Table of Contents

Appendix Table 1. Food and Beverage Groupings with Examples

General equation for fixed effect model of changes in nutritional profile of packaged food purchases at Walmart or other food retail chain

1

Creation of Inverse Probability Weights

Appendix Table 2. Descriptive statistics on inverse probability weights

Appendix Table 3. Year coefficients using different inverse probability weight approaches

Appendix Table 4. Household characteristics of the Nielsen Homescan sample in 2000 and 2013a (n=2,611,125)

References

Appendix Table 1. Food and Beverage Groupings From Nielsen Modules, With Examples^a

Food Group	Example of Nielsen modules
SSBs	Energy drinks, sports drinks, soda, fruit drinks
Fresh plain milk	Refrigerated milk
Coffee (grounds, beans, RTD)	Ground and whole bean coffee
Diet beverage	Carbonated soft drinks, <20 cal/100g
Tea (bags, loose, RTD)	Herbal tea, tea bags
Water and ice	Bottled water, bagged ice
Candy	Hard candy, chocolate, marshmallows
Savory snacks	Chips, pretzels, frozen hors d'oevres
GBD	Frozen cookie dough, ready-to-eat cookies,
RTE breads	Fresh bagels, fresh bread, frozen bread
RTE Breakfast	Ready-to-eat cereal, granola
Vegetables, fresh or frozen	Fresh carrots, frozen broccoli
Dairy-based dessert	Ice cream, sherbet, frozen novelties
Shelf-stable soups & stews	Ramen noodles, canned soup, soup mixes
Condiments & sauces	Ketchup, barbecue sauce, cocktail sauce
Alcohol	Wine, liquor, beer, malt beverage
Frozen entrees	Frozen dinners
Processed meat	Canned corned beef, lunch meat, bacon
100% Juice	Drinks containing 100% juice
Fats and oils	Butter, olive oil, margarines and spreads
Cheese	American Cheddar cheese, cream cheese, mozzarella
Fruit, fresh & frozen	Refrigerated fruit salad, fresh apples, frozen fruits
Sweeteners	Granulated sugar, molasses
Vegetables, canned or dried	Canned tomatoes, canned artichokes
Yogurt	Refrigerated yogurt
Eggs	Fresh eggs
Pasta & rice	Boxed spaghetti, instant rice, pasta macaroni
Desserts, prepare-at-home	Brownie mix, cake mix, muffin mix
RTE, prepared dishes	Refrigerated entrees, refrigerated chili
Potatoes and corn	Dehydrated mashed potatoes, canned corn
Beef/pork	Frozen ground beef, frozen veal
Shelf-stable creamers, evaporated or condensed milks	Powdered creamer, liquid creamer
Beans and legumes	Garbanzo beans, lima beans, pinto beans
Refrigerated sweetened dairy drinks	Refrigerated shakes, eggnog
Nut and fruit spreads	Peanut butter, honey, jams, jelly
Canned mixed dishes	Canned lasagna, canned chow mein
Nuts and seeds	Nuts, unshelled
Flours	All-purpose flour, white wheat flour
Frozen/refrigerated dairy-based toppings/condiments	Sour cream, whipping cream
Salad dressing	Refrigerated salad dressing, dressing mixes

Fruit, canned	Canned peaches
Frozen pizza	Frozen pizza
Seafood	Frozen breaded shrimp, frozen unbreaded fish
Cereals (requires cooking)	Wheat germ, hominy grits
Other fruit (dried, etc.)	Dried cranberries, raisins
Sweets, misc.	Chocolate syrup
Shelf-stable milks, milk substitutes and milk-based powders	Powdered milk
Shelf-stable Mexican-style products	Tortillas, Mexican shells
Poultry	Frozen poultry, canned turkey
Baby food	Baby cereal
Refrigerated/frozen dough products	Frozen pizza crust
Dry baking mix	Pancake mix, biscuit mix, pizza crust mix
Spices, seasoning, & extracts	Pepper, salt
Baking supplies	Baking powder, baking soda
Spreads and dips	Sandwich spread, garlic spread
^a Food groups were based on Nielson, which groups	

^a Food groups were based on Nielsen, which groups products into 624 modules according to the store section in which they are found (for example, tortilla chips and pretzels represent two modules.) We aggregated modules into 56 groups based on how products are consumed (for example, the tortilla chips and pretzel modules are in a group, "Savory Snacks."

