KUSUM L. AILAWADI, YU MA, and DHRUV GREWAL*

This article studies the impact of shopping at the warehouse club format on households' purchases of packaged food for the home. In addition to low prices, this format has several unique characteristics that can influence packaged food purchases. The empirical analysis uses a combination of households' longitudinal grocery purchase information, rich survey data, and detailed item-level nutrition information. After accounting for selection on observables and unobservables, the authors find a substantial increase in the total quantity (servings per capita) of packaged food purchases attributable to shopping at this format. Because there is no effect on the nutritional quality of purchases, this translates into a substantial increase in calories, sugar, and saturated fat per capita. The increase comes primarily from storable and impulse foods, and it is drawn equally from foods that have positive and negative health halos. The results have important implications for how marketers can create win–win opportunities for themselves and for consumers.

Keywords: packaged food purchases, nutrition, public health, warehouse club format, selection on observables and unobservables

Online Supplement: http://dx.doi.org/10.1509/jmr.16.0235

The Club Store Effect: Impact of Shopping in Warehouse Club Stores on Consumers' Packaged Food Purchases

As the incidence of obesity-related health problems increases around the globe, marketers of packaged food are facing heightened scrutiny. Packaged food purchased for home consumption accounts for a substantial percentage of consumers' total food intake. According to the Economic Research Service of the U.S. Department of Agriculture (USDA 2014), 74% of added sugar intake and 65% of fat intake comes from food consumed at home. Because the shopping environment influences purchase decisions, the characteristics of grocery retail formats where consumers buy food for their homes have important roles in determining what and how much they buy (and presumably consume).

The retail formats where U.S. consumers buy groceries range from convenience stores to supermarkets to supercenters and warehouse clubs. The "big box" supercenter and warehouse club (also called "club store") formats achieved the fastest growth rate of any retail category between 1992 and 2013. Total U.S. sales of these two formats grew from $40 billion to $420 billion (Hortascu and Syverson 2015), and as of 2013, the club store format was a $148 billion industry, with more than 1,600 stores. The format is also expanding outside the United States. At the end of 2016, Costco and Sam's Club operated 221 and 212 international locations, respectively, and smaller warehouse club retailers PriceSmart and Cost-U-Less had stores in the Latin American, Caribbean, and Pacific regions (Warehouse Club Intelligence Center 2017). Overall,

*Kusum L. Ailawadi is Charles Jordan 1911 TU'12 Professor of Marketing, Tuck School of Business, Dartmouth College (email: kusum.ailawadi@dartmouth.edu). Yu Ma is Associate Professor of Marketing and Bensadoun Scholar, Desautels Faculty of Management, McGill University (email: yu.ma@mcgill.ca). Dhruv Grewal is Toyota Chair of Commerce and Electronic Business and Professor of Marketing, Babson College (email: dgrewal@babson.edu). The authors thank SymphonyIRI for providing the panel and survey data used in this research. They thank Minha Hwang for sharing warehouse club store opening data. They also thank seminar participants at several universities, especially KU Leuven, Southern Methodist University, IDC Herzeliya, McGill University, University of Miami, Rice University, Syracuse University, and University of Hamburg, for their helpful suggestions. The second author was supported by grants from the Social Science and Humanities Research Council of Canada. Marnik Dekimpe served as associate editor for this article.

this format is a sizeable and growing part of the retail sector, making it important to understand how its characteristics influence purchase behavior.

Like other big box formats, club stores offer lower prices, stock higher-ticket durable products in addition to packaged goods, and stock a greater proportion of packaged foods than fresh foods. Inman, Shankar, and Ferraro (2004) document that club stores are most strongly associated with bulk and frozen foods, snacks, and pet food. However, this format is unique in three main respects. First, the package sizes sold at club stores are much larger than those in other formats, so shoppers must be willing and able to buy in bulk. Second, the density of club stores is lower than that of other formats. Third, club stores charge an annual membership fee. In the United States, basic membership costs around \$50, increasing to around \$100 for a higher tier that comes with rewards; it generally is open to both consumers and businesses.

Our examination of home-scan panel purchase data from Symphony IRI bears out these features of the format. As shown in the Appendix, on average, club store prices are 22.5% lower, package sizes are over three times larger, and travel distance is almost twelve times longer compared with traditional supermarkets. Furthermore, while both formats generally stock the same categories of packaged food, the number of items per category is much lower in club stores—only 15.5% of the number in supermarkets. However, there is no difference overall in the nutritional quality of specific packaged food items sold in the two formats, as measured by calories, saturated fat, and sugar per serving.

These characteristics have important ramifications for how shopping at club stores might affect households' total packaged food purchases. For instance, larger package sizes imply higher purchase quantities; membership fees and long travel distance become sunk costs that consumers justify with frequent visits and larger shopping baskets, respectively. Depending on whether the larger quantities purchased in club stores accelerate consumption or simply substitute for purchases in other formats, total packaged food purchases may or may not increase. Savings from the low prices in club stores may also enable consumers to afford healthier fresh food options. Therefore, the effect of shopping in club stores on a household's total quantity of packaged food purchases is an empirical question.

Of course, consumers who choose to shop at a club store can only buy what is available there, so the nutritional quality of the available assortment is important. One aspect of this is the relative emphasis on more calorie-dense packaged foods instead of fresh foods. Any substitution between these will be reflected in the total quantity of packaged food purchases. Another aspect is the quality of items within packaged foods. Because the assortment of packaged food in club stores is no different in nutritional quality than the assortment in other formats, shopping in club stores should affect the quality of packaged food purchases only if households selectively choose healthier or less healthy packaged food items when they shop there. This, too, is an empirical question.

We tackle these issues in depth by addressing three research questions in this article. First, how does shopping in a club store affect the quantity (e.g., servings per capita), nutritional quality (e.g., calories per serving), and quantity–quality combination (e.g., calories per capita) of the total packaged food purchased by households for home consumption, across all formats? Although quantity and quality are both important, it is the combination of the two that is of ultimate interest for consumer health. Second, does this impact, which we refer to as the "club store effect," differ by type of food, such as how storable it is or how prone to impulse purchase (which we refer to as its "impulse nature"), or how healthy it is perceived to be? Third, is the club store effect moderated by household characteristics such as family needs, financial ability, and health consciousness?

We use a unique data set compiled from four main sources in our empirical analysis. We combine longitudinal data, covering all packaged food and nonfood purchases from all retail formats for a large panel of U.S. households, with (1) item-level information on servings, calories, saturated fat, and sugar for all packaged food purchases, obtained from the nutritional labels on packages; (2) responses to a demographic, health, and wellness survey of panelists; and (3) a separate online survey that measures the storability, impulse nature, and perceived healthfulness of all packaged food categories.

In the next section, we present the conceptual underpinnings of our work based on relevant research. Next, we present the data and detail our modeling approach. Then, we present our results and conclude with a discussion of our main findings and their implications.

CONCEPTUAL AND EMPIRICAL BACKGROUND

Prior Empirical Research on the Health Impact of the Big Box Format

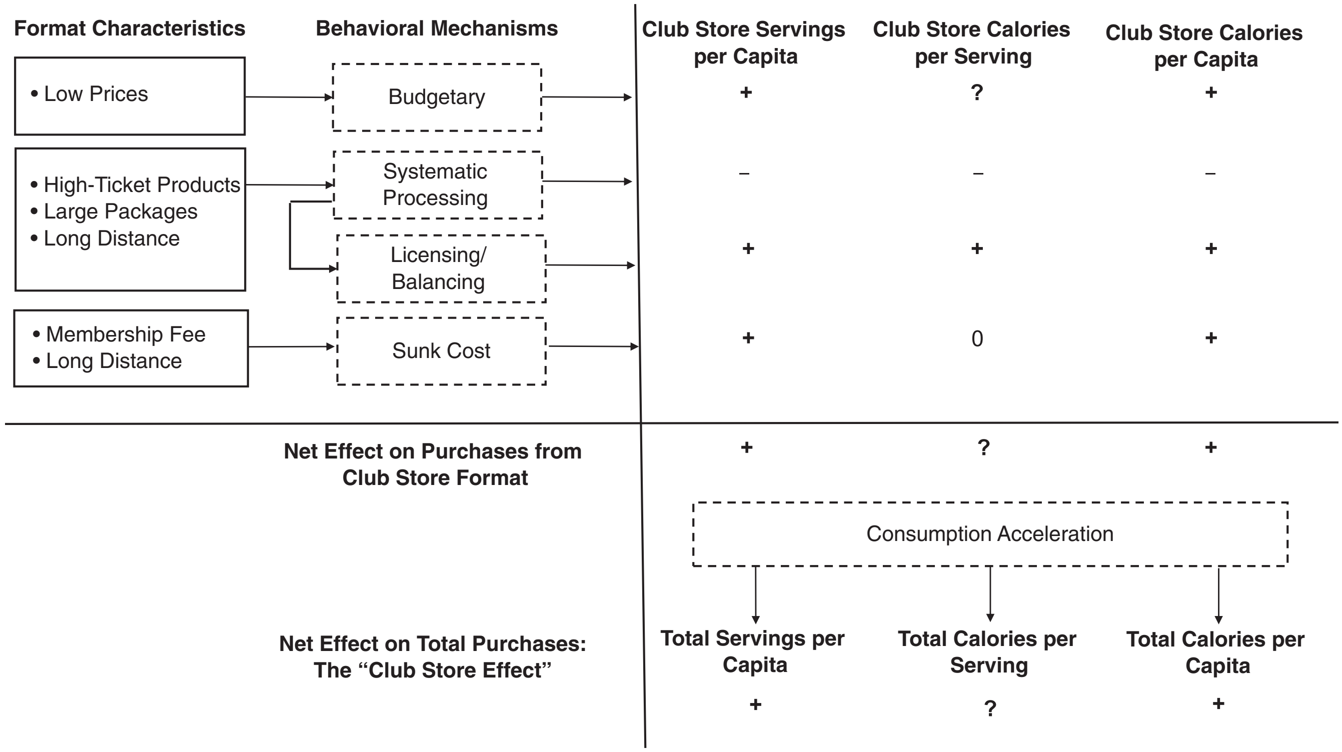
Courtemanche and Carden (2011) examine the impact of Walmart supercenter density on consumers' body mass index (BMI). They estimate a .24-unit increase in BMI and a 2.3 percentage point increase in obesity incidence for every additional supercenter per 100,000 households. Volpe, Okrent, and Leibtag (2013) estimate how the healthfulness of households' shopping baskets, which they approximate with various scoring methods, is affected by supercenter market share. They find a significantly negative elasticity of healthfulness with respect to supercenter share, ranging from –.1 to –.46, in most of the models they estimate.

