JOURNAL of FOOD DISTRIBUTION RESEARCH

Volume 55 / Issue 2 / July 2024



Published by

FDRS
FOOD DISTRIBUTION RESEARCH SOCIETY

http://www.fdrsinc.org

Journal of Food Distribution Research
Volume 55, Issue 2
July 2024

ISSN 0047-245X

The *Journal of Food Distribution Research* has an applied, problem-oriented focus on the flow of food products and services through wholesale and retail distribution systems. Related areas of interest include patterns of consumption, impacts of technology on processing and manufacturing, packaging and transport, data and information systems in the food and agricultural industry, market development, and international trade in food products and agricultural commodities. Business, agricultural, and applied economic applications are encouraged. Acceptable methodologies include survey, review, and critique; analysis and synthesis of previous research; econometric or other statistical analysis; and case studies. Teaching cases will be considered. Issues on special topics may be published based on requests or on the editors' initiative. Potential usefulness to a broad range of agricultural and business economists is an important criterion for publication.

The *Journal of Food Distribution Research* (JFDR) is a publication of the Food Distribution Research Society, Inc. (FDRS). The journal is published three times a year (March, July, and November). JFDR is refereed in its July and November issues. A third, non-refereed issue contains Research Reports and Research Updates presented at FDRS's annual conference. Members and subscribers also receive the Food Distribution Research Society Newsletter, normally published twice a year.

JFDR is refereed by a review board of qualified professionals (see Editorial Review Board, at left). Manuscripts should be submitted to the FDRS editors (see back cover for Manuscript Submission Guidelines).

The FDRS accepts advertising of materials considered pertinent to the purposes of the Society for both the journal and the newsletter. Contact the V.P. for Membership for more information.

Lifetime membership is $400; one-year professional membership is $45; three-year professional membership is $120; student membership is $15 a year; junior membership (graduated in last five years) is $15 and company/business membership is $140.

Food Distribution Research Society
http://www.fdrsinc.org/

Indexing and Abstracting
Articles in *Journal of Food Distribution Research* are indexed or cataloged in Ag Econ Search, Google Scholar, RePEc, Econ Papers, SCOPUS, IDEAS, and CAB International

Send change of address notifications to
Samuel Zapata
Texas AgriLife Extension Service
2401 E. Business 83
Weslaco, TX 78596
Phone: (956) 5581;
e-mail: samuel.zapata@ag.tamu.edu

Food Distribution Research Society

Journal of Food Distribution Research
Volume 55, Issue 2

Table of Contents

Food Distribution Research Society

Journal of Food Distribution Research
Volume 55, Issue 2, pp. 1–22

Assessing the Impacts of the Three Distinct Promotion Campaigns for Fluid Milk

Oral Capps, Jr.[a]ⓘ and Joshua Strine[b]

[a]*Executive Professor and Regents Professor, Department of Agricultural Economics, 600 John Kimbrough Blvd., Suite 371, 2124 Texas A&M University, College Station, TX 77840, USA*

[b]*Graduate Research Assistant, Department of Agricultural Economics, Krannert Building, 403 Mitch Daniels Blvd., Room # 651, Purdue University, West Lafayette, IN 47906, USA*

Abstract

Impacts of the "Got Milk?", "Milk Life," and "#GOTMILKCHALLENGE" campaigns concerning fluid milk consumption were analyzed using a time-varying parameter model over the period July 1995 to December 2022. The long-run promotion elasticity for fluid milk without consideration of individual campaigns was estimated to be 0.043. The individual advertising impacts were quite dynamic, changing within thematic periods, and these impacts were not uniform across themes. Unlike the "Milk Life" campaign, the "Got Milk?" and the "#GOTMILKCHALLENGE" campaigns were consistent with the hypothesis of advertising wearout. This work addressed the effectiveness of the overall generic message and the messages linked to the respective campaigns.