General Fixed Effect Model For Changes in Nutritional Profile of Packaged Food Purchases at Walmart or Other Food Retail Chains

$$\begin{split} Y_{it} = & \beta_0 + \beta_1 Year_y + \beta_2 Unemployment_{mq} + \beta_3 Market_{mq} + \beta_4 Walmkt_{mq} + \beta_5 FPrice_{mq} \\ & + \beta_6 WPrice_{mq} + \beta_7 Household_{hq} + \beta_8 Intx_{hq} + \epsilon_{hq} \end{split}$$

Where $Y_{i,t}$ represents nutrient outcome of retailer PFPS for the i^{th} household in quarter q at a particular retailer (Walmart or other food retail chains). Yeary denotes indicator variables for each year. Unemployment_{am} represents the quarter- and market-specific unemployment rate. $Market_h$ represents market dummies. $Walmkt_{mq}$ is the average number of Walmart stores per 100,000 people in market m in quarter q. $FPrice_{qm}$ represents the average price of packaged foods at other food retail chains and WPrice_{ham} represents the average price of packaged foods in quarter q, market m. Household_{hg} represents a vector of time-varying covariates, including: head of household education (\leq high school degree, some college, \geq college degree), household type (single adult, multiple adults with no kids, adult(s) with kid(s)) and household composition (numbers of women and men aged 19-49 and \geq 50y, and numbers of children aged 0-5y, 6-18y), income (low income $\leq 130\%$ federal poverty level; higher income $\geq 130\%$ federal poverty level), and race/ethnicity (non-Hispanic white, Hispanic, non-Hispanic black, and non-Hispanic other). $Intx_{hq}$ represents the interaction between key household SES characteristics (low income, household composition, race/ethnicity) and time. ε_{hq} represents unobserved time-varying characteristics. Separate models were analyzed for packaged food Walmart and other chain retailers.

Creation of Inverse Probability Weights

To deal with selectivity, or the idea that households who shop at Walmart might be different than households shopping at other food retail chains (FRC) every model includes a time-varying inverse probability weight for being a Walmart customer (i.e., being in the sample) or other FRC customer in a given quarter. The application of these inverse probability weights should create a more balanced sample between those who shop at Walmart and those who shop at other food retail chains. We use logistic regression to predict the probability of being a Walmart or other FRC customer in a given quarter, using market-level and household-level covariates which our previous work has shown to be associated with shopping at food retailers (LST, unpublished observations, 2015).

 $P(Walmart or other FRC customer)_{hq} = household size_h + householdtype_h + race_h + FPL_{hq} + WalmartAFP_{mq} + FoodstoreAFP_{mq +} + Walmartdensity_{my} + Unemploymentrate_{mq}$

Where h is household, q is quarter, and m is market.

FPL=Indicator variables for federal poverty level decile
Household size= indicator variables for number of individuals in household
Household type= indicator variable for household composition: single adult, multiple adult(s),
adult(s) with child(ren)
AFP= weighted average price of foods and beverages purchased

Walmart density= average annual Walmarts per 100,000 individuals

The weights are then stabilized^{1,2} by including in the numerator the overall probability of being a Walmart customer in a given quarter:

P(Walmart or other FRC customer_q)

 $P(Walmart or other FRC customer)_{hq} = household size_h + householdtype_h + race_h + FPL_{hq} + WalmartFPI_{mq} + FoodstoreFPI_{mq +} + Walmartdensity_{my} + Unemploymentrate_{mq}$

