Journal of Food Distribution Research
Volume 51, Issue 3, pp. 127–145

Economic and Sociodemographic Drivers Associated with the Decision to Purchase Food Items and Nonalcoholic Beverages from Vending Machines in the United States

Oral Capps, Jr.[a]ⓘ and Rejeana Marie Gvillo[b]

[a]*Executive Professor and Regents Professor, Department of Agricultural Economics,*
600 John Kimbrough Boulevard, Suite 371,
Co-Director of the Agribusiness, Food and Consumer Economics Research Center,
Texas A&M University,
College Station, TX 77843, USA

[b]*Senior Analyst, Farmer's Business Network,*
571 Duncan Loop Road,
Moundville, AL 35474, USA

Abstract

Using data extracted from BLS Consumer Expenditure Surveys between 2009 and 2012, we estimated a probit model concerning the decision made by household heads whether to purchase food items and nonalcoholic beverages from vending machines over a consecutive 2-week period. Key drivers associated with this decision are household income; urbanization; marital status; region; year; age; education level; hours worked; ethnicity; and expenditures made on potato chips and other snacks, candy and chewing gum, food away from home (excluding those made at vending machines), cola drinks, and tobacco products.

Keywords: BLS Consumer Expenditures Surveys, economic and sociodemographic factors, probit analysis, vending machines

ⓘCorresponding author: Tel: (979) 845-8492
Email: ocapps@tamu.edu

Introduction

The U.S. vending machine operators industry consists primarily of candy, food and snack, and hot and cold beverage sales, with a total projected revenue $7.7 billion in 2019. Revenue from 2009 to 2018 ranged between $7.9 billion and $8.9 billion (Figure 1), declining by 12.7% over the period. Total revenues are expected to decline by a further 2.5% in 2019 relative to 2018 (Zheng, 2019).

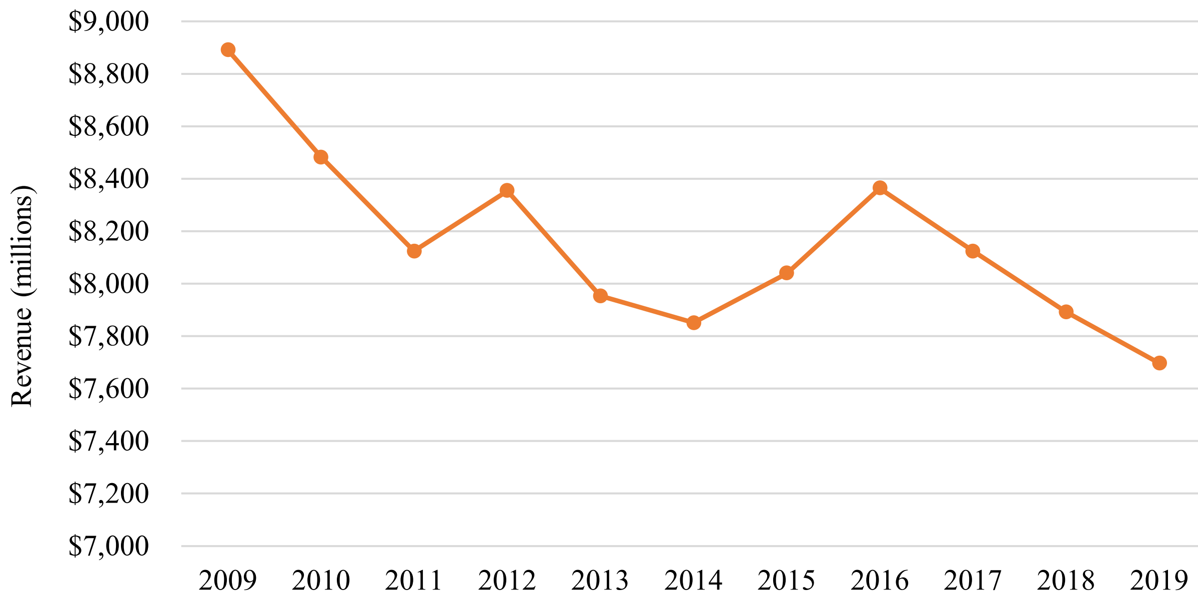

Figure 1. Revenue for U.S. Vending Machine Operators Industry, 2009–2019
Source: Zheng (2019).

The vending machine operators industry (NAICS code 454210) ranks 53rd in the retail trade industry by market size (in terms of revenue) and is the 567th largest industry in the United States (Zheng, 2019). Major U.S. companies include American Food and Vending, AVI Food Systems, and divisions of ARAMARK and Coca-Cola (Dun and Bradstreet, 2020). Among roughly 4.6 million vending machines currently in the United States, close to 60% of vending machine sales are for cold drinks, including soft drinks, juices, and other sugary options. Junk foods, such as soda and chips, typically make up the largest amount of industry revenue, but sales of healthy snacks and beverages are on the rise (Gaille, 2017).

Objectives

The ability to ascertain historical, current, and future patterns of food and beverage consumption is of extreme importance, yet our knowledge of the vending machine operators industry is meager at best. To fill this research void, this study develops a profile of vending machines users. We consider only purchases of food and nonalcoholic beverages made at vending machines.

We use 2009–2012 data from the Bureau of Labor Statistics (BLS) Consumer Expenditure Survey (CES) and a probit regression to accomplish the primary objective of this study.[1] In the probit specification, the dependent variable corresponds to a household's decision to purchase or not to purchase food items and nonalcoholic beverages from vending machines over a consecutive two-week period. The respective model centers attention on economic and sociodemographic factors as explanatory variables, including age, race, education level, income level, household size, employment status, ethnicity, gender, marital status, and region. From this analysis, public health officials, government policy makers, and industry stakeholders will have a better understanding of the factors associated with purchases of food items and nonalcoholic beverages from vending machines, which is presently lacking in the extant literature.