Both studies thus indicate a negative impact of supercenters on consumer health. However, the independent variables they use are aggregate measures of supercenter presence, not households' actual format choices. Furthermore, they do not provide insight into the sources of the change in healthfulness or the types of foods involved. Nor do they examine which types of household are prone to shop at big box stores or be affected by shopping there. Finally, they study supercenters; as we have noted, club stores differ from supercenters in important ways that are likely to affect purchase behavior. We therefore advance research in this area by focusing on the club store format while also addressing these other issues.

Behavioral Mechanisms for the Club Store Effect

Figure 1 shows how the characteristics of the club store format link to five theory-based behavioral mechanisms —budgetary, systematic processing, licensing/balancing, sunk costs, and consumption acceleration—by which shopping at a club store may affect total packaged food purchases and consumption. We cannot separate the five mechanisms empirically, but we use them to derive predictions about the club store effect on quantity, quality, and the combination of quantity and quality (for an example of this approach, see Geyskens, Gielens, and Dekimpe 2002). For simplicity, we refer here to servings per

Figure 1
MECHANISMS UNDERLYING THE CLUB STORE EFFECT ON TOTAL PACKAGED FOOD PURCHASES



Notes: "0" signifies that there is no reason to expect an effect, whereas "?" signifies that we cannot predict the direction of the effect.

capita, calories per serving, and calories per capita, but our empirical analysis also includes sugar and saturated fat per serving and per capita.

Budgetary mechanism. The lower prices in club stores have income and substitution effects (Courtemanche and Carden 2011). According to the income effect, savings from low prices can be used to buy more as well as healthier items in the club store or other formats. This may increase servings per capita and reduce calories per serving, but the latter effect may be quite limited since budget slack is often spent on impulse purchases, which tend to be more hedonic (Stilley, Inman, and Wakefield 2010). According to the substitution effect, lower prices encourage shoppers to switch away from more expensive products and stores. The prices of both fresh and packaged foods are lower at club stores, but the latter are more prominent, which implies more packaged food servings per capita, especially considering the larger package sizes in club stores. Overall, we expect the budgetary mechanism to have a positive effect on servings per capita; we cannot predict the effect, if any, on calories per serving; and we expect the effect on calories per capita to be positive, driven by the dominant positive effect on servings per capita.

Systematic processing mechanism. Dual-processing models of consumer decision making state that deliberate, or systematic, processing requires more motivation than heuristic, or automatic, processing. Grocery shopping is generally characterized by automaticity (Cohen and Babey 2012). However, because club stores require larger cash outlays to buy in bulk, sell higher-ticket durable products in addition to packaged goods, and entail greater travel distances, consumers are more involved and therefore more likely to be in a systematic processing mode when they visit this format (Celsi and Olson 1988; Chaiken 1980; Petty, Cacioppo, and Schumann 1983). The increased attention can improve choices, resulting in fewer servings per capita, calories per serving, and calories per capita.

Licensing and balancing mechanism. Systematic processing may increase self-control, but it can also lead to licensing and/or balancing (Khan and Dhar 2006; Mukhopadhyay and Johar 2009). When consumers exert self-control in the pursuit of one goal, they balance it by shifting to another goal within the same purchase or consumption episode. Similarly, exerting self-control and feeling virtuous licenses them to indulge later. Both effects can go across domains; that is, the self-control may be in one domain (e.g., health, savings), and the licensing or balancing may be in a very different domain (e.g., enjoyment, luxury). Furthermore, both effects may apply to club store shoppers. For example, as they make planned purchases in systematic processing mode, they may balance that form of self-control with other purchases during the same trip that are more indulgent. They also may feel licensed to buy less healthy items during subsequent trips to other formats. Therefore, this mechanism should have a positive effect on servings per capita, calories per serving, and calories per capita.

Sunk cost mechanism. The sunk cost effect is a "greater tendency to continue an endeavor once an investment in money, effort, or time has been made" (Arkes and Blumer 1985, p. 124). After investing in an annual club store membership, households may shop there frequently, wanting to get the most out of their investment. More visits result in more purchases, and annual renewal also keeps the sunk cost effect from weakening. Long travel distance may constitute a sunk cost as well, such that shoppers increase purchases to make the trip worthwhile. This should have a positive effect on servings and calories per capita though not on calories per serving.

Consumption acceleration mechanism. The preceding mechanisms pertain largely to increased purchases in the club store format. If households stockpile their extra purchases and consume them at a normal rate, club store shopping will simply substitute for purchases in other formats, and total servings and calories purchased (and consumed) will not increase. If trips to other formats are reduced, consumption may even decrease because of the reduced opportunity for impulse purchases. However, we consider this outcome unlikely. For one thing, consumers buy different product categories in different stores and frequently make fill-in trips. Fox, Montgomery, and Lodish (2004) report, for example, that shopping at mass stores does not substitute for trips to other formats. For another, greater household inventory is known to increase consumption (Ailawadi and Neslin 1998; Bell, Chiang, and Padmanabhan 1999), especially of large package sizes (Wansink 1996) that are easy and convenient to consume (Wansink and Deshpandé 1994) or that are salient (Chandon and Wansink 2002). Therefore, consumption acceleration is likely to transform the increased club store purchases into increased total purchases across all formats.

Overall club store effect. Overall, the behavioral mechanisms shown in Figure 1 suggest that the club store effect on total packaged food servings per capita and calories per capita across all formats should be positive. It is unclear what effect there will be on calories per serving.

Drivers of Club Store Patronage and Packaged Food Purchases

A key challenge in estimating the club store effect is that households choose whether to shop at the format, and at least some of the characteristics that drive this decision also influence their packaged food purchases. The resulting selection bias must be controlled for to estimate the effect of club store shopping. If we had such rich data that any variable reasonably expected to affect both behaviors would be observed, the selection problem would disappear (Germann, Ebbes, and Grewal 2015). We cannot make such a strong claim, but we account for as many common drivers of club store shopping and food purchases as possible.

The unique characteristics of the club store format determine the monetary, travel, inventory and other shopping costs associated with shopping there and thus the types of households that are more likely to patronize it. For example, household needs and storage space determine whether it is worthwhile to become a member of a club store and buy in bulk (Bell and Hilber 2006; Bhatnagar and Ratchford 2004). Lower travel distance to the format increases likelihood of patronage (Ailawadi, Pauwels, and Steenkamp 2008; Ma et al. 2011). Time spent away from home decreases at-home food needs and constrains time for shopping, likely reducing club store patronage. Financial ability determines whether the household can afford the membership fee and monetary outlay required to buy in bulk (Orhun and Palazzolo 2017) and whether there is easy access to the format (Talukdar 2008). Health consciousness and health status also could influence the decision to shop at a club store. If health-conscious households and those in good health are motivated to preserve their health (Moorman and Matulich 1993), they may exert self-control through rationing purchase quantities (Wertenbroch 1998) and be less likely to shop in club stores.

Many of these drivers also influence packaged food purchases. For example, household needs are clearly positively associated with servings per capita, and time spent away from home should be negatively associated with it. Financial ability should allow households to buy healthier fresh foods, thus reducing servings of packaged food per capita, calories per serving, and calories per capita from packaged foods. Similar effects should hold for health status and health consciousness (Andrews, Netemeyer, and Burton 2009; Ma, Ailawadi, and Grewal 2013).

Differences in Club Store Effects Across Category Types

The club store effect may vary across different food categories. We examine moderation by three category characteristics: storability, impulse nature, and perceived healthfulness.[1] For each one, we develop expectations according to which of the behavioral mechanisms discussed in the previous section are likely to be particularly influential, and in which direction, for servings per capita and calories per serving. The expectations then combine for calories per capita, as summarized in Table 1 and discussed next.

