39 years, standardized as a Z-score, was included as a covariate in the multivariable models.

Geographic variables

Geographic region (Northeast, Midwest, South and West) and urbanicity of each store location were included as covariates in the regression analysis. Both were included to control for differences in infrastructure and other aspects of the built environment unique to regions or to cities v. rural areas which could affect the price of either soda or milk. Urbanicity was based on county population size and was operationalized as a categorical variable with four levels: (i) large metropolitan area of 1 million or more residents; (ii) small metropolitan area of less than 1 million residents; (iii) micropolitan urban areas (centred on an urban area with population of 10 000 to 49 999); and (iv) non-core (all other areas smaller than micropolitan)[49].

Supermarket density

Supermarket density data were obtained from the National Center for Chronic Disease Prevention and Health Promotion of the Centers for Disease Control and Prevention, which compiled the number of chain and non-chain supermarkets within each census tract or within 0·5 miles (0·8 km) of the tract boundary[50]. Supermarket density was included as a covariate to control for potential differences in market competition that may affect product prices[51].

Statistical analysis

Bivariate associations between demographics (race, neighbourhood SES, region and urban class) and soda and

Table 1 Soda and milk prices at large chain supermarkets by census neighbourhood sociodemographics and geographic region, USA

	Number of stores	Soda ($US)*,†		Milk ($US)*,†		Soda:milk price ratio†,‡	
		Mean	SD	Mean	SD	Mean	SD
Overall (2004–2011)	1743	0·23	0·02	0·63	0·11	0·379	0·082
Neighbourhood SES tertile§							
Lowest tertile (least advantaged)	581	0·23	0·02	0·60	0·11	0·392	0·094
Middle tertile	582	0·22	0·02	0·61	0·11	0·379	0·08
Highest tertile (most advantaged)	580	0·23	0·02	0·65	0·10	0·363	0·069
Proportion black/Hispanic quintile							
Lowest (0 to 6·8 % black or Hispanic)	349	0·23	0·02	0·57	0·11	0·418	0·096
Second (6·8 to 13·0 %)	349	0·23	0·02	0·62	0·12	0·384	0·087
Middle (13·0 to 22·6 %)	347	0·23	0·02	0·62	0·11	0·380	0·077
Fourth (22·6 to 41·8 %)	350	0·23	0·02	0·65	0·09	0·355	0·057
Highest (41·8 to 98·2 %)	348	0·22	0·02	0·65	0·11	0·354	0·071
Urban classification							
Large metropolitan (population ≥1 million)	1165	0·23	0·02	0·64	0·10	0·363	0·069
Small metropolitan (population <1 million)	484	0·23	0·02	0·59	0·12	0·406	0·095
Micropolitan (population 10 000–50 000)	71	0·23	0·02	0·57	0·11	0·429	0·102
Non-core areas (population <10 000)	23	0·24	0·03	0·60	0·09	0·419	0·108
Region							
Northeast	379	0·22	0·02	0·57	0·10	0·403	0·083
Midwest	315	0·22	0·02	0·57	0·14	0·411	0·115
South	641	0·23	0·02	0·65	0·09	0·360	0·064
West	408	0·24	0·02	0·67	0·09	0·358	0·06

SES, socio-economic status.
*Prices are per 12 fl. oz (~355 ml) serving of soda and milk; all prices are inflation adjusted and reported in 2011 US dollars.
†Statistical testing for bivariate associations between soda and milk prices and the soda:milk price ratio with demographic characteristic was performed. Normal linear regression was used for neighbourhood SES (continuous Z-score) and proportion black/Hispanic (ordinal variable of quintiles), and ANOVA models were used for the urban classification and region. P values for all comparisons were <0·0001.
‡A ratio <1·0 indicates that the average price of a single serving of soda is less than the price of a serving of milk, while a ratio > 1·0 indicates that a serving of milk is more expensive than a serving of soda on average. The ratio can be interpreted as the price of soda as a percentage compared with that of milk; for example, a ratio of 0·42 indicates that the price of a serving of soda is 42 % of the price of a serving of milk. A higher soda:milk price ratio indicates a smaller difference between the two products and may be due to higher soda prices and/or lower milk prices.
§Neighbourhood SES was derived from the log of the median household income; the log of the median value of housing units; the percentage of households receiving interest, dividend or net rental income; the percentage of adults 25 years of age or older who had completed high school; the percentage of adults 25 years of age or older who had completed college; and the percentage of employed persons 16 years of age or older in executive, managerial or professional specialty occupations.

milk prices and the ratio were analysed using unadjusted normal linear models. SES was normally distributed and was converted to a Z-score for analysis. Race (measured by the proportion of individuals who were black or Hispanic) was heavily skewed and thus was categorized into quintiles according to the observed distribution and then entered as an ordinal variable in regression analysis. Region and urbanicity were treated as categorical variables.

To account for spatial dependence of prices between stores proximal to each other, a hierarchical model was used to model stores nested within states and counties. The hierarchical models used random intercepts for state and county to allow mean price to vary across space and account for correlation between prices within a given geographic area.