Previous Research

Much of the extant economic literature has centered attention on predominantly three topics: (i) the effect of price and promotion strategies on vending machine purchases; (ii) the nutritional content of foods in vending machines; and (iii) the availability of vending machines in public and private schools (Gvillo, 2014).

French et al. (1997), French et al. (2001), and Hua et al. (2017) investigated the effect of price and promotion strategies on purchases of low-fat snacks from vending machines. Reducing relative prices was effective in promoting lower-fat food choices, and vending machines provided a feasible way of implementing such nutrition interventions. When healthier vending snacks were available, promotional signs also were important to ensure purchases of those items in greater amounts.

Kubik et al. (2003) examined the association between dietary behaviors of young adolescents and purchases made at vending machines. Snacks procured from vending machines were negatively correlated with fruit consumption. Weicha et al. (2006) found that school vending machine use and fast food restaurant visits were associated with overall sugar-sweetened beverage intake. Additionally, French et al. (2003); Lytle et al. (2006); Finkelstein, Hill, and Whitaker (2008); and Pasch et al. (2011) noted that food items and beverages offered in vending machines at schools were high in fat and calories. Further, Cisse-Egbuonye et al. (2016) found that the food items most commonly available in vending machines were predominately foods of minimal nutritional value. Although few school food policies were reported that helped foster healthy food choices among students, Evans et al. (2005) found public support for restricting the availability of unhealthy foods in vending machines.

National data from the 2006 School Health Policies and Programs Study (SHPPS) revealed that 62.4% of middle schools and 85.8% of high schools had at least one vending machine available to students (O'Toole et al., 2007). Park et al. (2010) examined the prevalence of students buying snacks or beverages from school vending machines instead of buying school lunch. Based on data from the 2000 SHPPS, Wechsler et al. (2001) found that nearly all senior high schools, most

[1] These data were the most recent information available to us from the BLS at the time of this analysis. As such, this analysis serves as a benchmark for future analyses concerning vending machine purchases.

middle and junior high schools, and more than one-quarter of elementary schools had access to foods and beverages from vending machines.

To date, the extant literature has focused almost exclusively on vending machine product purchases and the potential health concerns related to such purchases. Unlike previous studies, we focus on the factors affecting the decision to purchase food and nonalcoholic beverages at vending machines. As such, we provide a unique contribution to the economic literature.

Model Development

Binary Choice Probit Model

Models of discrete choice such as probit and logit could be used to examine the factors influencing the decision to purchase food items and nonalcoholic beverages from vending machines. The use of the probit/logit analysis, particularly of binary choices, is well established in the economic literature (Maddala, 1983; McFadden, 1984; Pindyck and Rubinfeld, 1998). Capps and Kramer (1985) demonstrated that the probit and logit models yield similar results in the case of binary choice models. Additionally, since the logistic density function closely resembles the *t*-distribution with seven degrees of freedom (Hanushek and Jackson, 1977), the logit and probit formulations are quite similar. The only difference is that the logistic density has a slightly heavier tail than the standard normal density. In this study, we used a probit regression model.

The use of probit models is commonplace in economic analyses of the food industry (Byrne, Capps, and Saha, 1996; Alviola and Capps, 2010; Capps, Ahad, and Murano, 2017). The probit regression model in this analysis is a binary choice model, where the dependent variable takes on two values—0 for no vending expenditures made and 1 for positive vending expenditure made by reference person *i*. The reference person in the household is the household head who completed the survey. Mathematically, the probit model takes the following form:

(1) $$y_i = \boldsymbol{x}_i'\boldsymbol{\beta} + e_i,$$

where $\boldsymbol{y}_i = 1$ if any vending machine purchase was made by reference person *i*, $\boldsymbol{y}_i = \boldsymbol{0}$ if no vending machine purchase was made by reference person *i*, $\boldsymbol{x}_i'$ is a column vector of explanatory variables, $\boldsymbol{\beta}$ is a vector of parameters associated with the explanatory variables, e_i is the random error, and

(2) $$\Pr\left(y_i = 1 \mid \boldsymbol{x}_i'\right) = \Phi\left(\boldsymbol{x}_i'\boldsymbol{\beta}\right),$$

where Φ is the cumulative distribution function (CDF) of the standard normal distribution. Operationally, the decision to purchase food and nonalcoholic beverages at vending machines is denoted by

$$(3)\quad Vend\_Mach\_Purchase_i = \beta_0 + \beta_1 Age_i + \beta_2 Fam\_size_i + \beta_3 Fincaftx_i + \beta_4 HHhours_i + \beta_5 Male_i + \beta_6 Asian_i + \beta_7 Black_i + \beta_8 White_i + \beta_9 Hispanic_i + \beta_{10} College_i + \beta_{11} Northeast_i + \beta_{12} Midwest_i + \beta_{13} South_i + \beta_{14} Married_i + \beta_{15} Urban_i + \beta_{16} Tobacco_i + \beta_{17} Frsh\_Fruit\_Veg_i + \beta_{18} Candy_i + \beta_{19} Potato\_Chips_i + \beta_{20} Cola\_Drinks_i + \beta_{21} FAFH_i + \beta_{22} Nuts_i + \beta_{23} Jan_i + \beta_{24} Feb_i + \beta_{25} Mar_i + \beta_{26} Apr_i + \beta_{27} May_i + \beta_{28} Jun_i + \beta_{29} Jul_i + \beta_{30} Aug_i + \beta_{31} Sep_i + \beta_{32} Oct_i + \beta_{33} Nov_i + \beta_{34} Year\_2009_i + \beta_{35} Year\_2010_i + \beta_{36} Year\_2011_i + e_i .$$

Table 1 defines the dependent variable and the explanatory variables in the probit specification. Previous research generally depicts snack food and beverage items from vending machines as unhealthy (French et al., 2003; Evans et al., 2005; Lytle et al., 2006; Finkelstein, Hill, and Whitaker, 2008; Pasch et al., 2011; Cisse-Egbuonye et al., 2016). As such, we hypothesize that expenditures on tobacco products, candy, potato chips, and cola drinks—which are generally considered unhealthy foods (Chaloupka and Warner, 2000; Drewnowski, 2003; Dharmasena and Capps, 2011)—are positively related to the decision to purchase from vending machines. In contrast, expenditures on fresh fruits, vegetables, and nuts—typically regarded as healthy items (Drewnowski and Darmon, 2005; Jones, 2010)—are hypothesized to be negatively related to purchases made from vending machines.