Moderating effect of storability. We expect the club store effect on servings per capita to be more positive for storable foods. Our reasoning is as follows: Because buying in bulk makes more sense for such foods (Bhatnagar and Ratchford 2004), the increase in servings per capita through the budgetary mechanism should be greater and the decrease in servings per capita through systematic processing should be smaller. Further, many storable foods closely associated with the club store format, such as frozen foods and snacks, are particularly prone to flexible consumption. Frozen foods are salient in the freezer (Chandon and Wansink 2002), and snacks are convenient to consume (Wansink and Deshpandé 1994). There is no reason for storability to moderate the club store effect on nutritional quality. Combining quantity and quality, we expect the calories per capita effect to be more positive for storable foods.

Moderating effect of impulse nature. We expect the effect on servings per capita to be more positive for impulse categories. The rationale is as follows: Savings on planned purchases free up budget slack that can be spent on impulse purchases (Stilley, Inman, and Wakefield 2010), so the positive effect of the budgetary mechanism should be greater. Further, although the negative effect of systematic processing should be stronger for impulse categories, the positive effect of licensing and balancing should be stronger, too (Baumeister 2002; Inman, Winer, and Ferraro 2009). Finally, consumption acceleration may be stronger for impulse foods because they are often more tempting (Hofmann et al. 2010).

People tend to ignore nutrition information in hedonic categories bought on impulse (Balasubramanian and Cole 2002; Ehrich and Irwin 2005; Wansink and Chandon 2006), so licensing/balancing may increase calories per serving more in

[1]We also examined the hedonic nature of a category. Because the conceptual grounding and empirical results for this characteristic overlap substantially with impulse buying, we do not include them here. Results are available on request.

Table 1
EXPECTED MODERATION OF THE CLUB STORE EFFECT BY CATEGORY AND HOUSEHOLD CHARACTERISTICS

	Expected Moderation of Effect on...		
Variable	*Servings per Capita*	*Calories per Serving*	*Calories per Capita*
Category Characteristics			
Storability	+	0	+
Impulse buying	+	+	+
Perceived healthfulness	?	−	?
Household Characteristics			
Household needs	+	0	+
Time away from home	−	0	−
Health consciousness	−	−	−
Health status	−	−	−
Financial ability	?	?	?
Travel distance	?	0	?

Notes: "0" signifies that there is no reason to expect a moderating effect, whereas "?" signifies that we cannot predict the direction of the effect.

impulse categories. Combining quantity and quality, we expect a more positive calories-per-capita effect for impulse categories.

Moderating effect of perceived healthfulness. We cannot predict how perceived healthfulness moderates the club store effect on servings per capita. Budget slack could be allocated to healthier categories that tend to be more expensive, as well as to the impulse purchases of less healthy foods. Systematic processing likely increases purchases of packaged foods that possess a positive health halo (Inman, Winer, and Ferraro 2009), whereas licensing and balancing should increase purchases of foods with a negative health halo. Finally, consumers may accelerate consumption of foods that are perceived to be healthy but may also eat food that is less healthy (and more tasty) without guilt after having chosen healthy options (Chandon and Wansink 2011).

Consumers pay greater attention to nutrition information in categories with a positive health halo (Nikolova and Inman 2015), so systematic processing should reduce calories per serving more in such categories. Combining quantity and quality, we cannot predict the moderating effect of perceived healthfulness on calories per capita.

Differences in Club Store Effect Across Households

Moderating effect of household needs and time away from home. We expect the club store effect on servings per capita to be more positive for households with high needs and less positive for those whose members spend more time away from home. This is because cost savings from buying in bulk may make more sense for the former and less sense for the latter. Also, consumption acceleration is likely to be stronger for large households. One reason is that they purchase a greater variety of foods, given heterogeneous tastes of family members, and variety increases consumption (Kahn and Wansink 2004). Another is that people consume more in the presence of others (Wansink 2004). There is no reason for these variables to moderate the quality effect, so the quantity effect carries over to calories per capita.

Moderating effect of health consciousness and health status. We expect both health consciousness and health status to make the club store effect on servings per capita less positive. The reasoning is that health-conscious people systematically choose what they buy (Andrews, Netemeyer, and Burton 2009) and should be less vulnerable to licensing and to consumption acceleration. Unhealthy people (e.g., those with high BMI) may have a poor self-concept (Bodiba et al. 2008) that leaves them less able to self-regulate against licensing, balancing, and consumption acceleration.

Health consciousness and health status should make the effect on calories per serving less positive (or more negative) because healthy and health-conscious people may particularly avail themselves of the nutritional-quality benefits of systematic processing and be less vulnerable to licensing and balancing effects. Greater susceptibility to the health halo bias (Chernev 2011) also may push such households to accelerate their consumption of higher-quality foods. If we combine quantity and quality, health consciousness and health status should make the club store effect on calories per capita less positive.

Moderating effect of financial ability and travel distance. We cannot predict how financial ability and travel distance moderate the club store effect on servings per capita. Cost savings are less important for financially well-off families, so the budgetary mechanism may be weaker for them. However, their consumption acceleration may be higher because they are less concerned about conserving. Households that travel farther to shop at a club store may plan more, suggesting a more negative effect through systematic processing. But they may also perceive higher sunk costs, making the positive effect through sunk cost stronger.

There is no reason for travel distance to moderate the club store effect on calories per serving. Financial ability may moderate this effect, but the direction is difficult to predict. Financially well-off households can afford higher quality; they do not need to shop in club stores to afford it. Savings matter to financially constrained households, but they may use savings gained from shopping at a club store for other necessities instead of spending them on higher-quality packaged food. Putting quantity and quality together, we cannot predict the moderating effect of financial ability and travel distance on calories per capita.

DATA

Data Sources

We compile a unique data set from several sources for this research. The primary source is a nationwide home-scan panel data set from SymphonyIRI that records households' purchases from all grocery retail formats, including supermarkets, mass merchants, club stores, drug stores, and convenience stores. Households scan their purchases of all bar-coded packaged goods (food and nonfood) at home. We have data from January 2006 to December 2009.

The second source is a health survey administered by SymphonyIRI in November of each year from 2005 to 2008. It asks for information about each household member's health status and health-related perceptions and behaviors. These survey data allow us to control for a rich set of household variables that are relevant to shopping and food choices, beyond the basic demographics that are usually included with panel data.

The third source is a database of the nutrient content of individual stockkeeping units (SKUs) of packaged food, captured from the package nutrition labels. These data were available for 68.5% of SKUs in the SymphonyIRI data, accounting for almost 90% of purchases. For the remaining

SKUs, we impute nutritional content from other SKUs of the same brand. The nutritional data allow us to track the calories, sugar, and fat purchased by households.

Fourth, we conducted an online consumer survey on Amazon's Mechanical Turk to obtain perceived ratings of all 165 IRI food and beverage categories on the three category characteristics of storability, impulse nature, and perceived healthfulness. Each respondent rated one-fifth of the categories, and we obtained 501 responses for each category. Finally, we compiled information on the zip codes of store locations for the three major club store chains from their online store locators, and we obtained data on new store openings from the official news releases of the chains.

Variable Operationalization

Table 2 lists measures of all the variables in our empirical analysis. We begin with the outcome variables. Monthly servings per capita measure quantity of packaged food purchases. Nutritional quality is measured by calories, sugar, and saturated fat per serving. Calories, sugar, and saturated fat per capita combine quantity and quality. In addition, we measure households' total number of shopping trips and total packaged food spending per capita.

Next, we list all available explanatory variables related to the drivers discussed previously, as well as relevant control variables. For instance, household needs are assessed by the number of people in the household, their spending on nonfood categories in an initialization period (households with higher needs for personal care and general household products also have higher needs for food), and the frequency with which they eat in fast-food restaurants (the more they eat out, the less their need to purchase food to be consumed at home). Table 2 also lists the items used to measure the three category characteristics described earlier.

Sample Selection

For our main analysis, we start with all 15,321 households in the panel whose total recorded spending is at least $50 every month during the 2006–2007 period. After deleting households with no health survey data and trimming outliers (bottom 1% and top 5% of monthly calories per capita), we are left with 12,765 households.[2] This sample is split into three groups: those that shop at a club store at least once every three months, those who do so less frequently, and those who do not shop at club stores. The first group, comprising 2,513 households, is our treatment group. The third group of non–club store shoppers, comprising 4,144 households, is the control group pool.

For additional analyses, we identify households (1) who are initially non–club store shoppers with no club store within a 20-mile radius of their home zip code; (2) for whom a new club store opened in that radius between mid-2006 and mid-2009; and (3) for whom we have at least six months of valid purchase data before and after their local store opening. After trimming outliers using the same cutoffs as in the main sample, we have 265 households in this smaller sample. Of these, 83 households started shopping at the club store format after the new store opening and comprise a treatment group. We compare before–after changes of this treatment group with corresponding changes for two control groups in a difference-in-difference analysis.