Keywords: demand for fluid milk, econometric analysis, advertising wearout, and advertising/promotion campaigns for fluid milk

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

Introduction

Created by the Fluid Milk Promotion Act of 1990, The Milk Processor Education Program (MilkPEP) uses advertisements to promote and inform U.S. consumers of the dietary benefits of fluid milk. Funded by a conglomerate of milk processors, the program strives to maintain the reputation of and to increase the demand for milk. The “Got Milk?” campaign commenced in 1993, two years prior to the “Milk Mustache” campaign. The original Got Milk? campaign was developed and executed by Goodby, Silverstein, and Partners in 1993. In 1995, MilkPEP licensed the “Got Milk?” tagline and incorporated it into its “Milk Mustache” campaign. Hence, beginning in 1995, MilkPEP created the first of three consecutive national fluid milk promotional campaigns (Daddona, 2018). The “Got Milk?”, “Milk Life,” and “#GOTMILKCHALLENGE” campaigns ran from January 1995 to February 2014, from March 2014 to July 2020, and from August 2020 to the present, respectively. These three campaigns differed not only in tagline but also in approach.

The original “Got Milk?” campaign utilized a variety of television commercials that regularly involved subjects requiring milk to complement food choices. Posters with glossy photos of celebrities, from actors to athletes, sporting milk mustaches, and milk facts also were a mainstay of the “Got Milk?” campaign. In 2014, MilkPEP retired the “Got Milk?” campaign and replaced it with the “Milk Life” campaign (Schultz, 2014). Gone were the celebrities, as younger people in action became the focal point of the promotional efforts. Milk mustaches vanished and were replaced by a flow of milk encompassing the subject. Humor took a backseat in television spots as the “Milk Life” campaign conveyed a more sentimental and informative approach. Commercials promoted the nutritional attributes of milk and questioned those of plant-based alternatives. The “Milk Life” campaign did not have the same longevity as its predecessor. It was retired in February/March 2020 after just six years.

After a brief intermission, MilkPEP resurrected the “Got Milk?” campaign in 2020 to capitalize on stay-at-home orders brought about by the pandemic (Durbin, 2020). The current campaign known as the “#GOTMILKCHALLENGE” campaign, supplants milk mustaches with user-generated video clips, exploiting the use of social media. Participants demonstrate various talents while holding a glass of milk. While all three campaigns employ unique approaches, the ubiquitous theme revolves around the importance of drinking milk.

Objectives

The overarching goal of the “Got Milk?”, “Milk Life,” and “#GOTMILKCHALLENGE” campaigns is to increase the demand for fluid milk. The objectives of this study are fourfold: (i) to identify and assess the factors associated with the per capita consumption of fluid milk; (ii) to analyze the impacts of each of the three previously mentioned promotion campaigns concerning per capita consumption of fluid milk; (iii) to determine whether the impacts of the respective promotion campaigns vary over time; and (iv) if so, to determine whether the impacts of the respective campaigns exhibit advertising wearout, defined as declining effectiveness associated with increasing exposure.

Model Development

While these campaigns are of primary interest, other factors are likely to influence milk demand as well. These factors must be accounted for in a quantitative analysis of market demand to accurately isolate (or minimize confounding) the impacts of the three advertising and promotion campaigns. The econometric model concerning the per capita consumption of fluid milk in this analysis considers as explanatory factors: (i) the retail price of fluid milk; (ii) the retail prices of substitute/complementary products, in particular, the prices of other non-alcoholic beverages (bottled water, fruit juices, and plant-based alternatives to milk), (iii) disposable personal income; (iv) inflation; (v) population; (vi) changes in demographics or population dynamics, specifically regarding proportions of the population of children 0 to 5, 6 to 11, and 12 to 17 years of age; (vii) the retail price of cheese; (viii) the retail price of breakfast cereal; (ix) the percent of food expenditures in the away-from-home market; (x) seasonality; (xi) advertising and promotion expenditures for fruit drinks; (xii) generic demand-enhancing expenditures for fluid milk; and (xiii) the pandemic. This specification is consistent with previous work by Kaiser (2010), Davis et al. (2011), Davis et al. (2012), and Capps and Brown (2023). Through this specification, we filter out the effects of other factors and directly quantify the net impacts associated with the "Got Milk?" "Milk Life," and "#GOTMILKCHALLENGE" campaigns pertaining to fluid milk consumption.

Retail prices of fluid milk products capture own-price effects of consumption. Holding all factors invariant, as retail prices of fluid milk change, consumption of fluid milk is expected to change in the opposite direction. As economic theory suggests, prices of competing or complementary products as well as disposable personal incomes of consumers also may affect the consumption of fluid milk.

Historically, children and fluid milk consumption have been positively linked (Stewart, Dong, and Carlson, 2013). To capture the influence of children, we consider proportions of the U.S. population of preschool children (0 to 5 years of age), of elementary school and middle school children (6 to 11 years of age), and of adolescents (children 12 to 17 years of age).