The analysis was exploratory in nature and methodically tested for associations between price and outcomes and a number of demographic characteristics to understand what factors may explain differences between race and neighbourhood SES with price outcomes. Following work by others[52], the analysis began with an empty model which included only the random effects of state and county. Region and urbanicity were first added to the model to analyse the effects of geography on price, and the price of toilet paper was added to control for baseline costs of doing business at each specific store (Table 2). A description of variance decomposition was made possible by quantifying the amount of clustering by state and county (intraclass correlation coefficient)[52]. After examining the geographic effects on price, models performing the regression of price v. race and SES were analysed in the same manner, nesting stores within county and state (Table 3). For these models, the outcomes were price ratio, soda alone and milk alone. Race and SES effects were first modelled separately with no additional confounders. The exposures were then included in the same model with additional control for age, region, urbanicity, supermarket density and the price of milk (in the soda price analysis) or soda (in the milk price analysis).

Results

Descriptive results
The price of soda was always cheaper than the price of milk by a large margin across all observed neighbourhood characteristics (Table 1; $US 0·23 v. $US 0·63 per 12 fl. oz serving). Thus, the unadjusted difference between milk

Table 2 Hierarchical models examining geographical associations with the soda:milk price ratio, soda price and milk price, nested within county and state (*n* 1743), USA

Fixed effects	Soda:milk price ratio		Soda price		Milk price	
	Estimate*	95 % CI	Estimate	95 % CI	Estimate	95 % CI
Intercept	0·3690	0·318, 0·420	0·2365	0·228, 0·245	0·6434	0·585, 0·702
Region						
Northeast	0·0272	−0·037, 0·091	−0·0174	−0·0276, −0·0072	−0·0794	−0·154, −0·0050
Midwest	0·0716	0·0086, 0·135	−0·0127	−0·0227, −0·0026	−0·0994	−0·173, −0·0264
South	−0·0264	−0·085, 0·032	−0·0101	−0·0195, −0·0008	0·0196	−0·0486, 0·0878
West		Ref.		Ref.		Ref.
Urbanicity						
Large metropolitan (population ≥1 million)		Ref.		Ref.		Ref.
Small metropolitan (population <1 million)	0·0298	0·0187, 0·041	0·0042	0·0010, 0·0074	−0·0274	−0·0410, −0·0139
Micropolitan (population 10 000–50 000)	0·0499	0·0315, 0·068	0·0095	0·0039, 0·0151	−0·0361	−0·0590, −0·0131
Non-core areas (population <10 000)	0·0655	0·0377, 0·093	0·0179	0·0092, 0·0266	−0·0395	−0·0745, −0·0045
Additional control variable						
TP price	−0·0059	−0·0090, −0·0028	0·0049	0·0038, 0·0059	0·0187	0·0147, 0·0227
Random effects	Variance	SE	Variance	SE	Variance	SE
Between states	0·00377	0·00095	0·000074	0·000025	0·00500	0·00128
Between counties (in states)	0·00092	0·00014	0·000062	0·000013	0·00119	0·00018
Within counties	0·00246	0·000094	0·000298	0·000011	0·00429	0·00016
ICC (state level)	0·527		0·171		0·477	
ICC (county level)	0·129		0·231		0·115	

ICC, intraclass correlation coefficient (the proportion of variation in price that is accounted for by differences at the state and county levels); TP, toilet paper; ref., referent category.
*A higher soda:milk price ratio indicates a smaller difference between the two products and may be due to higher soda prices and/or lower milk prices; a negative coefficient indicates lower price of soda relative to milk; and positive coefficient indicates higher price of soda relative to milk. Soda and milk price are measured in 2011 US dollars.

Table 3 Hierarchical models nested within county and state performing the regression of soda, milk and soda:milk price outcomes *v.* neighbourhood socio-economic status (SES) and the proportion of individuals who are Hispanic or black (*n* 1743), USA

Outcome	Estimate	95 % CI	*P* value
Neighbourhood SES as the independent variable*			
A. Soda:milk price ratio†			
Unadjusted for covariates‡	−0·0059	−0·0087, −0·0030	<0·0001
Adjusted for race, age, region, urbanicity, supermarket density and TP price§	−0·0060	−0·0094, −0·0025	0·0006
B. Soda†			
Unadjusted for covariates‡	0·0037	0·0027, 0·0046	<0·0001
Adjusted for race, age, region, urbanicity, supermarket density, TP price and milk price§	0·0024	0·0012, 0·0035	<0·0001
C. Milk†			
Unadjusted for covariates‡	0·0176	0·0139, 0·0213	<0·0001
Adjusted for race, age, region, urbanicity, supermarket density, TP price and soda price§	0·0149	0·0104, 0·0193	<0·0001
Race as the independent variable (proportion non-white Hispanic or black)*			
A. Soda:milk price ratio†			
Unadjusted for covariates‡	−0·0024	−0·0047, −0·0001	0·0384
Adjusted for neighbourhood SES, age, region, urbanicity, supermarket density and TP price§	−0·0065	−0·0094, −0·0036	<0·0001
B. Soda†			
Unadjusted for covariates‡	−0·0027	−0·0034, −0·0019	<0·0001
Adjusted for neighbourhood SES, age, region, urbanicity, supermarket density, TP price and milk price§	−0·0014	−0·0023, −0·0004	0·0043
C. Milk†			
Unadjusted for covariates‡	−0·0030	−0·0060, 0·0000	0·0504
Adjusted for neighbourhood SES, age, region, urbanicity, supermarket density, TP price and soda price§	0·0072	0·0035, 0·0109	0·0001