These stabilized weights were then truncated at the upper and lower 0.3rd percentile to reduce the influence of large weights. Large weights occur when individuals who had very high likelihoods of being a Walmart or other FRC customer, were not customers; or when households who had very low likelihoods of being a Walmart or other FRC customer, were customers.³ In this sample, these large weights occurred primarily because the Walmart density variable was a very strong predictor of whether someone shopped at Walmart or not, especially for markets with no Walmart. Individuals who shopped at Walmart despite the presence of no Walmart in their market (or vice versa) tended to have more extreme weights. Progressively truncating these weights increases precision by eliminating the influence of large weights.³ We chose to truncate at the 0.3rd and 99.7th percentiles because this allowed us to remove outliers and achieve a more stable weight while still retaining the greatest sample possible. Results appeared robust to the use of various types of inverse probability weights (Appendix Tables 2 and 3).

Walmart		Mean	SD	Minimum	Maximum
	IPW	4.86	13.5	1.02	744
	Stabilized IPW	1.00	0.89	0.24	225
	Stabilized IPW, truncated				
	at 0.3^{rd} and 99.7^{th}	0.99	0.65	0.36	6.11
	percentiles				
Other FRCs					
	IPW	1.48	0.04	1.00	1.48
	Stabilized IPW	1.00	0.20	0.08	141
	Stabilized IPW, truncated				
	at 0.3^{rd} and 99.7^{th}	0.99	0.11	0.28	1.96
	percentiles				

Appendix Table 2. Descriptive Statistics on Inverse Probability Weights

IPW, inverse probability weights; FRC, food retail chain

			Using Differen	it inverse Proba	ionity weight	Approaches		
Energy	density (kcal/	-			TT 7 1			
	Other food retail chains				Walmart			
	No	Stabilized	Truncated IPW ^a	No	Stabilized	Truncated		
	IPW	IPW		IPW	IPW	IPW ^a		
	β (SE)	β (SE)	β (SE)	β (SE)	β (SE)	β (SE)		
2000	(SE) ref	ref	(SE) ref	(SL) Ref	ref	(SL) ref		
2000	101	101	ICI	Kei	101	101		
2001	-0.29	-0.29	-0.29	-8.66***	-9.62***	-9.79***		
	(0.21)	(0.21)	(0.21)	(0.83)	(1.60)	(1.52)		
2002	0.52*	0.51*	0.51*	-10.38***	-14.40***	-13.05***		
	(0.23)	(0.23)	(0.23)	(0.82)	(1.64)	(1.50)		
2003	-1.57***	-1.58***	-1.58***	-17.88***	-20.58***	-20.91***		
	(0.24)	(0.24)	(0.24)	(0.85)	(1.63)	(1.54)		
2004	-2.14***	-2.15***	-2.15***	-25.40***	-27.08***	-27.93***		
	(0.25)	(0.24)	(0.24)	(0.89)	(1.75)	(1.62)		
2005	-1.97***	-1.98***	-1.98***	-30.75***	-34.26***	-34.56***		
	(0.26)	(0.25)	(0.25)	(0.95)	(1.84)	(1.72)		
2006	-2.84***	-2.85***	-2.85***	-35.49***	-39.90***	-39.60***		
	(0.26)	(0.26)	(0.26)	(0.98)	(1.91)	(1.77)		
2007	-2.78***	-2.78***	-2.78***	-43.76***	-48.11***	-48.20***		
	(0.30)	(0.30)	(0.30)	(1.14)	(2.23)	(2.02)		
2008	-4.38***	-4.39***	-4.39***	-51.29***	-56.29***	-56.57***		
	(0.35)	(0.35)	(0.35)	(1.29)	(2.52)	(2.26)		
2009	-5.04***	-5.05***	-5.05***	-51.31***	-57.20***	-57.71***		
	(0.44)	(0.44)	(0.44)	(1.47)	(2.76)	(2.46)		
2010	-4.66***	-4.68***	-4.68***	-53.84***	-59.63***	-60.15***		
	(0.45)	(0.45)	(0.45)	(1.50)	(2.81)	(2.50)		
2011	-7.06***	-7.09***	-7.09***	-55.55***	-61.70***	-62.33***		
	(0.46)	(0.46)	(0.46)	(1.56)	(2.96)	(2.62)		
2012	-8.09***	-8.13***	-8.13***	-60.49***	-66.53***	-67.30***		
	(0.48)	(0.48)	(0.48)	(1.68)	(3.26)	(2.85)		
2013	-10.67***	-10.71***	-10.71***	-66.18***	-72.62***	-73.55***		
	(0.51)	(0.51)	(0.51)	(1.77)	(3.42)	(2.98)		
			Sugar (g/100) g)				
	Other food retail chains				Walmart			
	No	Stabilized	Truncated	No	Stabilized	Truncated		
	IPW	IPW	IPW	IPW	IPW	IPW		
	β	β	β	β	β	β		
2000	(SE)	(SE)	(SE)	(SE)	(SE)	(SE)		
2000	ref	ref	ref	ref	ref	ref		