We also must account for away-from-home eating and drinking trends. Roughly half of the share of the consumer dollar for food and beverages currently is spent away from home (USDA-ERS, n.d.). Further, fluid milk consumption is expected to be negatively impacted by the lack of availability of fluid milk products in away-from-home establishments as well as by the expanding availability of alternatives to fluid milk (plant-based alternatives) in the at-home market.

Importantly, generic marketing and promotion activities by fluid milk processors, dairy producers, and qualified programs (QPs) are expected to increase the consumption of fluid milk, holding all other factors constant. The generic fluid milk marketing, advertising, and promotion activities include all media activities, such as television, print, radio, outdoor, and web advertising by fluid milk processors and dairy farmers as well as health and nutrition educational programs, public relations, school milk programs, food service programs, retail programs, trade service communications, and other miscellaneous activities. At issue is whether consumer interest associated with generic advertising for fluid milk can be sustained by adjustments in creative approach or thematic appeal.

Additionally, we explore whether the promotion elasticities associated with the three distinct advertising campaigns or themes are constant or vary over time. We hypothesize that the respective promotion elasticities are not constant but exhibit inverted-U shaped patterns over time. This hypothesis is consistent with previous work by Kinnucan, Chang, and Venkateswaran (1993) in analyzing the impacts of five different fluid milk advertising themes in the New York City market over the period 1971–1984. Various theories labeled as life cycle, learning-based, information processing, and elaboration were related to explain the wear-out phenomenon. We hypothesize that each of three distinct advertising campaigns eventually lose their effectiveness as consumers assimilate the thematic information and find further repetitions superfluous.

In this light, we develop a single equation structural model presented in equation (1) as follows:

$$CDFLUIDMILK_t/POP_t = \alpha + \beta^* Z_t + \gamma_t * GW_t + e_t, \tag{1}$$

where $CDFLUIDMILK_t/POP_t$ corresponds to per capita consumption of fluid milk, defined as commercial disappearance divided by population; Z_t denotes a vector of exogenous variables, GW_t is advertising goodwill. Note that the parameters to be estimated are α, β, and γ; e_t is a random error term. Because γ has a subscript t, we allow for this parameter to vary over time.

The respective exogenous variables considered are: (i) the retail price of fluid milk, adjusted for inflation, denoted as $RETAIL\_PRICE\_FLUIDMILK_t/CPI\_NONALCBEV\_SA_t$ (the consumer price index for nonalcoholic beverages, seasonally adjusted; (ii) the consumer price index for cheese, seasonally adjusted, divided by the consumer price index for all items, seasonally adjusted denoted as $CPI\_CHEESE\_SA_t/CPI\_ALLITEMS\_SA_t$; the consumer price index for breakfast cereal, seasonally adjusted, divided by the consumer price index for all items, seasonally adjusted denoted as $CPI\_BREAKFAST\_CEREAL\_SA_t/CPI\_ALLITEMS\_SA_t$; real per capita disposable personable income denoted as $RPCDPI_t$; the percentage of the U.S. population corresponding to children 0 to 5 years of age (preschool) denoted as $PERCENT\_CHILDREN\_0TO5_t$; the percentage of the U.S. population corresponding to children 6 to 11 years of age (preadolescents) denoted as $PERCENT\_CHILDREN\_6TO11_t$; the percentage of the U.S. population corresponding to children 12 to 17 years of age (adolescents) denoted as $PERCENT\_CHILDREN\_12TO17_t$; the percent of sales from away-from-home eating establishments denoted by $FAFH\_PERCENT_t$; and promotion expenditures associated with fruit juices and drinks denoted by $JUICES\_AD\_D11_t/CPI\_ALLITEMS\_SA_t$ We also control for seasonality and the pandemic through the use of dummy variables.