METHOD

Viewing club store shopping as the "treatment," we need to estimate the treatment effect, which is a counterfactual: What are the total packaged food purchases of households that shop at club stores compared with what they would be if the households did not shop at club stores? We are interested in the average treatment effect (ATE) for all households, as well as in the average treatment effect on the treated (ATT), that is, the effect for club store shoppers. Further, we are interested in estimating the sustained effect of being a regular club store shopper on total purchases, not in the dynamics of weekly inventory, stockpiling, and consumption of individual categories, or in purchases at each format. Therefore, we aggregate weekly household–SKU–chain-level purchase data to the household level over the appropriate time period for our analyses.

As discussed previously, we observe a rich set of drivers of the decision to shop at a club store that also may affect packaged food purchases. We control for these to estimate our treatment effects, but there are likely other unobservable (to the researcher) drivers of the decision to shop at a club store. Some of these unobservables may also be correlated with packaged food purchases and must be controlled for to obtain a consistent estimate of the treatment effect. We estimate four models that control for unobservables to varying extents, thus enabling us to assess the robustness of the estimated treatment effects. Two of the models use the large sample and two use the smaller sample described in the previous section.

Main Sample Models

We compare the 2,513 "treatment" households who shop at club stores with a "control" group from the 4,144 households who do not, using two models. Model 1 is ordinary least squares (OLS), which is consistent under ignorability of treatment, that is, under the assumption that the unobservables are uncorrelated with treatment after conditioning on observables (Wooldridge 2002, p. 607). As shown in Equation 1, each outcome variable, Y_h, is aggregated over the full 2006–2007 period for each household. It is regressed on the treatment dummy (Club Shopper$_h$) and the observable explanatory variables (vector X_h) listed in Table 2.[3] Interactions of the treatment dummy with mean-centered household characteristics allow for potential heterogeneity in the treatment effect:

$$(1) \qquad Y_h = \alpha + X_h\beta + \gamma \text{Club Shopper}_h + \text{Club Shopper}_h \times \left(X_h - \bar{X}\right)\delta + \epsilon_h.$$

The ATE is estimated by γ, and the ATT is computed from γ, δ, and the X vector for the treatment group. We also estimated the models with quadratic and interaction terms between the observables as additional controls and found no substantive change in results.

[2]The home-scan panel comprises households, not businesses. Still, we trim the top 5% in case some people shop not just for personal consumption but also for small businesses. In addition, some households may not scan all their purchases; because we already use a minimum monthly spending threshold, we trim only 1% at the bottom. However, our results are robust to symmetric 5% trimming on both ends.

[3]We also estimated treatment effects using a propensity score–matching model and found very similar results. Complete estimates of that model are provided in Web Appendix A.

Table 2

VARIABLE OPERATIONALIZATION

Construct	Measures
Outcome Variables	
Packaged food quantity	Servings per capita: Total monthly purchases from all major formats (supermarkets, mass discounters, supercenters, club stores, and drug stores) of all Symphony/Information Resources Inc. IRI packaged food categories measured in number of servings, divided by weighted household size. Number of servings of an item is its total volume divided by customary serving size of the category in the *Code of Federal Regulations* (2011). Weighted household size is obtained by applying weights to household members of different age groups to account for their daily caloric requirements relative to 2000 calories per day for an adult (American Heart Association 2016). The weights for age groups of 0–5, 6–11, and 12–17 years are .55, .75, and .95, respectively.
Packaged food quality	Calories, saturated fat, and sugar per serving: Calories (or sugar or saturated fat) per capita divided by servings per capita.
Quantity × Quality	Calories, saturated fat, and sugar per capita: Total monthly purchases of calories, saturated fat, and sugar across all IRI packaged food categories, divided by weighted household size. Calories, saturated fat, and sugar in each item purchased are obtained from nutrition labels.
Shopping behavior	• Number of shopping trips: Total monthly shopping trips across all major formats. • Packaged food spending per capita: Total monthly spending across all major formats on all IRI packaged food categories, divided by weighted household size.
Theory-Based Independent Variables	
Household needs	• Family size: number of people in household. • Personal care spending per capita: Monthly spending on personal care products in initialization period,[a] divided by weighted household size (excluding health remedies, which are incorporated later in a robustness test). • General merchandise spending per capita: Monthly spending on general merchandise in initialization period,[a] divided by weighted household size. • Fast-food frequency: "On a weekly basis how often do you eat at fast-food restaurant?" (1 = "rarely/never" to 3 = "most days").
Time away from home	• Employment dummy 1: Equal to 1 if one household head is employed, 0 if both household heads are employed, and −1 otherwise. • Employment dummy 2: Equal to 1 if both household heads are employed, 0 if one household head is employed, and −1 otherwise.
Financial ability	Income: Household income (1 = "$10K per year or less" to 12 = "more than $100K").
Distance from format	• Supermarket distance: Distance (miles) from the center of household's home zip code to the center of the nearest supermarket's zip code. • Mass store distance: Distance (miles) to the nearest mass discounter or supercenter. • Club store distance: Distance (miles) to the nearest warehouse club store.
Health consciousness	• Nutrition interest: (a) "I often read nutrition labels on food." (1 = "disagree," and 3 = "agree"); (b) "How concerned are you about refined/processed food?" (1 = "not at all," and 3 = "very"); (c) "How concerned are you about transfat in food?" (1 = "not at all," and 3 = "very"). • Education: Highest level of education (1 = "grade school or less," and 8 = "postgraduate work") across household heads. • Healthy behaviors: (a) "On a weekly basis how often do you exercise?" (b) "On a weekly basis how often do you eat late at night?" (c) "On a weekly basis how often do you take multivitamins?" (all on three-point scales, with 1 = "most days," and 3 = "rarely/never"; a and c are reverse-coded).
Health status	• BMI: Average body mass index of household heads. • Perceived health: (a) "My health is..." (1 = "poor," and 4 = "excellent"); (b) "I'm much healthier than most people my age" (1 = "disagree," and 3 = "agree").
Control Variables	
Other control variables	• Single dummy: Equal to 1 if household head is single, −1 otherwise. • Age: Maximum age (in years) of the household heads (1 = 18–29, 2 = 30–34, 3 = 35–44, 4 = 45–54, 5 = 55–64, and 6 = ≥65). • Spending share at mass/supercenters: Share of total spending at the mass/supercenter formats in initialization period.[a] • Race dummy 1: Equal to 1 if African American, 0 if Hispanic, and −1 otherwise. • Race dummy 2: Equal to 1 if Hispanic, 0 if African American, and −1 otherwise. • Region dummy 1: Equal to 1 if Northeast, 0 if South or West, and −1 otherwise. • Region dummy 2: Equal to 1 if South, 0 if Northeast or West, and −1 otherwise. • Region dummy 3: Equal to 1 if West, 0 if Northeast or South, and −1 otherwise.
Instrumental Variable	
Storage space constraints	Rent dummy: Equal to 1 if residence is rented, −1 if owned.
Category Type	
Category characteristics	• Storability: "How easy is it to store extra quantities of this category?" (1 = "not at all easy," and 7 = "extremely easy"). • Impulse buying: "How often do you buy this category on a whim when you pass by it in the store?" (1 = "not at all often," and 7 = "extremely often"). • Perceived healthfulness (health halo): Overall perceived healthfulness of a category (1 = "not at all healthy," and 7 = "extremely healthy").

[a]The first three months of purchases are used as the initialization period.

Notes: All survey measures except the category characteristics were designed and administered by SymphonyIRI.

Model 2 does not rely on ignorability of treatment and instead uses an instrument to control for unobservables. The instrument must be relevant; that is, it must be able to predict the endogenous Club Shopper variable strongly enough. In addition, it should satisfy the exclusion restriction; that is, it should not affect Y directly, only indirectly via Club Shopper (see, e.g., Germann, Ebbes, and Grewal 2015; Rossi 2014). We believe that storage space at home can serve as a valid instrument. It is correlated with the Club Shopper variable: because one has to buy in bulk at a club store, space constraints make shopping in a club store less likely (Bhatnagar and Ratchford 2004). However, we cannot think of any reason why storage space should directly affect packaged food purchases after we account for its effect on club store shopping. Of course, storage space may be correlated with other household characteristics, such as income and household size, that also drive packaged food purchases. However, those variables, along with a rich set of other controls, are included in the model (and therefore are not in the error term of Equation 1). There is no reason for storage space to be correlated with variables, such as intrinsic preference for salty or sweet or fatty foods, or dietary restrictions due to health or personal reasons, that remain in the error term. Therefore, we argue that storage space satisfies the exclusion constraint.

Although we do not observe storage space, we know whether a household lives in a rented or owned residence (see Table 2). Not only are most rented dwellings apartments, which are much smaller than homes, but rented homes and apartments are each significantly smaller than their owned counterparts.[4] Therefore, the Rent dummy variable is a good measure of storage space constraints and is a theoretically valid instrument. We confirm that it does not affect packaged food purchases after we control for other observed variables, by estimating Model 1 for households without easy access to club stores who do not shop there, and including the Rent variable in the model. We find that the Rent variable is not significant for any of our purchase outcome variables, and therefore we have theoretical and empirical justification for its exogeneity. We test the strength of the instrumental variable using the incremental F-statistic in a first-stage regression. At 18, the F-statistic meets the hurdle of 10 (e.g., Bound, Jaeger, and Baker 1995) but is not as strong as we would like. Also, the instrumental variable model requires an additional assumption about the nature of self-selection for consistency when the treatment effect is heterogeneous: that households do not select into the treatment based on the idiosyncratic portion of their response to treatment (Wooldridge 2002, p. 626; Verbeek 2012, pp. 265–66). Because of these limitations, we view the instrumental variable model as only one of multiple models to estimate a robust club store effect.