TP, toilet paper.
*Neighbourhood SES is a standardized variable evaluated as a *Z*-score; the proportion black/non-white Hispanic is operationalized as an ordinal variable: quintiles cut-offs based on the distribution within the sample.
†A negative soda:milk price ratio indicates a larger differential between the two products and may be due to lower soda prices and/or higher milk prices. Soda and milk prices are reported in 2011 US dollar amounts and thus effect estimates can be interpreted as the change in dollars.
‡While no covariates were included, results are adjusted for county and state via nesting.
§Age is the proportion of individuals aged 20 to 39 years, standardized as a *Z*-score; urbanicity is a categorical variable with values of large metropolitan area (population ≥1 million), small metropolitan area (<1 million), micropolitan area (population 10 000–50 000) and non-core area (population <10 000); supermarket density is an ordinal variable with values from 1 to 4 indicating one to four supermarkets in the census tract area, respectively, and a value of 5 indicating five or more supermarkets.

and soda was $US 0·40/serving or soda was 37·9 % of the price of milk, on average. Milk prices had more dispersion than soda prices and tended to drive the variability in milk: soda price differentials across supermarkets. The Midwest and Northeast regions also had the lowest price differential, due to the lower price of milk in these regions. Large metropolitan areas had the highest differential, driven by prices of milk that were 7–13 % higher on average than in other areas.

Compared with the highest-SES neighbourhoods, the lowest-SES neighbourhoods had slightly lower prices of soda, much lower prices of milk and a smaller differential between the price of soda and milk (Fig. 1). Areas with the highest concentration of blacks and Hispanics had a larger difference in the price between soda and milk than areas with fewer minorities, due to significantly higher milk prices in areas with more Hispanic/black persons. The Spearman rank correlation coefficient between race (as a proportion) and SES was 0·41 ($P < 0·001$).

Model results

After adjusting for region, urbanicity and toilet paper cost at the store, variance decomposition of the price ratio suggests that most of the store-level variability was accounted for by state-level factors (53 %), 13 % was accounted for by county-level factors (within each state) and the remaining 34 % was due to price ratio variation between stores (within each county; Table 2). The variation in milk price was similar, with 48 % explained at the state, 12 % at the county level, and the remaining 40 % was between stores. Soda, on the other hand, had much less variation explained by the state (17 %) and county (23 %) levels, with the majority explained by between-store variation within counties (60 %).

Table 3 shows the estimated associations of neighbourhood SES and proportion Hispanic/black with three outcomes: price of soda relative to milk, soda alone and milk alone. The direction of the adjusted results was similar to the gradient seen in the descriptive results of Table 1. In all adjusted models, a lower price of soda relative to milk was associated with a higher concentration of Hispanic/black and higher SES. For example, after adjusting for region, urbanicity, supermarket density, toilet paper price and age distribution, the price of soda relative to milk was 0·65 % lower for a one quintile increase in the proportion of Hispanic/black and 0·60 % lower for a one standard deviation increase in SES (Table 3, A). In order to assess whether the association between the price ratio and proportion Hispanic/black and the price ratio and SES was consistent across urbanicity, *post hoc* interactions were tested with county urbanicity. The difference in the price of soda relative to milk persisted, although it was marginally weaker in rural areas with lower concentrations of Hispanic/black (P for interaction = 0·05, df = 3; data not shown) due to the lower price of milk in those areas.

The association between SES and the price of soda relative to milk did not differ by urbanicity (P for interaction = 0·89, df = 3; data not shown).

When the outcome was price of soda alone and milk alone, adjusted results were similar to the descriptive results reported in Table 1. Lower SES and higher proportions of Hispanic/black persons were associated with lower soda prices (Table 3, B). Although statistically significant, the magnitude of the association was very small: $US 0·002 higher and $US 0·001 lower price of 12 fl. oz of soda per one standard deviation increase in SES and one quintile increase in the proportion Hispanic or black, respectively. In higher-SES areas and neighbourhoods with a larger proportion of Hispanic/black individuals, the price of milk was also higher (Table 3, C). The magnitude was somewhat larger but also small: $US 0·015 higher (or 2·4 % of the average price per serving) and $US 0·007 higher (or 1·1 % of the average price per serving) price of 12 fl. oz of milk per one standard deviation increase in SES and one quintile increase in the proportion Hispanic/black, respectively.

Discussion

The present study explored the association of the price of soda and milk with neighbourhood sociodemographic characteristics and found that although there did not appear to be a strong association between SES and the relative price of soda compared with milk, there was evidence that stores in neighbourhoods with a higher proportion of black and non-white Hispanics sell soda that is both cheaper overall and less expensive compared with milk than in comparable neighbourhoods with fewer non-white Hispanics and blacks. The absolute price of soda was lower in low-SES neighbourhoods compared with high-SES neighbourhoods and while there was a slightly reduced differential between the price of milk and soda in low-SES neighbourhoods, a serving of soda remained just 38 % of the price of a serving of milk. The present study is one of the first to examine price variation between stores located within a large geographic area and variation in price according to neighbourhood-level demographics.