Advertising goodwill, GW_t is defined as

$$GW_t = \sum_{k=0}^{m} \delta_k * \ln AD_{t-k} \tag{2}$$

where AD_{t-k} pertains to advertising and promotion expenditures in period t-k, m is the length of the distributed lag process, and the δ_ks are the lag weights. Upon substitution of equation (2) into equation (1) we arrive at the model specification given as:

$$\ln \quad CDFLUIDMILK_t/POP_t = \alpha + \beta * Z_t + \gamma_t * \sum_{k=0}^{m} \delta_k * \ln AD_{t-k} + e_t, \qquad (3)$$

The generic advertising and promotion expenditures for fluid milk correspond to the combined efforts of Dairy Management, Inc. (DMI), MilkPEP, and Qualified Programs (QPs). The set of AD variables in equations (2) and (3) correspond to real generic demand-enhancing promotion expenditures for fluid milk, seasonally adjusted, denoted as $DMI\_MILKPEP\_QP\_A\_D11_t/CPI\_ALLITEMS\_SA_t$. To be consistent with economic theory, advertising must be subject to diminishing marginal returns (Simon and Arndt, 1980). As such, we adopt the logarithmic transformation in equations (2) and (3). The double log model is consistent with the diminishing marginal returns hypothesis. The weights of the goodwill variable δs are assumed to be time invariant. The contemporaneous impact (short-run elasticity) of advertising and promotion on the part of DMI, MilkPEP, and QPs across the campaigns is given by δ_0, while the cumulative impact (long-run elasticity) is given by $\sum_{k=0}^{m} \delta_k$.

Importantly, in vetting the impacts of marketing or generic promotional expenditures, carryover effects likely are evident. Previous studies support the hypothesis that demand-enhancing activities have carryover or lagged effects (e.g., Nerlove and Waugh, 1961; Williams, Capps, and Palma, 2008; Kaiser, 2010; Williams and Capps, 2019; Williams, Capps, and Dang, 2010; and Williams and Capps, 2020). However, economic theory provides relatively little guidance as to the structure and length of these dynamic processes. To capture the dynamics of these carryover effects, we use a polynomial distributed lag process (Almon, 1965) in the model specification. This approach is consistent with the quantitative evaluation of checkoff programs in general (Forker and Ward, 1993; Kaiser et al., 2005; Capps, Bessler, and Williams, 2016; Capps and Brown, 2023).

In equation (3), the lag weights used in the construction of the goodwill variable are estimated jointly with α and β. In this estimation, we rely on the Almon (1965) procedure, with head and tail constraints, and we assume the lag weights to follow a second-degree polynomial. Aside from the distribution of the lag weights, another key issue is the length of the lag structure associated with the respective real and seasonally adjusted promotion expenditures for fluid milk. We follow the conventional procedure of using statistical criteria like the Akaike Information Criterion (AIC), the Schwarz Loss Criterion (SLC), or the Hannan-Quinn Criterion (HQC) in allowing the data to suggest the optimal number of lags (m) to include in the specification. We account for the dynamics of promotion expenditures associated with fruit juices and drinks in precisely the same manner.

The coefficient associated with advertising goodwill is expressed as a time-varying parameter:

$$\gamma_t = f(T), \qquad (4)$$

where T corresponds to a time trend and *f* corresponds to the specified functional form. Based on equation (4), we can test the hypothesis concerning whether the impacts of advertising goodwill are constant or varying. Additionally, if we ascertain that the impacts are time sensitive, then we are in position to ascertain if the advertising wearout hypothesis holds.

Empirical Specification

The empirical version of equations (3) and (4) is specified as:

ln $CDFLUIDMILK_t/POP_t$, = a_0 + a_1 *ln ($RETAIL\_PRICE\_FLUIDMILK_t$ / $CPI\_NONALCBEV\_SA_t$) + a_2*ln $RPCDPI_t$ +a_3*ln ($CPI\_CHEESE\_SA_t$/ $CPI\_ALLITEMS\_SA_t$) + a_4 *ln ($CPI\_BREAKFAST\_CEREAL\_SA_t$,/ $CPI\_ALLITEMS\_SA_t$) + a_5 *ln($PERCENT\_CHILDREN\_0TO5)_t$, + a_6 * ln($PERCENT\_CHILDREN\_6TO11)_t$ + a_7 *ln($PERCENT\_CHILDREN\_12to17)_t$ + a_8 *ln $FAFH\_PERCENT_t$ + a_9 *$MILKLIFE_t$ + a_{10}*#$GOTMILKCHALLENGE_t$ + $\sum_{k=11}^{21} a_k$*@SEAS(k) + a_{22}*D2020m04 + a_{23}*D2020m05 + a_{24} *Pandemic_JuntoDec2020 + a_{25} *Pandemic_2021+ a_{26} *Pandemic_2022 + $\sum_{l=1}^{3} \gamma_{lt}$*$GW_t$*$THEME_l$ +v_t. (5)

In this specification, the subscript t represents monthly observations over the period January 1995 to December 2022. Consequently, the number of observations available for analysis is 336.