The model is estimated on the same sample as Model 1. We first estimate a logit model of the decision to shop in a club store using the X vector in Equation 1 and the Rent instrument. Then, we use the predicted propensity and its interactions with the mean-centered X variables as instruments for Club Shopper_h and interaction variables (Wooldridge 2002, p. 626).

[4]According to the U.S. Census Bureau, around 70% of renters lived in multifamily apartment buildings in 2007, and the average size of an apartment (1,200 square feet) was less than half the average size of a single-family home (2,500 square feet). In addition, the average size of an apartment for rent was only 1,121 square feet compared with the average apartment for sale, which was 1,577 square feet. Even within single-family homes, the area of a rented home is about 26% smaller than that of an owned home (Harvard University Joint Center for Housing Studies 2011).

Small Sample Difference-in-Difference (D-in-D) Models

Two additional models control for time-invariant unobservables and all unobservables, respectively. They compare before–after changes for households who choose to shop at a club store after a new one opens in the vicinity, with corresponding changes for two control groups. In Model 3, the 182 households who do not shop at the format even after a club store opens in their vicinity serve as a control group pool for the 83 treatment households who do. We match them on the observable household characteristics, using nearest-neighbor propensity score matching without replacement and a caliper of .05. We are able to match 68 of the treatment households, and we estimate Equation 2 on the 136 matched treatment and control households:

$$
\begin{aligned}
(2)\qquad Y_{ht} = {} & \alpha + X_h\beta + \gamma_1\text{Club Shopper}_h \\
& + \gamma_2\text{After}_{ht} + \gamma_3\text{Club Shopper}_h \times \text{After}_{ht} \\
& + \delta_1\text{Club Shopper}_h \times \left(X_h - \bar{X}\right) + \delta_2\text{After}_{ht} \\
& \times \left(X_h - \bar{X}\right) + \delta_3\text{Club Shopper}_h \\
& \times \text{After}_{ht} \times \left(X_h - \bar{X}\right) + \epsilon_{ht}
\end{aligned}
$$

The outcome variables, Y_{ht}, are aggregated to two time periods for each household, one before and the other after the new store opening in their vicinity. After_{ht} is a dummy variable that equals 1 if the time is after a club store opens near household h and 0 otherwise; γ_3 is the D-in-D estimate of the club store effect. Interactions of the mean-centered household characteristics with the D-in-D term permit heterogeneity in the treatment effect. We also include interactions with Club Shopper_h and After_{ht} for completeness.

This model controls for observables, as in Model 1; in addition, the before–after difference controls for time-invariant unobservables. Still, it is possible that the remaining unobservable reasons that control households decided not to shop at the club store may not satisfy ignorability of treatment. Model 4, described later, addresses that concern.

Club stores open at different times in different locations. Therefore, for some months, households in areas where a club store will open later offer a potential control group for treatment households located near the stores that opened earlier. For example, a club store may open in location A in July 2007 and in location B in June 2008. The period from January 2006 to May 2008 for households in location B who have not shopped at club stores until June 2008 are a potential control group pool for treatment households in Location A. Of the 83 original treatment households, 70 are in locations where a club store opened early enough to correspond to control households with a later store opening, and we have 892 potential control households. Of these 892 households, 238 shop at a club store once it opens in their vicinity, so they clearly do not systematically differ from the treatment group on unobservables. Therefore, for the period before their own store opening, these 238 households serve as a control group for another D-in-D analysis, which is also estimated using Equation 2.

Moderation of Treatment Effect by Household and Category Characteristics

All our models include household characteristics and their potential interactions with the treatment term.[5] Considering the many potential interactions, we follow an empirical block-by-block selection process (Bijmolt et al. 2005; Van Heerde et al. 2013). Specifically, we test interactions with variables relating to one driver (e.g., household needs) at a time, retaining in the final model those variables that are significant from each block for at least one of the outcome variables. To examine the impact of category characteristics, we do a median split of categories into "high" and "low" groups on storability, impulse buying, and perceived healthfulness, and estimate models separately for the high and low groups.

RESULTS

The Average Club Store Effect

Treatment effects from main sample Models 1 and 2. The first column of data in Table 3 shows the ATE for all our outcome variables, estimated from Model 1.[6] After controlling for all the observable characteristics, we find that club store shopping is associated with significantly higher quantity (more servings per capita) for the average household; it has a small but significant association with quality (fewer calories, sugar, and fat per serving), but the higher quantity effect far exceeds the improved quality effect, so that calories, sugar, and saturated fat per capita increase significantly. Further, club store shopping is not associated with a reduction in either total spending or total shopping trips; both increase by significant, though small, amounts.

The second column of data of Table 3 shows the treatment effect from the instrumental variable model. The magnitudes of the treatment effects for all the per-capita variables (servings, calories, sugar, and saturated fat) are similar to those in Model 1, even though the standard errors are much higher, making the effects on sugar and servings per capita not significant. The treatment effects for the per-serving quality variables are not robust.

Treatment effects from D-in-D models 3 and 4. The rightmost two columns of Table 3 show the D-in-D estimates from Models 3 and 4. The treatment effects for all the per-capita variables (servings, calories, sugar, and saturated fat) are robust in the D-in-D models. Among the eight estimates (four outcomes across two models), all are directionally consistent, and most have magnitudes comparable with Model 1. They are also statistically significant, with two exceptions: the servings and sugar per capita effects fall short of significance in Model 3. We note that the magnitude of the sugar per capita effect is much higher in the D-in-D models than in Model 1, but it is also much less precise. Still, this robustness is reassuring, given the different samples and methods and the smaller sample size of the D-in-D estimates.[7]

The ATEs for quality (i.e., calories, sugar, and saturated fat per serving) are not significant in the D-in-D models. Across the four models, therefore, we conclude that there is little to no effect of club store shopping on quality per serving and focus the remainder of our discussion on the per-capita outcomes. We also focus the remainder of our discussion on the estimates from Models 1 and 2 that are based on the large sample.

Moderation of Club Store Effect by Observable Household Characteristics

Table 4 contains estimates of the interactions between household characteristics and the treatment variable, selected using the empirical process described previously. The table shows that the pattern of interactions is very similar for the OLS (Model 1) and instrumental variable (Model 2) estimates. Consistent with expectations regarding household needs, the club store effect on calories, sugar, and fat per capita is more positive for large families and heavy buyers of general household merchandise. It is less positive for heavy users of personal care products, which makes sense to the extent that use of personal care products is indicative of households who are more concerned with their physical well-being and appearance. Finally, the effect on fat (though not calories and sugar) is more positive for those who travel farther to the store and for high-BMI households. This tendency of overweight households to be less careful about fat than sugar is consistent with Ma, Ailawadi, and Grewal (2013). Other interactions are not significant for the per-capita outcomes, although they meet our selection criteria.

The Club Store Effect Across Category Types

Table 5 summarizes treatment effects for the high versus low groups on each category characteristic, again for Models 1 and 2. The results are largely robust across the two models and consistent with our expectations in Table 1. First, most of the increase in the per-capita outcomes occurs through storable foods. In Model 1, the servings, calories, and sugar treatment effects are several times higher for the high-storability group than for the low-storability group. The one exception is fat, for which the increase comes equally from both groups. The effects for high- versus low-impulse foods are also in line with our expectations. They are positive for both groups but significantly larger for the high-impulse group. Finally, the difference in club store effects for foods that are perceived as more or less healthy is more nuanced and appears consistent with a health halo bias. Club store shoppers increase servings more from foods that are perceived to be healthy; they increase sugar and fat less from the healthful group; and the increase calories per capita equally from both groups. In Model 2, the pattern is similar, except that, just as the overall treatment effect is not significant for servings and sugar per capita, nor are the corresponding effects for the high- and low-healthfulness groups.

Additional Test

Effect on need-based products. To further demonstrate that our results do not simply arise from households with large unobservable requirements for all types of products, including food, being more likely to shop at club stores, we estimate the

[5]The correlations between all the variables in our analysis are provided in Web Appendix B. Neither the correlations nor the variance inflation factors raise any cause for multicollinearity concerns.

[6]We also compute the ATT and find very similar patterns, except that the ATT is generally slightly smaller than the ATE. To conserve space, we only report the ATE. The two would be identical if the treatment effect were homogeneous. The small difference in our analysis is attributable to the fact that only a few interactions with the treatment variable are significant, as we discuss next.

[7]A comparison of the main sample and the small sample available for D-in-D analysis, as well as results for household and category moderation in the D-in-D models, are available in Web Appendix C.