It is known that demand for and purchases of a particular food or beverage are influenced by the price of the item and the price of potential substitutes[53], and that differences in price affect the behaviour of some individuals more than others. Specifically, those with lower income spend a greater proportion of their income on food[35] and are more sensitive to price changes and relative prices between products than those who are wealthier[34]. The aforementioned conditions and the findings from the current study, that one serving of milk was 2·5 times more expensive than a serving of soda, may partly explain higher soda consumption among low-SES individuals[3]. While soda was also cheap in higher-SES
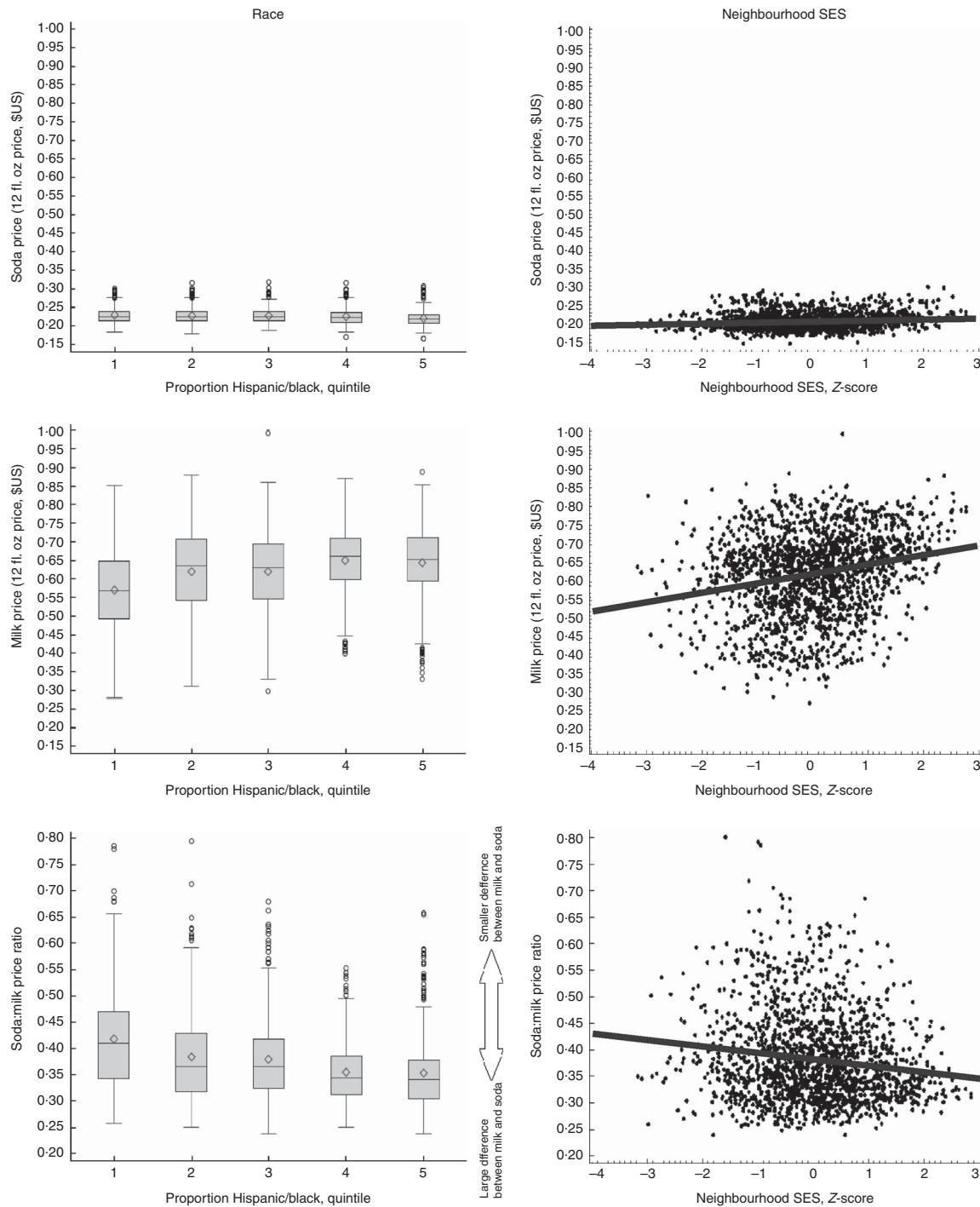
Fig. 1 Soda price (top row), milk price (middle row) and the soda:milk price ratio (bottom row) by race (proportion Hispanic/black) and neighbourhood socio-economic status (SES) in large chain supermarkets (*n* 1743), USA. The box-and-whisker plots (left column) display the distribution of soda price, milk price and the soda:milk price ratio by the proportion of Hispanic/black quintile: the bottom and top edge of the box represent the first and third quartiles (interquartile range (IQR)); the line (diamond) within the box represents the median (mean); the ends of the bottom and top whiskers represent the minimum and maximum values; and the circles represent outliers (values above the 75th percentile plus 1·5 times the IQR and values below the 25th minus 1·5 times the IQR). The scatter plots (right column) display the bivariate association between beverage prices and neighbourhood SES Z-score. The lines in the scatter plots are regression lines from the unadjusted bivariate comparison between price and SES. Prices (per 12 fl. oz (~355 ml) serving of soda and milk; reported in 2011 US dollars) are unadjusted for covariates

areas, higher-SES individuals may be less affected by the large price gap due to their ability to afford more expensive beverages. This conjecture relies on the assumption that lower-SES individuals are primarily living in lower-SES neighbourhoods and higher-SES individuals in higher-SES neighbourhoods, a trend that has been growing as

levels of income inequality have risen during the past 30 years[54].