Appendix Table 3. Year Coefficients Using Different Inverse Probability Weight Approaches

2001	-0.07***	-0.07***	-0.07***	-1.08***	-1.27***	-1.27***
	(0.02)	(0.02)	(0.02)	(0.09)	(0.18)	(0.17)
2002	0.00	0.00	0.00	-1.45***	-1.90***	-1.79***
	(0.02)	(0.02)	(0.02)	(0.08)	(0.18)	(0.17)
2003	-0.27***	-0.27***	-0.27***	-2.31***	-2.71***	-2.79***
	(0.02)	(0.02)	(0.02)	(0.09)	(0.19)	(0.17)
2004	-0.51***	-0.51***	-0.51***	-3.23***	-3.59***	-3.65***
	(0.02)	(0.02)	(0.02)	(0.09)	(0.20)	(0.18)
2005	-0.66***	-0.66***	-0.66***	-3.64***	-4.16***	-4.19***
	(0.02)	(0.02)	(0.02)	(0.10)	(0.20)	(0.19)
2006	-0.77***	-0.77***	-0.77***	-3.95***	-4.55***	-4.54***
	(0.02)	(0.02)	(0.02)	(0.10)	(0.21)	(0.20)
2007	-0.82***	-0.82***	-0.82***	-4.66***	-5.36***	-5.37***
	(0.03)	(0.03)	(0.03)	(0.12)	(0.25)	(0.23)
2008	-0.95***	-0.95***	-0.95***	-5.39***	-6.12***	-6.15***
	(0.03)	(0.03)	(0.03)	(0.13)	(0.28)	(0.25)
2009	-0.90***	-0.90***	-0.90***	-5.29***	-6.09***	-6.14***
	(0.04)	(0.04)	(0.04)	(0.15)	(0.30)	(0.27)
2010	-0.90***	-0.90***	-0.90***	-5.33***	-6.14***	-6.18***
	(0.04)	(0.04)	(0.04)	(0.15)	(0.31)	(0.28)
2011	-1.08***	-1.08***	-1.08***	-5.56***	-6.45***	-6.50***
	(0.04)	(0.04)	(0.04)	(0.16)	(0.33)	(0.29)
2012	-1.28***	-1.28***	-1.28***	-6.11***	-7.01***	-7.08***
	(0.04)	(0.05)	(0.05)	(0.17)	(0.36)	(0.32)
2013	-1.71***	-1.71***	-1.71***	-6.81***	-7.78***	-7.86***
	(0.05)	(0.05)	(0.05)	(0.18)	(0.38)	(0.33)
			Saturated fat (g/	/100g)		
		er food retail ch		N .T	Walmart	T 1
	No IPW	Stabilized IPW	Truncated IPW	No IPW	Stabilized IPW	Truncated IPW
	β	β	β	β	β	β
	(SE)	(SE)	(SE)	(SE)	(SE)	(SE)
2000	ref	ref	ref	ref	ref	ref
2000	101	101	101	101	101	101
2001	0.01	0.01*	0.01*	-0.12***	-0.14***	-0.14***
	(0.01)	(0.00)	(0.00)	(0.02)	(0.04)	(0.03)
2002	0.06***	0.05***	0.05***	-0.13***	-0.19***	-0.16***
	(0.01)	(0.01)	(0.01)	(0.02)	(0.04)	(0.03)
2003	0.05***	0.05***	0.05***	-0.22***	-0.25***	-0.24***
	(0.01)	(0.01)	(0.01)	(0.02)	(0.04)	(0.03)
2004	0.04***	0.04***	0.04***	-0.30***	-0.34***	-0.33***
	(0.01)	(0.01)	(0.01)	(0.02)	(0.04)	(0.04)
2005	0.02**	0.02***	0.02***	-0.45***	-0.49***	-0.49***