Park et al. (2010) found that age, race, and Hispanic ethnicity were key factors of students buying snacks or vegetables from school vending machines. Therefore, we hypothesize that younger household heads and households of Hispanic ethnicity are more likely to purchase food and nonalcoholic beverages at vending machines. We also expect race to influence the decision to purchase from vending machines. Further, because education level often is positively associated with health consciousness (Alviola and Capps, 2010), we hypothesize that this sociodemographic factor is inversely related to the decision to purchase from vending machines. We hypothesize that the number of hours worked and expenditures on food away from home are positively related to the decision to purchase from vending machines in accordance with the opportunity cost of time (Byrne, Capps, and Saha, 1996). Hill and Lynchehaun (2002) and Dharmasena and Capps (2014) identified various cultural and socio-economic factors—including age, ethnicity, income, education, gender, presence of children, marital status, region, and race—influencing consumer preferences. Hence, we hypothesize that household income, household size, gender, marital status, and region are also determinants of the decision to purchase food and nonalcoholic beverages from vending machines. Finally, given the coverage of the data over 2009–2012, we capture seasonal trends through the use of monthly dummy variables and year-to-year trends through the use of

Table 1. Description and Descriptive Statistics of the Dependent Variable and Explanatory Variables included the Probit Regression

Variable	Definition	Mean
Age	Age of the reference person in the household (the household head who completed the survey)	50
Asian	= 1 if the race of the reference person is Asian; 0 otherwise	0.0429
Black	= 1 if the race of the reference person is Black; 0 otherwise	0.1178
Candy	Consecutive 2-week expenditure on candy and chewing gum	$3.13
Potato_Chips	Consecutive 2-week expenditure on potato chips and other snacks	$4.28
Cola_Drinks	Consecutive 2-week expenditure on cola drinks	$3.15
College	= 1 if the reference person has recorded at least some college education; 0 otherwise	0.6291
Fam_size	Number of members in the consumer unit (CU)	2.52
Fincaftx	Amount of CU income after taxes in past 12 months	$60,064
FAFH	Consecutive 2-week expenditure on food away from home, excluding monies spent at a vending machine	$91.15
Female	1 if the reference person is female; 0 otherwise (reference/base category)	0.5394
Frsh_Fruit_Veg	Consecutive 2-week expenditure on fresh fruits and fresh vegetables	$18.68
HHhours	Total number of hours usually worked per week by the reference person and spouse	41
Hispanic	= 1 if the reference person is Hispanic; 0 otherwise	0.1279
Male	= 1 if the reference person is male; 0 otherwise	0.4606
Married	= 1 if the reference person is married; 0 otherwise	0.5213
Midwest	= 1 if the CU resides in the Midwest; 0 otherwise	0.2399
Month_i	= 1 for recorded month *i* of CU vending machine expenditure; 0	
Jan	otherwise	0.0893
Feb		0.0789
Mar		0.0866
Apr		0.0914
May		0.0886
Jun		0.0910
Jul		0.0804
Aug		0.0818
Sep		0.0840
Oct		0.0860
Nov		0.0801
Dec	(reference/base category)	0.0619
No College	= 1 if the reference person has recorded no college education; 0 otherwise (reference/base category)	0.3709
Non-Hispanic	= 1 if the reference person is non-Hispanic; 0 otherwise (reference/base category)	0.8721

Variable	Definition	Mean
Non-Married	= 1 if the reference person is not married; 0 otherwise (reference/base category)	0.4787
Northeast	= 1 if the CU resides in the Northeast; 0 otherwise	0.1927
Nuts	Consecutive 2-week expenditure for nuts	$1.53
Other Races	= 1 if the race of the reference person is not white, Black, or Asian; 0 otherwise (reference/base category)	0.0194
Rural	= 1 if the CU resides in a rural area, 0 otherwise (reference/base category)	0.0535
South	= 1 if the CU resides in the South; 0 otherwise	0.3520
Tobacco	Consecutive 2-week expenditure on tobacco products	$9.46
Urban	= 1 if the CU resides in an urban area; 0 otherwise	0.9465
Vend_Mach_Purchase	= 1 if a food item or nonalcoholic beverage is purchased; 0 otherwise (dependent variable in the probit model)	0.2040
West	= 1 if the CU resides in the West; 0 otherwise (reference/base category)	0.2154
White	= 1 if the reference person is white; 0 otherwise	0.8199
$Year_i$	= 1 for recorded year *i* of CU vending machine expenditure; 0	
Year_2009	otherwise	0.2502
Year_2010		0.2533
Year_2011		0.2448
Year_2012	(reference/base category)	0.2517

Source: Calculations by the authors using EVIEWS v. 11 (IHS Global, Inc., 2020).

yearly dummy variables. We hypothesize that seasonal differences and year-to-year differences are evident in the decision to purchase food and nonalcoholic beverages from vending machines.

Data

The source of data for this analysis is the Consumer Expenditure Survey (CES), available from the Bureau of Labor Statistics (BLS). This survey includes two separate surveys—the Interview Survey and the Diary Survey. While both surveys provide information on American consumers' buying habits, the Diary Survey is of interest for this analysis.