The positive association between SES and the price of soda relative to milk was contrary to our original expectation. This appears to have been driven in part by higher milk prices in urbanized/metropolitan areas, leading to a lower relative price of soda. The high price of milk in these areas may be due to higher property costs, higher cost of living and higher costs of doing business. However, it would be expected that the price of soda would also be higher if these factors were the main price drivers, but instead soda prices were slightly lower in large metropolitan areas. Thus, higher milk prices in cities were likely due to milk price regulation and distribution costs. The price of milk is regulated at the federal and/or state level by setting price floors as well as allowable relative increases in milk price within marketing areas ('regulated price differentials')[55,56]. Regulated price differentials are specified for every county within a milk marketing area and generally increase the closer a county is to a major consumption area, resulting in higher prices in more populated areas such as cities[55]. Lower milk prices outside large metropolitan areas may be due to lower distribution costs in dairy-producing areas, which are typically rural.

The present study found milk more expensive than soda in neighbourhoods with a higher proportion of black and Hispanic individuals. However, effect sizes were small and thus may suggest that price has only a low to moderate influence on disparities in soda consumption between white and Hispanic/black populations. A systematic review by Andreyeva et al. in 2010 found that the elasticity of soft drinks is 0·79 (95 % CI 0·33, 1·24) or higher, depending on the items used to define soft drinks, while the elasticity of milk is slightly weaker at 0·59 (95 % CI 0·40, 0·79)[53]. If we apply the above elasticities to our results, consumption may be 1·9 % higher and 2·7 % lower, for soda and milk respectively, as concentrations of Hispanics and blacks rise by 53 percentage points (from 4 % as the median in the first quintile to 57 % in the fifth). Black and Hispanic adults, however, obtain roughly 58 % more of their energy from sugar-sweetened beverages than others[2]; much greater than the 2 % that may be explained from soda price differences found in the present study. It is likely that price is just one of many factors that contribute to higher consumption of sugar-sweetened beverages among minorities; for example, advertising, product promotion and physical access may promote soda consumption in minority communities[57–61].

The price data used for the present study were limited to a subset of large chain supermarkets and superstores. Supermarkets dominate retail food sales, accounting for 63 % of all retail food sales in 2009–2012[62]. Not included were smaller/non-chain grocers and convenience/corner stores. Very little information has been published on the price of milk and soda at individual stores throughout the USA, thus it is impossible to know how much our results generalize to other food retail venues.

Due to the small number of products in the data set, these analyses were limited to a single unhealthy beverage category (sugar-sweetened soda) and a single healthy beverage category (milk). Because diet soda was equivalent in price to sugared soda (for equivalent brand and size), results would have been roughly equivalent if the unhealthy category had used diet soda rather than regular soda (see 'Methods' section). Diet soda was not considered for the healthier beverage category because it offers no nutritional value and is not widely viewed as a healthful product. If available data exist, future work examining differences in unhealthy v. healthy beverages could include a wider range of unhealthy beverages (sugar-sweetened sports drinks and flavoured waters) and healthier beverages (water, 100 % juices), thereby improving the representativeness of the healthy and unhealthy beverage categories. Furthermore, we limited our analyses to branded products only in order to increase comparability across beverage categories and across stores. This likely had little effect on the price of soda, given that 95 % of soda sales were for branded items; however, 47 % of milk sales were for non-branded items.

Results of the present study are consistent with previous research that found, in general, that healthier foods are more expensive than unhealthy foods[36–40]. For example, one study showed that milk costs roughly five times more than sugar on a per calorie basis[37]. Overall cost differentials between milk v. soda can in part be explained by differences in production and distribution[63]. Absolute soda prices in our study were similar to prices from the Council for Community and Economic Research (C2ER, formerly known as ACCRA)[64,65]. The absolute price of branded milk in our study was similar to Nielsen Homes can branded milk data ($US 0·56 v. $US 0·63 in our study)[66] but was higher than C2ER data[64] ($US 0·42 per serving v. $US 0·63 in our study), due to C2ER's inclusion of private label and branded whole milk, as opposed to only branded items included in the current study. Nevertheless, available data from C2ER affirm that soda is much cheaper than milk (average values from C2ER suggest one-half the price of milk). It is unlikely that branded milk strongly affected the association of milk price with race and SES characteristics because the price of branded and non-branded items would likely vary similarly across stores and geographic areas due to underlying pricing structures that affect all milk and dairy products (e.g. production, distribution, perishability/refrigeration, etc.) within a given store.