Packag	ed Food Purchas	es at Walmart a	Appendix nd Other Food Re 2000 to 201 Smith Tallie e	L3	ges in Nutritiona	I Profile From
	(0.01)	(0.01)	(0.01)	(0.02)	(0.04)	(0.04)
2006	0.02	0.02*	0.02*	-0.52***	-0.59***	-0.57***
2000	(0.01)	(0.01)	(0.01)	(0.02)	(0.04)	(0.04)
2007	-0.02	-0.02*	-0.02*	-0.69***	-0.74***	-0.73***
2007	(0.01)	(0.01)	(0.01)	(0.03)	(0.05)	(0.05)
2008	-0.05***	- 0.05 ***	-0.05***	- 0.88 ***	- 0.95 ***	- 0.93 ***
	(0.01)	(0.01)	(0.01)	(0.03)	(0.06)	(0.05)
2009	0.03*	0.03*	0.03*	-0.82***	-0.91***	-0.90***
	(0.01)	(0.01)	(0.01)	(0.04)	(0.07)	(0.06)
2010	0.09***	0.09***	0.09***	-0.87***	-0.96***	-0.95***
	(0.01)	(0.01)	(0.01)	(0.04)	(0.07)	(0.06)
2011	0.07***	0.07***	0.07***	-0.91***	-1.01***	-1.00***
	(0.01)	(0.01)	(0.01)	(0.04)	(0.07)	(0.06)
2012	0.06***	0.06***	0.06***	-1.02***	-1.12***	-1.11***
	(0.02)	(0.01)	(0.01)	(0.04)	(0.08)	(0.07)
2013	-0.03	-0.03	-0.03	-1.14***	-1.25***	-1.23***
	(0.02)	(0.01)	(0.01)	(0.04)	(0.08)	(0.07)
			Sodium (mg/1	.00g)		
	Othe	er food retail cl	nains	<u> </u>	Walmart	
	No	Stabilized	Truncated	No	Stabilized	Truncate
	IPW	IPW	IPW	IPW	IPW	IPW
	β	β	β	β	β	β
	(SE)	(SE)	(SE)	(SE)	(SE)	(SE)
2000	ref	ref	ref	ref	ref	ref
2001	0.46	0.47	0.47	1.98	3.60	3.65
	(0.48)	(0.47)	(0.47)	(1.41)	(2.16)	(2.10)
2002	0.09	0.09	0.09	4.63***	4.43*	5.32*
	(0.51)	(0.51)	(0.51)	(1.40)	(2.26)	(2.08)
2003	-1.20*	-1.21*	-1.21*	2.18	5.61*	5.20*
	(0.54)	(0.55)	(0.55)	(1.45)	(2.28)	(2.14)
2004	-2.32***	-2.32***	-2.32***	-3.83*	-0.40	-1.38
	(0.55)	(0.55)	(0.55)	(1.52)	(2.44)	(2.26)
2005	-2.12***	-2.14***	-2.14***	-8.78***	-4.87	-5.89*
	(0.57)	(0.58)	(0.58)	(1.62)	(2.56)	(2.40)
2006	-5.21***	-5.25***	-5.25***	-12.73***	-7.92**	-9.05***
	(0.59)	(0.59)	(0.59)	(1.66)	(2.66)	(2.46)
2007	-6.73***	-6.76***	-6.76***	-20.04***	-13.60***	-15.45***
	(0.68)	(0.69)	(0.69)	(1.93)	(3.12)	(2.84)
2008	-5.76***	-5.76***	-5.76***	-18.95***	-12.74***	-15.06***
	(0.79)	(0.81)	(0.81)	(2.20)	(3.57)	(3.19)
2009	-9.64***	-9.64***	-9.64***	-20.17***	-13.32***	-16.07***
	(0.98)	(1.00)	(1.00)	(2.49)	(3.98)	(3.52)
2010	-11.27***	-11.29***	-11.29***	-29.14***	-21.92***	-24.77***