The Diary Survey comprises several data files; for this study, we use the expenditure and family files. The expenditure files consist of a "diary" of expenditures in which the respondent records information for two consecutive 1-week periods. The family files contain demographic information and characteristics of the respondents, typically referred to as consumer units (CUs). The BLS defines a CU as comprising either (i) all members of a particular household who are related by blood, marriage, adoption, or other legal arrangements; (ii) a person living alone or sharing a household with others or living as a roomer in a private home or lodging house or in permanent living quarters in a hotel or motel, but who is financially independent; or (iii) two or more persons living together who use their income to make joint expenditure decisions. In essence, the term CU is synonymous with the term household.

For each household, there are two weekly observation periods. In this study, we merge the respective expenditures for the two consecutive 1-week periods for each household. The time period corresponds to 2009–2012, the most recent data available to us at the time of this analysis. Nonetheless, the most recent CES data are for 2018 (U.S. Bureau of Labor Statistics, 2020). As such, this analysis provides a baseline or benchmark study concerning vending machine expenditures made by U.S. households that could help as a reference for future studies using more recent data.

The expenditure files do not contain quantity or price information, only information on household expenditures over two consecutive weeks. Several vending machine expenditures are recorded in the Diary Survey, including breakfast, lunch, dinner, and snacks purchased from vending machines as well as tobacco or alcohol purchased from vending machines. Here, we focus exclusively on food and nonalcoholic beverage purchases at vending machines.

The dataset used in this study consists of 23,333 observations compiling 4 years of data from 2009 to 2012. Each observation corresponds to a unique household identification number. Thus, the dataset is equivalent to a cross-sectional representation of U.S. households across the 4-year period from 2009 to 2012. Prior to data cleaning, the original sample size was 27,225 observations. We dropped households with insufficient information and removed outliers associated with income and various food expenditures.[2]

In Table 1, we summarize the descriptive statistics (mean values only) for the sample of households included in our analysis. About 20% of the sample, or 4,670 of the 23,333 households in the sample, had nonzero (positive) vending expenditures associated with food items and nonalcoholic beverages over a 2-week period. Across all households, the average amount spent over two consecutive weeks at vending machines for food and nonalcoholic beverages was $1.39. For those households that made vending machine purchases, the average amount spent over a two-week period was $6.82.

The average age of the respondent (*Age*) in the sample was 50. Household size (*Fam_size*) was about 2.5, and the average income (*Fincaftx*) was roughly $60,000. Household hours worked (*HHhours*) combined for all members was, on average, 41. About 63% of the sample had at least some college education (*College*), slightly more than 46% of the sample were male, nearly 95% were in urban areas, about 52% were married, and nearly 13% were of Hispanic ethnicity. Further, roughly 82% of the sample were white, nearly 12% were Black, and about 4% were Asian. About 19% of the sample were located in the Northeast, 24% were located in the Midwest, 35% were located in the South, and almost 22% were located in the West.

On average, consecutive two-week expenditures on food-away-from-home, excluding vending machine expenditures, amounted to $91.15. Consecutive two-week average expenditures over the 2009–2012 period for nuts, potato chips and other snacks, candy and chewing gum, cola drinks,

[2] Households were dropped from the dataset if income or expenditures exceeded the mean value ± 3 times the corresponding standard deviation.

fresh fruits and vegetables, and tobacco products were $1.53, $4.28, $3.13, $3.15, $18.68, and $9.46, respectively.

Our sample of households is representative of the U.S. population during the 2009–2012 period. To support this contention, we compare the sociodemographic characteristics of our sample with population statistics provided by the U.S. Census Bureau for 2010 (U.S. Census Bureau, 2012, 2020; DeNavas-Walt, Proctor, and Smith, 2011). According to the 2010 Census, average household income was $58,500, slightly below the average income of our sample (see Table 2); household size was 2.34, in line with our average household size of 2.52. Further, similar percentages of race, region, age, gender, ethnicity, and marital status are evident. However, our sample had a much lower percentage of households in rural areas and a much higher percentage of households in urban areas compared to the 2010 Census. Finally, in our sample, the percentage of households whose heads received at least some college education was 63%, compared to 55% from the 2010 Census. Aside from population density and education of the household head, our sample of households matches up well to the U.S. population as represented by the 2010 Census.

Table 2. Representativeness of the Sample to the US Population According to the 2010 U.S. Census

Sociodemographic Characteristic	2010 U.S. Census	Sample
White (%)	80.17	81.99
Black (%)	13.34	11.78
Asian (%)	5.02	4.29
Other (%)	1.46	1.94
Household size	2.34	2.52
Age > 25	47	50
Northeast (%)	17.92	19.27
Midwest (%)	21.68	23.99
South (%)	37.10	35.2
West (%)	23.30	21.54
Household income	$58,500	$60,064
Female (%)	50.87	53.94
Male (%)	49.13	46.06
Hispanic (%)	16.27	12.79
Not Hispanic (%)	83.73	87.21
Married (%)	56.58	52.13
Not married (%)	43.42	47.87
Rural (%)	19.27	5.35
Urban (%)	80.73	94.65
At least some college education (%)	55.24	62.91
No college education (%)	44.76	37.09

Sources: DeNavas-Walt, Proctor, and Smith (2011), U.S. Census Bureau (2012, 2020).

Specification Tests

A concern in this analysis is that that the explanatory variables in the probit specification associated with expenditures on food away from home, nuts, potato chips and other snacks, candy and chewing gum, cola drinks, fresh fruits and vegetables, and tobacco products may be endogenous. If so, then the estimated coefficients are inconsistent (Greene, 2012). Using the Durbin–Wu–Hausman test (Guo et al., 2018), we reject the null hypothesis that the respective expenditure variables in the set of explanatory variables are exogenous.