Table 3

AVERAGE CLUB STORE TREATMENT EFFECTS

Outcome Variable	*Model 1 (n = 6,657)*		*Model 2 (n = 6,657)*		*Model 3 (n = 272)*		*Model 4 (n = 616)*	
Quantity								
Servings per capita	82.21***	(8.19)	87.93	(74.49)	85.70	(58.82)	103.26**	(47.77)
Quality								
Calories per serving	−3.58***	(.52)	3.96	(4.75)	.80	(4.36)	−.39	(3.08)
Saturated fat per serving (g)	−.02*	(.01)	.20**	(.09)	.04	(.08)	.01	(.05)
Sugar per serving (g)	−.38***	(.05)	−.44	(.47)	.06	(.33)	−.06	(.30)
Quantity × Quality								
Calories per capita	3,410.5***	(403.3)	7,824.8**	(3,721.9)	5,804.3**	(2,503.0)	4,950.2**	(2123.9)
Saturated fat per capita (g)	68.3***	(6.8)	183.5***	(63.9)	99.5**	(39.2)	67.0 **	(33.8)
Sugar per capita (g)	159.3***	(38.1)	69.9	(344.9)	378.5	(238.1)	418.2**	(183.5)
Shopping Behavior								
Total food spending per capita ($)	3.72***	(.96)	12.50	(8.90)	10.84**	(4.84)	8.35**	(3.91)
Number of shopping trips	2.41***	(.16)	3.00*	(1.60)	2.62***	(.90)	1.39**	(.63)

*$p < .10$.
**$p < .05$
***$p < .01$.
Notes: Standard errors are reported in parentheses. All p-values are two-sided. Model 1: Regression on club shopping dummy, X variables, and interactions. Model 2: Instrumental variable model. Model 3: D-in-D, club shoppers versus nonshoppers near new store openings. Model 4: D-in-D, club shoppers near earlier openings versus eventual shoppers near later openings.

treatment effect for health remedies. The logic is as follows: The use of health remedies is need-based. It is unlikely that households use more of these products simply because they are purchased at lower prices, in larger packages, and stockpiled at home. Therefore, club store shopping should not increase purchases of such products. If we saw a significant treatment effect for such categories, it would put in question our results, suggesting that we did not adequately control for unobservable household requirements.

We therefore identify all the health remedy categories in the IRI database (adult incontinence remedies, analgesics, cold and allergy relief, gastrointestinal remedies, family planning, feminine hygiene products, first aid products, and health and sleeping remedies) and create a new outcome variable: a household's purchase volume per capita of these categories.[8] Of course, we ensure that expenditures on these categories are not included in the observable variables on which the groups are matched. As expected, we find an insignificant treatment effect on this outcome variable in all our models. This further increases confidence in the validity of our estimates for packaged food purchases.

Stable results in later time period. Models 1 and 2 are based on a sample of households with valid data for the 2006–2007 period. We estimate those models for a sample with valid data during the 2008–2009 period to assess the stability of our results, although we realize that the financial crisis during this period may have made households spend less overall. The estimated treatment effects are very similar.[9]

DISCUSSION

The club store format's low prices, large package sizes, membership fees, and low store density together result in a significant, substantial, and robust increase in the total quantity of packaged food purchased by households who shop at this format, compared with what their purchases would have been had they not shopped at this format. Because there is little to no association of club store shopping with the quality of packaged food purchased, the quantity increase translates into a significant, substantial, and robust increase in calories, sugar, and saturated fat purchased by these households.

Although the magnitudes of the treatment effects vary across models, the average is more than 3,000 calories and more than 60 grams of saturated fat per person per month. This is substantial, seen, for example, in light of the commonly cited rule of thumb that 100 additional calories per day translate into a weight gain of 10 pounds in a year.[10] Essentially, households buy more quantity in the club store format and do not offset that with proportionally fewer purchases in other formats. This is also reflected in our finding that club store shopping is not associated with a reduction in total shopping trips.

Most of the effect on calories, sugar, and saturated fat comes from storable foods. Households increase purchases of storable products, but not perishable ones, when they shop at club stores, presumably because they intend to store them for consumption at a normal rate, but then they appear to use them faster. The club store effect is also derived more from high-impulse foods. This result also makes sense considering our finding that club store shopping does not substitute for trips to other formats. Consequently, opportunities to make impulse purchases are not reduced, nor is total spending. With regards to heterogeneity, our main results show that households with larger families are especially prone to the club store effect. Income and health consciousness do not significantly limit the club store effect. However, to the extent that expenditure on

[8]Because units of volume are not comparable across categories, we multiply each category's volume by its average price per unit volume and then sum across categories (for a similar calculation, see Ma et al. 2011).

[9]Results are available in Web Appendix E.

[10]Of course, weight gain, just like weight loss, is self-limiting (Katan and Ludwig 2010).

Table 4

INTERACTION ESTIMATES IN OLS AND INSTRUMENTAL VARIABLE MODELS

	OLS Regression (Model 1)				Instrumental Variable Regression (Model 2)			
Independent Variable	Servings per Capita (R^2 = .20)	Calories per Capita (R^2 = .22)	Saturated Fat per Capita (g) (R^2 = .22)	Sugar per Capita (g) (R^2 = .13)	Servings per Capita	Calories per Capita	Saturated Fat per Capita (g)	Sugar per Capita (g)
Club shopper	82.21*** (8.19)	3,410.49*** (403.29)	68.28***(6.83)	159.27*** (38.06)	87.93 (74.49)	7,824.75** (3,721.87)	183.53*** (63.90)	69.88 (344.87)
Interactions: Club shopper × . . .								
Family size	.71 (6.31)	913.72*** (310.70)	10.47** (5.26)	73.93** (29.32)	13.81 (14.99)	2,306.77*** (749.14)	30.98** (12.86)	140.11** (69.42)
Personal care spending per capita	−.93 (.66)	−65.57** (32.52)	−1.23** (.55)	−6.65** (3.07)	−2.65* (1.48)	−230.50*** (73.92)	−2.69** (1.27)	−21.40*** (6.85)
General merchandise spending per capita	2.14*** (.43)	132.92*** (21.12)	1.65*** (.36)	10.52*** (1.99)	7.11*** (.93)	446.83*** (46.47)	7.04*** (.80)	29.03*** (4.31)
Fast-food frequency	20.21 (16.07)	538.90 (791.33)	−.98 (13.40)	84.67 (74.67)	48.80 (41.49)	2,631.38 (2,073.24)	68.69* (35.60)	202.41 (192.11)
Income					−9.98 (6.20)	−569.28* (309.98)	−8.22 (5.32)	−7.43 (28.72)
Club store distance	1.26 (.88)	41.07 (43.34)	1.44* (.73)	4.46 (4.09)	2.23 (2.78)	233.75* (138.89)	5.02** (2.38)	−2.02 (12.87)
Nutrition interest					−23.45 (28.21)	−1,163.72 (1,409.56)	2.58 (24.20)	−51.96 (130.61)
Education	−7.28 (5.38)	−375.72 (264.89)	−4.77 (4.48)	−3.47 (25.00)				
Healthy behaviors	−14.17 (18.44)	−602.74 (908.11)	−10.11 (15.37)	−21.54 (85.69)	−8.89 (42.36)	−529.32 (2,116.47)	−7.06 (36.34)	−32.45 (196.11)
BMI	.88 (1.41)	92.52 (69.60)	2.25* (1.18)	2.83 (6.57)	−.74 (3.30)	213.16 (165.03)	8.12*** (2.83)	4.25 (15.29)

*$p < .10$.
**$p < .05$
***$p < .01$.
Notes: Standard errors are reported in parentheses. All p-values are two-sided. For complete model estimates including the main effects and control variables, see Web Appendix D.

Table 5
AVERAGE TREATMENT EFFECT ACROSS CATEGORY TYPES

	Regression				*Instrumental Variable*			
	Servings per Capita	*Calories per Capita*	*Saturated Fat per Capita (g)*	*Sugar per Capita (g)*	*Servings per Capita*	*Calories per Capita*	*Saturated Fat per Capita (g)*	*Sugar per Capita (g)*
High storability	75.51*** (7.34)	2,964.32*** (268.55)	35.03*** (3.31)	164.42*** (32.23)	53.12 (66.66)	4,878.55** (2456.87)	75.10** (30.71)	–40.34 (291.83)
Low storability	6.70*** (1.78)	446.17** (191.53)	33.25*** (4.82)	–5.15 (12.68)	34.81** (16.47)	2,946.19* (1767.62)	108.43** (44.54)	110.21 (115.89)
High impulse nature	52.78*** (4.97)	2,137.84*** (290.57)	50.89*** (5.36)	46.46* (24.74)	91.97** (45.49)	6,093.82** (2695.17)	148.62*** (50.45)	–80.41 (225.10)
Low impulse nature	29.43*** (5.17)	1,272.66*** (169.64)	17.39*** (2.37)	112.81*** (23.47)	–4.04 (46.84)	1,730.93 (1538.11)	34.91 (21.48)	150.28 (211.42)
Healthy	48.07*** (4.42)	1,644.96*** (217.39)	24.12*** (2.95)	47.16** (20.26)	64.21 (40.13)	3,778.19* (1992.18)	47.39* (27.10)	87.87 (183.65)
Unhealthy	34.14*** (5.63)	1,765.54*** (259.65)	44.16*** (5.01)	112.11*** (28.80)	23.71 (51.00)	4,046.56* (2384.38)	136.14*** (46.86)	–17.99 (261.07)

*$p < .10$.
**$p < .05$
***$p < .01$.
Notes: Standard errors are reported in parentheses. All p-values are two-sided.

personal care products is a surrogate for concern about physical appearance and wellness, households with such concerns are less prone to the club store effect.