American Journal of Preventive Medicine

	(1.00)	(1.03)	(1.03)	(2.55)	(4.06)	(3.59)
2011	-14.08***	-14.11***	-14.11***	-31.27***	-23.90***	-27.07***
	(1.02)	(1.04)	(1.04)	(2.65)	(4.28)	(3.75)
2012	-17.68***	-17.69***	-17.69***	-36.09***	-27.35***	-31.16***
	(1.08)	(1.11)	(1.11)	(2.87)	(4.70)	(4.08)
2013	-19.89***	-19.88***	-19.88***	-37.90***	-28.80***	-33.00***
	(1.14)	(1.17)	(1.17)	(3.02)	(4.94)	(4.28)

^a IPWs truncated at the upper and lower 0.3rd percentile

Note: Boldface indicates statistical significance, p<0.05, p<0.01; p>0.01; p

Appendix Table 4. Household Characteristics of the Nielsen Homescan Sample in 2000 and 2013^a (n=2,611,125)

```````````````````````````````	2000	2013	$p^{c}$
Household composition			_
Single adult, % ^b	25%	25%	0.070
Multiple adults, no children <18y, %	46%	52%	< 0.001
Adults, with children <18,%	28%	22%	< 0.001
Race/ethnicity			
Non-Hispanic white,%	85%	81%	< 0.001
Hispanic, %	5%	5%	0.243
Non-Hispanic black, %	8%	9%	< 0.001
Non-Hispanic other, %	2%	4%	< 0.001
Maximum household education			
High school or less, %	23%	17%	< 0.001
Some college,%	31%	29%	< 0.001
≥College degree, %	46%	54%	< 0.001
Income ^d			
Federal poverty level ≤130%	6%	10%	< 0.001
>130%	94%	90%	< 0.001
Walmart stores/100,000 people, mean (SD)	0.2 (0.3)	1.0 (0.6)	< 0.001
Unemployment rate, mean (SD)	3.9 (0.9)	7.8 (1.6)	< 0.001
Walmart average food price, mean (SD)	72.2 (12.3)	103.1 (4.4)	< 0.001
Other FRCs average food price, mean (SD)	84.6 (8.4)	117.6 (6.5)	< 0.001

^a University of North Carolina calculation based in part on data reported by Nielsen through its Homescan Services for all food categories, Including beverages and alcohol for the 2000-2013 periods, for the U.S. market. Copyright © 2014, The Nielsen Company.

^b Column percents.

^c For categorical variables, from proportions testing and for continuous variables, from t-tests, comparing households in the sample in 2000 and 2013.

^d Supplemental Nutrition Assistance Program (SNAP) eligibility based on household income using percent of annual federal poverty level

# References

- Hogan JW, Lancaster T. Instrumental variables and inverse probability weighting for causal inference from longitudinal observational studies. *Stat Methods Med Res.* 2004;13(1):17-48. http://dx.doi.org/10.1191/0962280204sm351ra.
- Cole SR, Hernán MA. Constructing inverse probability weights for marginal structural models. *Am J Epidemiol*. 2008;168(6):656-664. <u>http://dx.doi.org/10.1093/aje/kwn164</u>.
- Seaman SR, White IR. Review of inverse probability weighting for dealing with missing data. *Stat Methods Med Res.* 2013;22(3):278-295.

http://dx.doi.org/10.1177/0962280210395740.