Hence, to mitigate the endogeneity issue associated with each of these right-side expenditure variables, we employ a two-stage Tobit procedure, which we choose to deal with the issue of the censored response of the right-side expenditure variables. We incorporate instrument variables to circumvent the endogeneity issue (Sargan, 1958; Davidson and MacKinnon, 1993). In the first stage, each of the expenditure categories are expressed as a function of the sociodemographic variables (the instrument variables): hours worked, region, urbanization, race, Hispanic, education level, gender, and marital status as well as income, income squared, family size, family size squared, the interaction of income and family size, and monthly dummy variables. From this first-stage estimation process, we subsequently obtain predicted values of unconditional expenditures by way of calculating $Ey = G'_{i}\gamma F(z) + \sigma f(z)$, where $z = G'\gamma/\sigma, f(z)$ is the normal density function with standard deviation σ, and $F(z)$ is the cumulative normal distribution function (McDonald and Moffitt, 1980); G corresponds to the column vector of the aforementioned instrument variables, and γ represents the vector of parameters associated with the set of instrument variables. In turn, these predicted values (Ey) were used as the explanatory variables for expenditures related to nuts, fresh fruits and vegetables, tobacco products, candy and chewing gum, cola drinks, potato chips, and food away from home in the probit regression.[3]

We used variance inflation factors, condition indices, and variance proportions to examine potential collinearity issues in the probit model (Belsley, Kuh, and Welsch, 1980). No degrading collinearity issues were evident from this examination.

Results

Upon mitigating the endogeneity issues previously discussed, the estimation of the probit model was done using a maximum likelihood procedure from the software package EVIEWS v. 11 (IHS Global, Inc., 2020). Table 3 reports the parameter estimates, standard errors, and associated p-values of the respective explanatory variables in the probit model. The goodness-of-fit statistic, McFadden's R^2, is 0.0670. We tested the overall significance of the probit regression model using a likelihood ratio test. Specifically, we tested the null hypothesis that all estimated coefficients, except the intercept coefficient, are jointly equal to 0. The p-value associated with the likelihood ratio test (Table 3) suggests the null hypothesis is rejected and, therefore, at least one of the estimated coefficients is statistically different from 0.

[3] To conform to space limitations, details associated with the first-stage Tobit equations are available from the authors upon request.

Table 3. Parameter Estimates, Standard Errors, and *p*-Values Associated with the Estimation of the Binary Probit Regression

Variable	Coefficient	Std. Error	*z*-Statistic	*p*-Value
C	-0.8538	0.1023	-8.34	**0.0000**
Fincaftx	-5.30E-07	1.78E-07	-2.97	**0.0029**
White	-0.0025	0.0688	-0.04	0.9707
Black	0.0426	0.0743	0.57	0.5665
Asian	-0.0427	0.0828	-0.52	0.6063
Urban	-0.0941	0.0431	-2.18	**0.0291**
Married	-0.0653	0.0259	-2.52	**0.0117**
Northeast	0.0221	0.0312	0.71	0.4800
Midwest	0.1705	0.0294	5.81	**0.0000**
South	0.0643	0.0277	2.33	**0.0200**
Jan	-0.0565	0.0515	-1.10	0.2732
Feb	-0.0192	0.0524	-0.37	0.7133
Mar	0.0316	0.0508	0.62	0.5345
Apr	0.0094	0.0505	0.19	0.8519
May	0.0647	0.0504	1.28	0.1999
Jun	0.0462	0.0503	0.92	0.3582
Jul	0.0550	0.0516	1.07	0.2868
Aug	0.0232	0.0517	0.45	0.6544
Sep	0.0877	0.0510	1.72	0.0853
Oct	0.0227	0.0510	0.45	0.6557
Nov	-0.0136	0.0521	-0.26	0.7933
Year_2009	0.0988	0.0272	3.64	**0.0003**
Year_2010	0.0212	0.0274	0.78	0.4382
Year_2011	-0.0285	0.0277	-1.03	0.3028
Age	-0.0089	0.0007	-12.66	**0.0000**
Fam_size	0.0080	0.0080	1.00	0.3170
College	0.0493	0.0220	2.23	**0.0254**
Male	-0.0113	0.0197	-0.57	0.5665
HHhours	0.0037	0.0004	8.73	**0.0000**
Hispanic	0.0719	0.0312	2.30	**0.0212**
Nuts	-0.0026	0.0025	-1.06	0.2883
Potato_Chips	0.0062	0.0015	4.22	**0.0000**
Candy	0.0028	0.0014	1.94	**0.0519**
FAFH	0.0018	8.20E-05	21.70	**0.0000**
Cola_Drinks	0.0114	0.0016	6.97	**0.0000**
Frsh_Fruit_Veg	-0.0003	0.0005	-0.52	0.6060
Tobacco	0.0023	0.0003	7.80	**0.0000**
McFadden R^2	0.0670			
LR statistic	1,582			
Prob(LR statistic)	0.0000			
Observations with Dep = 0	18,573	Total observations	23,333	
Observations with Dep = 1	4,760			

Notes: Bold *p*-values indicate statistical significance at the 0.05 level.

The level of statistical significance chosen for this analysis is 0.05. All estimated coefficients statistically different from 0 are in bold. The key drivers associated with the decision to purchase food items and nonalcoholic beverages from vending machines are: (i) income; (ii) urbanization; (iii) marital status; (iv) region; (v) year; (vi) age; (vii) education level; (viii) hours worked; (ix) ethnicity; and (x) expenditures made on potato chips and other snacks, candy and chewing gum, food away from home excluding those made at vending machines, cola drinks, and tobacco products.

Households with lower incomes and in rural areas with individuals who are not married are more likely to purchase food and nonalcoholic beverages at vending machines than households with higher incomes and in urban areas with married individuals. Household heads who are Hispanic and college-educated and households in the Midwest and the South also are more likely to purchase food and nonalcoholic beverages at vending machines than household heads who are not Hispanic and not college-educated and households in the West. Moreover, households with younger heads who work more hours are more likely to purchase food and nonalcoholic beverages at vending machines than households with older heads who work fewer hours. As hypothesized, households that expend more on potato chips and other snacks, candy and chewing gum, food away from home, cola drinks, and tobacco products are more likely to purchase food and nonalcoholic beverages at vending machines.