While literature exists on the demographics of soda and milk consumers[2,3,20,21], little was previously known about how area-level demographics relate to the price of these products, which may affect purchasing and consumption patterns. Although it does not appear that lower-SES individuals are exposed to lower prices of soda relative to milk, there appears to be an association with

the concentration of Hispanic and black individuals. As obesity rates are higher within black and Hispanic populations[16], future research may examine if this is due in part to pricing structures that make unhealthy foods more affordable for minority populations compared with their white counterparts.

Conclusions

To our knowledge, the present study is the first to link stores to their immediate neighbourhoods to understand the relationship between neighbourhood demographics and prices of individual products; specifically, soda and milk.

High consumption of soda and other sugar-sweetened beverages has been linked to adverse health outcomes, most notably obesity and type 2 diabetes. If soda consumption can be reduced and replaced by a healthy alternative, such as milk or water, the increasing prevalence of obesity and diabetes may be slowed[7,10,67,68]. Since higher prices discourage consumption and soda is priced well below many of its potential substitutes, policy makers may wish to take measures such as taxing sugar-sweetened beverages or reforming milk price regulations. Decreasing soda consumption would be especially beneficial in black and Hispanic neighbourhoods where prices are lowest, soda consumption is highest[2] and obesity rates are greatest[16].

Acknowledgements

References

1. Euromonitor (2015) Statistics: Total volume, soft drink (carbonates), 2014 sales, per capita. http://www.portal.euromonitor.com/portal/statistics/tab (accessed December 2015).
2. Ogden CL, Kit BK, Carroll MD *et al.* (2011) *Consumption of Sugar Drinks in the United States, 2005–2008. NCHS Data Brief* no. 71. Hyattsville, MD: National Center for Health Statistics.
3. Han E & Powell LM (2013) Consumption patterns of sugar sweetened beverages in the United States. *J Acad Nutr Diet* **113**, 43–53.
4. US Department of Agriculture & US Department of Health and Human Services (2015) Dietary Guidelines for Americans, 2015–2020, 8th ed. http://health.gov/dietaryguidelines/2015/guidelines/ (accessed June 2016).
5. Cocal-Cola (2014) Product Nutrition Information. http://productnutrition.thecoca-colacompany.com/ (accessed December 2014).
6. Malik VS & Hu FB (2012) Sweeteners and risk of obesity and type 2 diabetes: the role of sugar-sweetened beverages. *Curr Diab Rep* **12**, 195–203.
7. Schulze MB, Manson JE, Ludwig DS *et al.* (2004) Sugar-sweetened beverages, weight gain, and incidence of type 2 diabetes in young and middle-aged women. *JAMA* **292**, 927–934.
8. de Koning L, Malik VS, Rimm EB *et al.* (2011) Sugar-sweetened and artificially sweetened beverage consumption and risk of type 2 diabetes in men. *Am J Clin Nutr* **93**, 1321–1327.
9. Malik VS, Popkin BM, Bray GA *et al.* (2010) Sugar-sweetened beverages and risk of metabolic syndrome and type 2 diabetes: a meta-analysis. *Diabetes Care* **33**, 2477–2483.
10. Ludwig DS, Peterson KE & Gortmaker SL (2001) Relation between consumption of sugar-sweetened drinks and childhood obesity: a prospective, observational analysis. *Lancet* **357**, 505–508.
11. Forshee RA, Anderson PA & Storey ML (2008) Sugar-sweetened beverages and body mass index in children and adolescents: a meta-analysis. *Am J Clin Nutr* **89**, 441–442.
12. Malik VS, Schulze MB & Hu FB (2006) Intake of sugar-sweetened beverages and weight gain: a systematic review. *Am J Clin Nutr* **84**, 274–288.
13. Dhingra R, Sullivan L, Jacques PF *et al.* (2007) Soft drink consumption and risk of developing cardiometabolic risk factors and the metabolic syndrome in middle-aged adults in the community. *Circulation* **116**, 480–488.
14. Nguyen S, Choi HK, Lustig RH *et al.* (2009) Sugar sweetened beverages, serum uric acid, and blood pressure in adolescents. *J Pediatr* **154**, 807–813.
15. de Koning L, Malik VS, Kellogg MD *et al.* (2012) Sweetened beverage consumption, incident coronary heart disease, and biomarkers of risk in men. *Circulation* **125**, 1735–1741, S1.
16. Ogden CL, Carroll MD, Kit BK *et al.* (2014) Prevalence of childhood and adult obesity in the United States, 2011–2012. *JAMA* **311**, 806–814.
17. Evans JM, Newton RW, Ruta DA *et al.* (2000) Socio-economic status, obesity and prevalence of type 1 and type 2 diabetes mellitus. *Diabet Med* **17**, 478–480.
18. McWilliams JM, Meara E, Zaslavsky AM *et al.* (2009) Differences in control of cardiovascular disease and diabetes by race, ethnicity, and education: US trends from 1999 to 2006 and effects of medicare coverage. *Ann Intern Med* **150**, 505–515.
19. Centers for Disease Control and Prevention (2014) *National Diabetes Statistics Report: Estimates of Diabetes and Its Burden in the United States, 2014.* Atlanta, GA: US Department of Health and Human Services.
20. Krebs AL (2013) Beyond the gallon jug: how can innovation change the path of fluid milk? CoBANK Knowledge Exchange, June 2013 issue. http://www.cobank.com/Knowledge-Exchange/~/media/Files/Unsearchable%20Files/Knowledge%20Exchange/KEFluidMilkReportJune2013.pdf (accessed June 2016).
21. Sebastian RS, Goldman JD, Enns CW *et al.* (2010) *Fluid Milk Consumption in the United States: What We Eat in America, NHANES 2005–2006. FSRG Dietary Data Brief* no. 3. Beltsville, MD: US Department of Agriculture,