Implications for Marketers

Marketers must balance their goal of profitably making food affordable and convenient for consumers against public health concerns. Grocery retailers are increasing their proportions of healthy products and improving their in-store merchandising (Lee 2016). Major players like Walmart are urging suppliers to reformulate products to make them healthier (Walmart 2011). Suppliers, for their part, are trying to balance revenues and profits with the imperative to improve healthfulness (*The Economist* 2012; Slining, Ng, and Popkin 2013). As club stores and their suppliers find ways to address the significant impact that the characteristics of this format have on consumers' packaged food intake, there may be some unique win–win opportunities. The opportunities we suggest next, though motivated by our research, are speculative because we have not tested them.

First, large package sizes are a hallmark of this format that is unlikely to change, yet the relative largeness of a package need not be the same across all products. Healthier product packages could be made even larger, and the less healthy product packages could be shrunk a little. Of course, given consumers' vulnerability to the health halo bias, it is important to highlight appropriate serving sizes.

Second, large households may have to buy more variety to satisfy the heterogeneous tastes of multiple family members, and shopping at club store requires them to buy large quantities of each variety, which then get consumed faster. The problem could be mitigated if products at club stores were packaged as bundles of different items instead of only large quantities of one item. Large families could then satisfy some of their needs with these bundles instead of having to buy large quantities of each variety. Decisions about which items to bundle could be made on the basis of basket analysis and household purchase histories, which are easily available to club store retailers from their member databases.

Third, club stores sell fresh produce, also in large packages, which can dissuade many shoppers from buying it for fear of spoilage. Bundling regular-sized packages of different produce items might be more desirable and feasible because it would encourage more shoppers to buy the produce packages. Bundling private-label or unbranded fresh products would be easy for the retailer, whereas bundling packaged products from different manufacturers might be more difficult. There might even be a benefit to club store retailers if their own spoilage costs decreased because of produce bundling.

Finally, we do not find a robust effect of club store shopping on calories, sugar, and saturated fat per serving, which is in line with the fact that packaged foods in club stores currently do not differ on these dimensions from packaged foods in supermarkets. However, because of the large package sizes, improvements in nutritional quality can have a bigger impact on consumer health in this format than in others. A package of a healthier product bought in a club store would account for much more of a family's intake than a package of the same product bought in other grocery formats. So reductions in unhealthy ingredients could go further in club stores. This is likely to appeal to the format's clientele, who are more educated and health conscious. The improvements could even be accompanied by a small price increase because club store shoppers have higher incomes, and the format's current price advantage is substantial.

Implications for Researchers

Our work suggests some important directions for researchers. First, we present five behavioral mechanisms by which club store shopping may influence packaged food purchases and consumption. Our overall results and the differences across category types support some role for each mechanism, but we have not teased apart their individual impacts. Now that our work has documented overall treatment effects, researchers can shed more light on the underlying process with a combination of lab experiments and field data.

Second, our many robustness tests engender confidence in the validity of the effects we document. Nevertheless, if researchers have access to data on a large sample of households before and after a warehouse club store opening, a larger-scale D-in-D analysis than the one we were able to conduct would be useful. It could more precisely identify the effects, especially for sugar per capita and the quality (per serving) variables. Such an analysis could also identify the specific types of products whose purchases are shifted to club stores from other formats and therefore provide additional insight into how the impact of club store shopping on food purchases occurs. Such a D-in-D analysis would be particularly valuable given the limitations of instrumental variable analysis for estimating heterogeneous treatment effects.

Third, if data were available for a longer period after households first start shopping in the club store format, researchers could study the impact of learning. Consumers may learn not to buy extra packaged food because they waste it or eat too much, or they may get used to buying and eating more.

Fourth, we observe purchases, not consumption, of packaged foods. Our analysis approach aggregates data over time, thus accounting for the likelihood that households who buy extra in one week or store may reduce purchases in subsequent weeks and other stores. However, it is possible that some of the food bought in bulk may spoil and be discarded. We cannot rule this out, but it is important to note our finding that the club store effect occurs mainly in storable foods, not perishable ones, so spoilage should be less likely to account for it. It is also possible that the increase we find in purchases of packaged food substitutes for away-from-home food in restaurants. We have good controls for away-from-home intake in our fast-food frequency, working status, and income variables. Fast-food frequency is a direct measure, and working status and income are strong correlates of away-from-home eating because high-income people and those who work outside the house, especially in dual-employment households, are more likely to eat out. However, we cannot study substitution with our data, so it would be helpful for future research to track away-from-home food.

Finally, we study the club store format here, but new formats are being introduced by retailers all over the world, as are new initiatives like home delivery, click-and-collect, and self-scanning in stores. All of these have the potential to affect food-for-home purchases and offer interesting avenues for research. We hope that our work will spur more such research.

APPENDIX: CHARACTERISTICS OF MAJOR GROCERY FORMATS

	Average Value Relative to Supermarket Format		
Variable	*Club Stores*	*Mass Merchants*	*Supercenters*
Marketing Mix			
Price per serving	77.5%[a]	83.5%	81.6%
SKUs per category	15.5%	55.5%	110.9%
Number of servings per SKU	312.7%	102.1%	105.4%
Percentage of SKUs on promotion	70.4%	53.6%	76.3%
Distance			
Distance to nearest store	1192.7%	629.8%	787.9%
Nutritional Quality			
Calories per serving	99.8%	99.9%	99.7%
Sugar per serving	98.0%	99.6%	99.2%
Saturated fat per serving	99.9%	101.1%	100.7%

[a]Read as: "The weighted average price per serving in the club store format is 77.5% of the corresponding price in the supermarket format."

Notes: Averages of each marketing-mix and nutritional-quality variable are first computed for each packaged food category in each format. For marketing mix, we divide the category averages for each of the three listed formats by the corresponding averages for supermarkets to get indices relative to supermarkets; we report the weighted average index across categories (weighted by category sales). For nutritional quality, we first take the weighted average across categories for each format and then compute the indices relative to supermarkets.

REFERENCES

Ailawadi, Kusum, and Scott Neslin (1998), "The Effect of Promotion on Consumption: Buying More and Using It Faster," *Journal of Marketing Research*, 35 (3), 390–98.

Ailawadi, Kusum, Koen Pauwels, and Jan-Benedict Steenkamp (2008), "Private Label Use and Store Loyalty," *Journal of Marketing*, 72 (November), 19–30.

American Heart Association (2016), "Dietary Recommendations for Healthy Children," http://www.heart.org/HEARTORG/HealthyLiving/Dietary-Recommendations-for-Healthy-Children_UCM_303886_Article.jsp.

Andrews, J. Craig, Richard G. Netemeyer, and Scot Burton (2009), "The Nutrition Elite: Do Only the Highest Levels of Caloric Knowledge, Obesity Knowledge, and Motivation Matter in Processing Nutrition Ad Claims and Disclosures?" *Journal of Public Policy & Marketing*, 28 (1), 41–55.

Arkes, Hal R., and Catherine Blumer (1985), "The Psychology of Sunk Cost," *Organizational Behavior and Human Decision Processes*, 35 (February), 124–40.

Balasubramanian, Siva K., and Catherine Cole (2002), "Consumers' Search and Use of Nutrition Information: The Challenge and Promise of the Nutrition Labeling and Education Act," *Journal of Marketing*, 66 (July), 112–27.

Baumeister, Roy F. (2002), "Yielding to Temptation: Self-Control Failure, Impulsive Purchasing, and Consumer Behavior," *Journal of Consumer Research*, 28 (4), 670–76.

Bell, David R., Jeongwen Chiang, and V. Padmanabhan (1999), "The Decomposition of Promotional Response: An Empirical Generalization," *Marketing Science*, 18 (4), 504–26.

Bell, David R., and Christian A.L. Hilber (2006), "An Empirical Test of the Theory of Sales: Do Household Storage Constraints Affect Consumer and Store Behavior?" *Quantitative Marketing and Economics*, 4 (2), 87–117.

Bhatnagar, Amit, and Brian T. Ratchford (2004), "A Model of Retail Format Competition for Non-Durable Goods," *International Journal of Research in Marketing*, 21 (1), 39–59.

Bijmolt, Tammo, Harald van Heerde, and Rik Pieters (2005), "New Empirical Generalizations on the Determinants of Price Elasticity," *Journal of Marketing Research*, 42 (May), 141–56.

Bodiba, P., S.N. Madu, J.O. Ezeokana, and O.A.U. Nnedum (2008), "The Relationship Between Body Mass Index and Self-Concept Among Adolescent Black Female University Students," *Curationis*, 31 (1), 77–84, http://www.curationis.org.za/index.php/curationis/article/view/917/854.

Bound, John, David A. Jaeger, and Regina M. Baker (1995), "Problems with Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogenous Explanatory Variable Is Weak," *Journal of the American Statistical Association*, 90 (430), 443–50.

Celsi, Richard L., and Jerry C. Olson (1988), "The Role of Involvement in Attention and Comprehension Processes," *Journal of Consumer Research*, 15 (2), 210–24.