No differences across months are evident in the likelihood of purchasing food and nonalcoholic beverages. Relative to 2012, the likelihood of purchasing at vending machines was statistically the same in 2010 and 2011. However, the likelihood of purchasing at vending machines was higher in 2009 relative to 2012.

Marginal effects, exhibited in Table 4, provide insight as to how changes in the right-side variables affect the probability of purchasing from a vending machine. To calculate the marginal effect for any explanatory variable, the estimated coefficient associated with that variable is multiplied by the standard normal density function, $f(\boldsymbol{x}_i'\boldsymbol{\beta})$. Because the marginal effects vary from observation to observation, they are calculated at the sample means for each of the explanatory variables in the probit model. We highlight the marginal effects for the statistically significant sociodemographic binary variables as well as for the continuous variables associated with the decision to purchase food and nonalcoholic beverages at vending machines.

As the household head ages each year, the probability of purchasing food and nonalcoholic beverages at vending machines is lower by about 0.2%. For college-educated household heads, the probability of purchasing at vending machines is higher by 1.3% relative to noncollege educated individuals. For household heads of Hispanic ethnicity, the probability of purchasing at vending machines is higher by 1.9% relative to individuals of non-Hispanic ethnicity. For households with married individuals, the probability of purchasing at vending machines is lower by 1.8% relative to households without married individuals. Households in urban areas are 2.5% less likely to purchase food items and nonalcoholic beverages at vending machines than those in nonurban areas. CUs in the Midwest are 4.6% more likely to make purchases at vending machines than those in the West. Similarly, households in the South are 1.7% more likely to make purchases at vending machines than those in the West. The likelihood of purchasing at vending machines was higher in 2009 by 2.7% relative to 2012.

Table 4. Marginal Effects and Elasticities Associated with the Probit Regression Estimates

Variable	Marginal Effects	Elasticities
Age	**-0.0024**	
White	-0.0007	
Black	0.0115	
Asian	-0.0115	
College	**0.0133**	
Fam_size	0.0022	
Hispanic	**0.0194**	
Male	-0.0030	
Married	**-0.0176**	
Midwest	**0.0460**	
Northeast	0.0060	
South	**0.0174**	
Jan	-0.0152	
Feb	-0.0052	
Mar	0.0085	
Apr	0.0025	
May	0.0174	
Jun	0.0125	
Jul	0.0148	
Aug	0.0062	
Sep	0.0237	
Oct	0.0061	
Nov	-0.0037	
Urban	**-0.0254**	
Year_2009	**0.0267**	
Year_2010	0.0057	
Year_2011	-0.0077	
Candy	**0.000752**	**0.0125**
Potato_Chips	**0.001669**	**0.0380**
Cola_Drinks	**0.003084**	**0.0516**
Fincaftx	**-0.000000143**	**-0.0456**
FAFH	**0.000480**	**0.2324**
Frsh_Fruit_Veg	-0.000069	-0.0068
Nuts	-0.000712	-0.0058
HHhours	**0.000996**	**0.2166**
Tobacco	**0.000618**	**0.0311**

Notes: Bold values are associated with statistically significant coefficients of the respective sociodemographic indicator variables as well as the nondiscrete variables.
Source: Calculations by the authors at the sample means of the data.

We also provide the elasticity or the percentage change in the probability of purchasing at vending machines attributed to a 1% change in the respective continuous variables (except for age and family size) in the probit model. The elasticity is always the product of the marginal effect and the ratio of the relevant continuous explanatory variable to the dependent variable. In our study, the appropriate value of the dependent variable is the probability that a food or nonalcoholic beverage purchase at a vending machine will be made. This probability is calculated at the sample means.

If household income were to change by 1%, the probability of purchasing at vending machines would decrease by 0.05%. Moreover, if the number of hours worked by household heads were to change by 1%, the probability of purchasing at vending machines would change by 0.22%. A 1% change in household expenditures related to candy and chewing gum, potato chips and other snacks, cola drinks, food away from home, and tobacco products yields a 0.01% change, a 0.04% change, a 0.05% change, a 0.23% change, and a 0.03% change, respectively, in the probability of purchasing at vending machines.

About 20% of the survey respondents purchased food and nonalcoholic beverages at vending machines. Hence, in the derivation of the prediction success (Table 5), the cutoff probability for classification purposes is 0.20. That is, we predict that the *i*th reference person will purchase at a vending machine if the probability of doing so exceeds 0.20. In agreement with Greene (2012, p. 658), "in general any prediction rule will make two types of errors; it will incorrectly classify zeros as one and ones as zeros." For binary choice models, to the best of our knowledge, no benchmark exists regarding correct classifications. Within the sample, the probit model correctly classifies the decision to not make purchases with 62.21% accuracy (11,555 out of 18,573). Within the sample, the probit model correctly classifies the decision to purchase with 65.48% accuracy (3,117 out of 4,760). Overall, within the sample, the model correctly classifies all decisions 14,672 out of 23,333 times, with 62.88% accuracy.

Table 5. Expectation-Prediction Evaluation of the Probit Model

	Dep = 0	**Dep = 1**	**Total**
P(Dep = 1) ≤ C	11,555	1,643	13,198
P(Dep = 1) > C	7,018	3,117	10,135
Total	18,573	4,760	23,333
Correct	11,555	3,117	14,672
% Correct	62.21	65.48	62.88

Notes: Success cutoff: C = 0.2040029. Dep = 0 indicates nonpurchase of food items or nonalcoholic beverages from vending machines. Dep = 1 indicates purchase of food items or nonalcoholic beverages from vending machines.