Agricultural Research Service, Beltsville Human Nutrition Research Center, Food Surveys Research Group; available at http://www.ars.usda.gov/SP2UserFiles/Place/80400530/pdf/DBrief/3_milk_consumption_0506.pdf

22. Afeiche M, Williams PL, Mendiola J et al. (2013) Dairy food intake in relation to semen quality and reproductive hormone levels among physically active young men. Hum Reprod 28, 2265–2275.

23. Chan JM, Stampfer MJ, Ma J et al. (2001) Dairy products, calcium, and prostate cancer risk in the Physicians' Health Study. Am J Clin Nutr 74, 549–554.

24. Nachman KE & Smith TJ (2015) Hormone use in food animal production: assessing potential dietary exposures and breast cancer risk. Curr Environ Health Rep 2, 1–14.

25. Choi HK, Willett WC, Stampfer MJ et al. (2005) Dairy consumption and risk of ype 2 diabetes mellitus in men. Arch Intern Med 165, 997–1003.

26. Liu S, Choi HK, Ford E et al. (2006) A prospective study of dairy intake and the risk of type 2 diabetes in women. Diabetes Care 29, 1579–1584.

27. Pereira MA, Jacobs DR, Van Horn L et al. (2002) Dairy consumption, obesity, and the insulin resistance syndrome in young adults: the CARDIA Study. JAMA 287, 2081–2089.

28. Elwood PC, Givens DI, Beswick AD et al. (2008) The survival advantage of milk and dairy consumption: an overview of evidence from cohort studies of vascular diseases, diabetes and cancer. J Am Coll Nutr 27, issue 6, 723S–734S.

29. Elwood PC, Pickering JE, Givens DI et al. (2010) The consumption of milk and dairy foods and the incidence of vascular disease and diabetes: an overview of the evidence. Lipids 45, 925–939.

30. Kratz M, Baars T & Guyenet S (2013) The relationship between high-fat dairy consumption and obesity, cardiovascular, and metabolic disease. Eur J Nutr 52, 1–24.

31. Maersk M, Belza A, Holst JJ et al. (2012) Satiety scores and satiety hormone response after sucrose-sweetened soft drink compared with isocaloric semi-skimmed milk and with non-caloric soft drink: a controlled trial. Eur J Clin Nutr 66, 523–529.

32. Dove ER, Hodgson JM, Puddey IB et al. (2009) Skim milk compared with a fruit drink acutely reduces appetite and energy intake in overweight men and women. Am J Clin Nutr 90, 70–75.

33. Fletcher JM, Frisvold DE & Tefft N (2010) The effects of soft drink taxes on child and adolescent consumption and weight outcomes. J Public Econ 94, 967–974.

34. Powell LM & Chaloupka FJ (2009) Food prices and obesity: evidence and policy implications for taxes and subsidies. Milbank Q 87, 229–257.

35. Frazao E, Andrews M, Smallwood D et al. (2007) Food Spending Patterns of Low-Income Households: Will Increasing Purchasing Power Result in Healthier Food Choices? Economic Information Bulletin no. 29–4. Washington, DC: US Department of Agriculture, Economic Research Service.

36. Drewnowski A (2010) The cost of US foods as related to their nutritive value. Am J Clin Nutr 92, 1181–1188.

37. Drewnowski A & Darmon N (2005) Food choices and diet costs: an economic analysis. J Nutr 135, 900–904.

38. Lipsky LM (2009) Are energy-dense foods really cheaper? Reexamining the relation between food price and energy density. Am J Clin Nutr 90, 1397–1401.

39. Andrieu E, Darmon N & Drewnowski A (2006) Low-cost diets: more energy, fewer nutrients. Eur J Clin Nutr 60, 434–436.

40. Monsivais P, McLain J & Drewnowski A (2010) The rising disparity in the price of healthful foods: 2004–2008. Food Policy 35, 514–520.

41. Information Resources, Inc. (2015) Academic Data Set. http://www.iriworldwide.com/solutions/Academic-Data-Set (accessed June 2015).

42. Bronnenberg BJ, Kruger MW & Mela CF (2008) Database Paper – The IRI Marketing Data Set. Market Sci 27, 745–748.

43. Bleich SN, Wolfson JA, Vine S et al. (2014) Diet-beverage consumption and caloric intake among US adults, overall and by body weight. Am J Public Health 104, e72–e78.

44. Nettleton JA, Lutsey PL, Wang Y et al. (2009) Diet soda intake and risk of incident metabolic syndrome and type 2 diabetes in the Multi-Ethnic Study of Atherosclerosis (MESA). Diabetes Care 32, 688–694.