Chaiken, Shelly (1980), "Heuristic Versus Systematic Information Processing and the Use of Source Versus Message Cues in Persuasion," *Journal of Personality and Social Psychology*, 39 (5), 752–66.

Chandon, Pierre, and Brian Wansink (2002), "When Are Stockpiled Products Consumed Faster? A Convenience-Salience Framework of Post-Purchase Consumption Incidence and Quantity," *Journal of Marketing Research*, 39 (8), 321–35.

Chandon, Pierre, and Brian Wansink (2011), "Is Food Marketing Making Us Fat? A Multi-Disciplinary Review," *Foundations and Trends in Marketing*, 5 (3), 113–96.

Chernev, Alexander (2011), "The Dieter's Paradox," *Journal of Consumer Psychology*, 21 (2), 178–83.

Code of Federal Regulations (2011), "Reference Amounts of Foods Customarily Consumed Per Eating Occasion," 21 C.F.R. §101.12.

Cohen, Deborah, and Susan Babey (2012), "Contextual Influences on Eating Behaviors: Heuristic Processing and Dietary Choices," *Obesity Reviews*, 13 (9), 766–79.

Courtemanche, Charles, and Art Carden (2011), "Supersizing Supercenters? The Impact of Walmart Supercenters on Body Mass Index and Obesity," *Journal of Urban Economics*, 69 (March), 165–81.

The Economist (2012), "Food for Thought: Food Companies Play and Ambivalent Part in the Fight Against Flab," (December 15), https://www.economist.com/news/special-report/21568064-food-companies-play-ambivalent-part-fight-against-flab-food-thought.

Ehrich, Kristine R., and Julie R. Irwin (2005), "Willful Ignorance in the Request for Product Attribute Information," *Journal of Marketing Research*, 42 (August), 266–77.

Fox, Edward J., Alan L. Montgomery, and Leonard M. Lodish (2004), "Consumer Shopping and Spending Across Retail Formats," *Journal of Business*, 77 (April), S25–60.

Germann, Frank, Peter Ebbes, and Rajdeep Grewal (2015), "The Chief Marketing Officer Matters," *Journal of Marketing*, 79 (May), 1–22.

Geyskens, Inge, Katrijn Gielens, and Marnik Dekimpe (2002), "The Market Valuation of Internet Channel Additions," *Journal of Marketing*, 66 (2), 102–19.

Harvard University Joint Center for Housing Studies (2011), "Renter Demographics," http://www.jchs.harvard.edu/sites/jchs.harvard.edu/files/ahr2011-3-demographics.pdf.

Hofmann, Wilhelm, Guido M. van Koningsbruggen, Wolfgang Stroebe, Suresh Ramanathan, and Henk Aarts (2010), "As Pleasure Unfolds: Hedonic Responses to Tempting Food," *Psychological Science*, 21 (2), 1863–70.

Hortascu, Ali, and Chad Syverson (2015), "The Ongoing Evolution of US Retail: A Format Tug-of-War," *Journal of Economic Perspectives*, 29 (4), 89–111.

Inman, J. Jeffrey, Venkatesh Shankar, and Rosellina Ferraro (2004), "The Roles of Channel-Category Associations and Geodemographics in Channel Patronage," *Journal of Marketing*, 68 (2), 51–71.

Inman, J. Jeffrey, Russell S. Winer, and Rosellina Ferraro (2009), "In-Store Decision Making: The Role of Category-Level and Shopping Trip-Level Factors," *Journal of Marketing*, 73 (September), 19–29.

Kahn, Barbara, and Brian Wansink (2004), "The Influence of Assortment Structure on Perceived Variety and Consumption," *Journal of Consumer Research*, 30 (4), 519–33.

Katan, Martijn, and David Ludwig (2010), "Extra Calories Cause Weight Gain—But How Much?" *Journal of the American Medical Association*, 303 (1), 65–66.

Khan, Uzma, and Ravi Dhar (2006), "Licensing Effect in Consumer Choice," *Journal of Marketing Research*, 43 (May), 259–66.

Lee, Bruce Y. (2016), "Grocery Stores Taking More Active Roles in Health Promotion," *Forbes* (March 28), http://www.forbes.com/sites/brucelee/2016/03/28/grocery-stores-taking-more-active-roles-in-health-promotion/#6726c9c94d05.

Ma, Yu, Kusum Ailawadi, Dinesh Gauri, and Dhruv Grewal (2011), "An Empirical Analysis of the Effect of Gasoline Prices on Consumer Shopping Behavior," *Journal of Marketing*, 75 (2), 18–35.

Ma, Yu, Kusum Ailawadi, and Dhruv Grewal (2013), "Soda Versus Cereal and Sugar Versus Fat: Drivers of Healthful Food Intake and the Impact of Diabetes Diagnosis," *Journal of Marketing*, 77 (3), 101–20.

Moorman, Christine, and Erika Matulich (1993), "A Model of Consumers' Preventive Health Behaviors: The Role of Health Motivation and Health Ability," *Journal of Consumer Research*, 20 (September), 208–28.

Mukhopadhyay, Anirban, and Gita V. Johar (2009), "Indulgence as Self-Reward for Prior Shopping Restraint: A Justification-Based Mechanism," *Journal of Consumer Psychology*, 19 (July), 334–45.

Nikolova, Hristina D., and J. Jeffrey Inman (2015), "Healthy Choice: The Effect of Simplified Point-of-Sale Nutritional Information on Consumer Choice Behavior," *Journal of Marketing Research*, 52 (6), 817–35.

Orhun, A. Yesim, and Mike Palazzolo (2017), "Frugality Is Hard to Afford," Chicago Booth Kilts Center for Marketing, Nielsen Data Set Paper 2-031; Ross School of Business Paper 1309, https://ssrn.com/abstract=2742431.

Petty, Richard E., John T. Cacioppo, and David Schumann (1983), "Central and Peripheral Routes to Advertising Effectiveness: The Moderating Role of Involvement," *Journal of Consumer Research*, 10 (2), 135–46.

Rossi, Peter (2014), "Even the Rich Can Make Themselves Poor: A Critical Examination of IV Methods in Marketing Applications," *Marketing Science*, 33 (5), 655–72.

Slining, Meghan, Shu Wen Ng, and Barry Popkin (2013), "Food Companies' Calorie-Reduction Pledges to Improve U.S. Diet," *American Journal of Preventive Medicine*, 44 (2), 174–84.

Stilley, Karen, J. Jeffrey Inman, and Kirk Wakefield (2010), "Spending on the Fly: Mental Budgets, Promotions, and Spending Behavior," *Journal of Marketing*, 74 (May), 34–47.

Talukdar, Debabrata (2008), "Cost of Being Poor: Retail Price and Consumer Price Search Differences Across Inner-City and Suburban Neighborhoods," *Journal of Consumer Research*, 35 (3), 457–71.

USDA (2014), "Average Daily Intake of Food by Food Source and Demographic Characteristics," USDA Economic Research Service, https://www.ers.usda.gov/data-products/food-consumption-and-nutrient-intakes/food-consumption-and-nutrient-intakes/#Nutrient Intake Estimates [accessed July 24, 2017].

Van Heerde, Harald J., Maarten J. Gijsenberg, Marnik G. Dekimpe, and Jan-Benedict E.M. Steenkamp (2013), "Price and Advertising Effectiveness Over the Business Cycle," *Journal of Marketing Research*, 50 (April), 177–93.

Verbeek, Marno (2012), *A Guide to Modern Econometrics*, 4th ed. West Sussex, UK: John Wiley & Sons.

Volpe, Richard, Abigail Okrent, and Ephraim Leibtag (2013), "The Effect of Supercenter-Format Stores on the Healthfulness of Consumers' Grocery Purchases," *American Journal of Agricultural Economics*, 95 (3), 568–89.

Walmart (2011), "Walmart Launches Major Initiative to Make Food Healthier and Healthier Food More Affordable," press release (January 20), http://corporate.walmart.com/_news_/news-archive/2011/01/20/walmart-launches-major-initiative-to-make-food-healthier-healthier-food-more-affordable.

Wansink, Brian (1996), "Can Package Size Accelerate Usage Volume?" *Journal of Marketing*, 60 (3), 1–14.

Wansink, Brian (2004), "Environmental Factors That Increase the Food Intake and Consumption Volume of Unknowing Consumers," *Annual Review of Nutrition*, 24 (July), 455–79.

Wansink, Brian, and Pierre Chandon (2006), "Can 'Low-Fat' Nutrition Labels Lead to Obesity?" *Journal of Marketing Research*, 43 (November), 605–17.

Wansink, Brian, and Rohit Deshpandé (1994), "Out of Sight, Out of Mind: The Impact of Household Stockpiling on Usage Rates," *Marketing Letters*, 5 (1), 91–100.

Warehouse Club Intelligence Center (2017), "2017 Warehouse Club Industry Guide," http://www.clubintelligencecenter.com/hhc/documents/wcf_470_20170127.pdf.

Wertenbroch, Klaus (1998), "Consumption Self-Control by Rationing Purchase Quantities of Virtue and Vice," *Marketing Science*, 17 (4), 317–37.

Wooldridge, Jeffrey (2002), *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.