Concluding Remarks

To date, the extant literature has focused almost exclusively on vending machine product purchases and the potential health concerns related to such purchases. The purpose of this study was to examine economic and sociodemographic factors that influence individuals' decision to purchase food items and nonalcoholic beverages from a vending machine. Using data extracted from BLS Consumer Expenditure surveys over the period from 2009 to 2012, a probit model was estimated incorporating instrumental variables to address endogeneity issues. Results from this study could help vending machine operators to increase sales by targeting those individuals more likely to purchase food and nonalcoholic beverages from vending machines. Lower-income households with younger household heads who reside in rural areas, are Hispanic, are college-educated, reside in the Midwest and the South, and work more hours are more likely to make purchases from vending machines. Additionally, households that expend more on potato chips and other snacks, candy and chewing gum, food away from home, cola drinks, and tobacco products are more likely to purchase food and nonalcoholic beverages at vending machines.

This research provides a benchmark for future studies concerning purchases from vending machines. Additional research with more current data should be undertaken to examine whether the results of this study are robust. Because of public health concerns, the Food and Drug Administration (FDA) sets rules regarding calorie disclosure required by the Affordable Care Act enacted in March 2010. Moreover, the U.S. Department of Agriculture (USDA) proposed rules regarding the items allowed in school vending machines in order to help students make healthier snack choices (Vending Market Watch, 2019). Beginning in 2014, the vending machines industry was required to provide calorie counts of their snack foods and beverages.

Understanding why more educated individuals are choosing to purchase from a vending machine even though the items for sale are more or less unhealthy is another area to explore. As well, future research efforts should incorporate experiments, seeing whether having healthier options available for purchase in a vending machine leads to healthier items actually being purchased.

There are several limitations to this study. We are not able to discern the impacts of price on the decision to purchase food items and nonalcoholic beverages as prices were not available. Another limitation is that the data used in our analysis are self-reported. As such, measurement error may exist attributed to self-reporting. Further, exploring the use of a Tobit model or a Heckman two-step model is warranted to obtain information on unconditional and conditional demands of items purchased from vending machines. While this study has limitations, we have answered a question that had not previously been addressed, namely what economic and sociodemographic factors affect purchases of food items and nonalcoholic beverages from vending machines.

Acknowledgment

We gratefully acknowledge comments made by anonymous reviewers. Any remaining errors or omissions are the sole responsibility of the authors.

References

Alviola, P.A., and O. Capps, Jr. 2010. "Household Demand Analysis of Organic and Conventional Fluid Milk in the United States Based on the 2004 Nielsen Homescan Panel." *Agribusiness: An International Journal* 26(3): 369–388.

Belsley, D.A., K. Kuh, and R.E. Welsch. 1980. *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. New York, NY: John Wiley & Sons.

Byrne, P.J., O. Capps, Jr., and A. Saha. 1996. "Analysis of Food-Away-from-Home Expenditure Patterns for U.S. Households 1982–1989." *American Journal of Agricultural Economics* 78(3): 614–627.

Capps, Jr., O., A. Ahad, and P.S. Murano. 2017. "Understanding Spending Habits and Buying Behavior of the American Muslim Community: A Pilot Study." *Journal of Food Distribution Research* 48(3): 51–74.

Capps, Jr., O., and R.A. Kramer. 1985. "Analysis of Food Stamp Participation Using Qualitative Choice Models." *American Journal of Agricultural Economics* 67(1): 49–59.

Chaloupka, F.J., and K.E. Warner. 2000. "The Economics of Smoking." In A.J. Culyer and J.P. Newhouse, eds. *Handbook of Health Economics*, Vol. 1. Amsterdam, Netherlands: Elsevier, pp. 1539–1627.

Cisse-Egbuonye, N., S. Liles, K.E. Schmitz, N.O.F. Kassem, V.L. Irvin, and M.F. Hovell. 2016. "Availability of Vending Machines and School Stores in California Schools." *Journal of School Health* 86(1): 48–53.

Davidson, R., and J.G. MacKinnon. 1993. *Estimation and Inference in Econometrics.* New York, NY: Oxford University Press.

DeNavas-Walt, C., B.D. Proctor, and J.C. Smith. 2011. *U.S. Census Bureau, Current Population Reports, P60-239, Income, Poverty, and Health Insurance Coverage in the United States: 2010*. Washington, DC: U.S. Government Printing Office.

Dharmasena, S., and O. Capps, Jr. 2011. "Intended and Unintended Consequences of a Proposed National Tax on Sugar-Sweetened Beverages to Combat the U.S. Obesity Problem." *Health Economics* 21(6): 669–694.

Dharmasena, S., and O. Capps, Jr. 2014. "Unraveling Demand for Dairy-Alternative Beverages in the United States: The Case of Soymilk." *Agricultural and Resource Economics Review* 43(1): 140–157.

Drewnowski, A. 2003. "Fat and Sugar: An Economic Analysis." *Journal of Nutrition* 133(3): 838S–840S.

Drewnowski, A., and N. Darmon. 2005. "The Economics of Obesity: Dietary Energy Density and Energy Cost." *American Journal of Clinical Nutrition* 82(1 Suppl): 265S–273S.

Dun and Bradstreet. 2020. "Vending Machine Operators Industry Insights." Available online: https://www.dnb.com/business-directory/industry-analysis.vending-machine-operators.html.

Evans, D.W., E.A. Finkelstein, D.B. Kamerowm and J.M. Renaud. 2005. "Public Perceptions of Childhood Obesity." *American Journal of Preventative Medicine* 28(1): 26–32.

Finkelstein D.M., E.L. Hill, and R.C. Whitaker. 2008. "School Food Environments and Policies in U.S. Public Schools." *Pediatrics* 122(1): 251–259.

French, S.A., R.W. Jeffrey, M. Story, K.K. Breitlow, J.S. Baxter, P. Hannan, and M.P. Snyder. 2001. "Pricing and Promotion Effects on Low-Fat Vending Snack Purchases: The CHIPS Survey." *American Journal of Public Health* 91(1): 112–117.