45. US Department of Labor, Bureau of Labor Statistics (2014) Consumer Price Index – All Urban Consumers, Series Id: CUUR0000SA0. http://www.bls.gov/cpi/home.htm (accessed August 2014).

46. Grocery Headquarters (2014). Wiping out the competition. http://www.groceryheadquarters.com/2014/11/paper-products (accessed June 2014).

47. Babey SH, Diamant AL, Hastert TA et al. (2008) Designed for Disease: The Link Between Local Food Environments and Obesity and Diabetes. California Center for Public Health Advocacy, PolicyLink and the UCLA Center for Health Policy Research. April 2008. http://healthpolicy.ucla.edu/publications/Documents/PDF/Designed%20for%20Disease%20The%20Link%20Between%20Local%20Food%20Environments%20and%20Obesity%20and%20Diabetes.pdf (accessed June 2016).

48. Diez Roux AV, Stein Merkin S, Arnett D et al. (2001) Neighborhood of residence and incidence of coronary heart disease. N Engl J Med 345, 99–106.

49. US Department of Agriculture, Economic Research Service (2013) Urban Influence Codes: Documentation. http://www.ers.usda.gov/data-products/urban-influence-codes/documentation.aspx (accessed June 2015).

50. Grimm KA, Moore LV & Scanlon KS (2013) Access to healthier food retailers – United States, 2011. MMWR Suppl 62, issue 3, 20–26.

51. Lamichhane AP, Warren J, Puett R et al. (2013) Spatial patterning of supermarkets and fast food outlets with respect to neighborhood characteristics. Health Place 23, 157–164.

52. Merlo J, Chaix B, Yang M et al. (2005) A brief conceptual tutorial of multilevel analysis in social epidemiology: linking the statistical concept of clustering to the idea of contextual phenomenon. J Epidemiol Community Health 59, 443–449.

53. Andreyeva T, Long MW & Brownell KD (2010) The impact of food prices on consumption: a systematic review of research on the price elasticity of demand for food. Am J Public Health 100, 216–222.

54. Fry R & Taylor P (2012) The Rise of Residential Segregation by Income. Pew Research Center Social & Demographic Trends. http://www.pewsocialtrends.org/2012/08/01/the-rise-of-residential-segregation-by-income/ (accessed June 2016).

55. Jesse E & Cropp B (2008) Basic Milk Pricing Concepts for Dairy Farmers. University of Wisconsin Extension – Cooperative Extension, Publication a3379. http://future.aae.wisc.edu/publications/a3379.pdf (accessed June 2016).

56. Manchester AC & Blayney D (2001) Milk Pricing in the United States. Agriculture Information Bulletin no. AIB-761. Washington, DC: US Department of Agriculture, Economic Research Service; available at http://www.ers.usda.gov/media/306483/aib761_1_.pdf

57. Powell LM, Wada R & Kumanyika SK (2014) Racial/ethnic and income disparities in child and adolescent exposure to food and beverage television ads across the US media markets. Health Place 29, 124–131.

58. Grier SA & Kumanyika SK (2008) The context for choice: health implications of targeted food and beverage marketing to African Americans. Am J Public Health 98, 1616–1629.

59. Powell L, Rimkus L, Isgor Z *et al.* (2012) *Exterior Marketing Practices of Fast-Food Restaurants – A BTG Research Brief*. Chicago, IL: Bridging the Gap Program, Health Policy Center, Institute for Health Research and Policy, University of Illinois at Chicago; available at www.bridgingthegap research.org

60. Fleischhacker SE, Evenson KR, Rodriguez DA *et al.* (2011) A systematic review of fast food access studies. *Obes Rev* **12**, e460–e471.

61. Larson NI, Story MT & Nelson MC (2009) Neighborhood environments: disparities in access to healthy foods in the US. *Am J Prev Med* **36**, 74–81.

62. US Department of Agriculture, Economic Research Service (2014) Food Expenditures – Table 15. http://www.ers.usda. gov/data-products/food-expenditures.aspx (accessed June 2016).

63. US Department of Agriculture, Economic Research Service (2015) Milk Cost of Production Estimates – 2010 Base. http://www.ers.usda.gov/data-products/milk-cost-of-production-estimates.aspx (accessed December 2015).

64. Duffey KJ, Gordon-Larsen P, Shikany JM *et al.* (2010) Food price and diet and health outcomes: 20 years of the CARDIA Study. *Arch Intern Med* **170**, 420–426.

65. Wada R (2015) Associations between soda prices and intake: evidence from 24-h dietary recall data. *Food Policy* **55**, 54–60.

66. Zhuang Y, Dimitri C & Jaenicke EC (2009) Consumer choice of private label or national brand: the case of organic and non-organic milk. Paper presented at the *AAEA & ACCI 2009 Joint Annual Meeting*, Milwaukee, WI, USA, 26–28 July 2009.

67. Hu FB & Malik VS (2010) Sugar-sweetened beverages and risk of obesity and type 2 diabetes: epidemiologic evidence. *Physiol Behav* **100**, 47–54.

68. Berkey CS, Rockett HR, Field AE *et al.* (2004) Sugar-added beverages and adolescent weight change. *Obes Res* **12**, 778–788.