French, S.A., R.W. Jeffrey, M. Story, P. Hannan, and M.P. Snyder. 1997. "A Pricing Strategy to Promote Low-Fat Snack Choices through Vending Machines." *American Journal of Public Health* 87(5): 849–851.

French, S.A., M. Story, J.A. Fulkerson, and A.F. Gerlach. 2003. "Food Environment in Secondary Schools: A La Carte, Vending Machines, and Food Policies and Practices." *American Journal of Public Health* 93(7): 1161–1167.

Gaille, B. 2017. "25 Intriguing Vending Machine Sales Statistics." *Brandon Gaille Small Business and Marketing Advice*. Available online: https://brandongaille.com/23-intriguing-vending-machine-sales-statistics/.

Greene, W.H. 2012. *Econometric Analysis*, 7th ed. New York, NY: Pearson.

Guo, Z., H. Kang, T.T. Cai, and D.S. Small. 2018. "Testing Endogeneity with High Dimensional Covariates." *Journal of Econometrics* 207(1): 175–187.

Gvillo, R.M., 2014. "Milk Advertising, Vending Machine Purchases, and Their Health Implications." Ph.D. Dissertation, Texas A&M University.

Hanushek, J., and E. Jackson. 1977. *Statistical Methods for Social Scientists*. New York, NY: Academic Press.

Hill, H., and F. Lynchehaun. 2002. "Organic Milk: Attitudes and Consumption Patterns." *British Food Journal* 104(7): 526–542.

Hua. S.V., L. Kimmel, M. Van Emmenes, R. Taherian, G. Remer, A. Millman, and J.R. Ickovics. 2017. "Health Promotion and Healthier Products Increase Vending Purchases: A Randomized Factorial Trial." *Journal of the Academy of Nutrition Dietetics* 117(7): 1057–1065.

IHS Global, Inc. 2020. EVIEWS v. 11. Available online: https://www.eviews.com .

Jones, E. 2010. "An Economic Analysis of Fresh Fruit and Vegetable Consumption: Implications for Overweight and Obesity among Higher- and Lower-Income Consumers." *Journal of Food Distribution Research* 41(2): 86–112.

Kubik, M.Y., L.A. Lytle, P.J. Hannan, C.L. Perry, and M. Story. 2003. "The Association of the School Food Environment with Dietary Behaviors of Young Adolescents." *American Journal of Public Health* 93(7): 1168–1173.

Lytle, L.A., M.Y. Kubik, C. Perry, M. Story, A.S. Birnbaum, and D.M. Murray. 2006. "Influencing Healthful Food Choices in School and Home Environments: Results from the TEENS Study." *Preventative Medicine* 43(1): 8–13.

Maddala, G.S. 1983. *Limited Dependent and Qualitative Variables in Econometrics.* Cambridge, UK: Cambridge University Press.

McDonald, J.F., and R.A. Moffitt. 1980. "The Uses of Tobit Analysis." *Review of Econometrics and Statistics* 62(2): 318–321.

McFadden, D. 1984. "Econometric Analysis of Qualitative Response Models." In Z. Griliches and M. Intriligator, eds. *Handbook of Econometrics,* Vol. 2. Amsterdam, Netherlands: North Holland, pp. 1395–1457.

O'Toole, T.P., S. Anderson, C. Miller, and J. Guthrie. 2007. "Nutrition Services and Foods and Beverages Available at School: Results from the School Health Policies and Programs Study." *Journal of School Health* 77(8): 500–521.

Park, S., W.M. Sappenfield, Y. Huang, B. Sherry, and D.M. Bensyl. 2010. "The Impact of the Availability of School Vending Machines on Eating Behavior During Lunch: The Young Physical Activity and Nutrition Survey." *Journal of the American Dietetic Association* 110(10): 1532–1536.

Pasch, K.E., L.A. Lytle, A.C. Samuelson, K. Farbakhsh, M.Y. Kubik, and C.D. Patnode. 2011. "Are School Vending Machines Loaded with Calories and Fat: An Assessment of 106 Middle and High Schools." *Journal of School Health* 81(4): 212–218.

Pindyck, R.S., and D.L. Rubinfield. 1998. *Econometric Models and Economic Forecasts,* 4th ed. New York, NY: McGraw-Hill.

Sargan, J.D. 1958. "The Estimation of Economic Relationships Using Instrument Variables." *Econometrica* 26(3): 393–415.

U.S. Department of Labor, Bureau of Labor Statistics. 2020. *Consumer Expenditure Survey, Diary Survey*. Available online: https://www.bls.gov.

U.S. Census Bureau. 2012. *2010 Census of Population and Housing*: *Population and Housing Unit Counts, CPH-2-1, United States Summary.* Washington, DC: U.S. Government Printing Office.

U.S. Census Bureau. 2020. *Current Population Reports.* Available online: https://www.census.gov/main/www/cprs.html.

Vending Market Watch. 2019. "Vending Machine Operators Industry in the United States." Market Research Report, July.

Wechsler, H., N.D. Brener, S. Kuester, and C. Miller. 2001. "Food Service and Foods and Beverages Available at School: Results from the School Health Policies and Programs Study 2000." *Journal of School Health* 71(7): 313–324.

Wiecha, J.L., D. Finkelstein, P.J. Troped, M. Fragala, and K.E. Peterson. 2006. "School Vending Machine Use and Fast-Food Restaurant Use Are Associated with Sugar-Sweetened Beverage Intake in Youth." *Journal of the American Dietetic Association* 106(10): 1624–1630.

Zheng, Q. 2019. "Vending Machine Operators in the U.S." IBIS World Industry Report 45421, July 2019. Available online at: http://www.ibisworld.com/industry-statistics/market-size/vending-machine-operators-united-states.