The dependent variable labeled as $CDFLUIDMILK_t/POP_t$, denotes the commercial disappearance of fluid milk per capita in the United States. Hence, we account for fluid milk consumption as well as population in the analysis. The commercial disappearance of fluid milk corresponds to estimated fluid milk product sales available from the Agricultural Marketing Service, USDA. These sales data measured in pounds are dispositions (deliveries) of fluid milk products in consumer type packages from milk processing (bottling) plants to outlets in Federal Order marketing areas. These outlets include food stores, convenience stores, warehouse stores/wholesale clubs, non-food stores, schools, the food service industry, and home delivery.

$RETAIL\_PRICE\_FLUIDMILK_t$ denotes the retail price of fluid milk; $CPI\_NONALCBEV\_SA_t$ denotes the consumer price index of nonalcoholic beverages. The retail price of whole milk in terms of dollars per gallon is a proxy for the price of fluid milk. By dividing by CPI_NONALCBEV_SAt, the real price of fluid milk indirectly considers the price of nonalcoholic beverages. RPCDPIt denotes real per capita disposable income, measured in 2017 dollars. CPI_CHEESE_SAt and CPI_BREAKFAST_CEREAL_SAt denote the seasonally adjusted consumer price index of cheese and related products and the seasonally adjusted consumer price index of breakfast cereals. CPI_ALLITEMS_SAt denotes the seasonally adjusted consumer price index for all items. The consumer price index of cheese and related products adjusted for inflation (CPI_CHEESE_SAt / CPI_ALLITEMS_SAt) reflects the substitution of cheese for fluid milk. The consumer price index of breakfast cereals adjusted for inflation (CPI_BREAKFAST_CEREAL_SAt,/ CPI_ALLITEMS_SAt) reflects the complementarity of breakfast cereals with fluid milk.

The explanatory variables labeled as PERCENT_CHILDREN_0TO5, PERCENT_CHILDREN_6TO11, and PERCENT_CHILDREN_12TO17 represent the percentage of the U.S. population that falls within the specified age brackets. These measures control for population dynamics among preschool children, elementary and middle school children, and adolescents. FAFH_PERCENT denotes the percent of sales from away-from-home eating establishments.

The explanatory variables $MILKLIFE_t$, and $\#GOTMILKCHALLENGE_t$ denote the "Milk Life" and "#GOTMILKCHALLENGE" campaigns, respectively. Both are dummy variables, and the reference or base category is the original "Got Milk?" campaign. The respective campaigns are mutually exclusive and exhaustive. The coefficients associated with these variables capture how much higher or lower, on average, per capita consumption of fluid milk is relative to the "Got Milk?" campaign. These coefficients do not capture the effects of the three distinct campaigns.

Dummy variables are included in the model specification to account for seasonality. The variables labeled @SEAS(k), k =11,12..., 21, represent the 11 months of each calendar year, respectively. The month of December is excluded to avoid the dummy variable trap and corresponds to the base or reference month.

The World Health Organization formally declared COVID-19 a pandemic on March 11, 2020. Two days later on March 13, 2020, the Trump Administration declared COVID-19 a national emergency. We adopt this period to indicate the start of market disruption attributed to COVID-19. That said, we acknowledge that initial consumer reaction to the pandemic could have happened before March 11, 2020, given that the first COVID-19 case in the United States could be traced back to January 21, 2020, and given that the CDC expressed a warning of a looming pandemic on February 25, 2020. In this analysis, the dummy variables D2020m04 (defined as 1 if April 2020 and 0 otherwise) and D2020m05 (defined as 1 if May 2020 and 0 otherwise) represent the months immediately following the pandemic. We also consider dummy variables associated with the pandemic for the remainder of 2020 (defined as 1 if June, July, August, September, October, November, and December 2020 and 0 otherwise) as well as consider the impacts of the pandemic for calendar years 2021 (defined as 1 for months in calendar year 2021 and 0 otherwise) and 2022 (defined as 1 for months in calendar year 2022 and 0 otherwise).

$THEME_l$ corresponds to dummy variables to indicate theme changes in advertising copy. $Theme_1$ corresponds to the "Got Milk?" campaign, and $Theme_1 = 1$ if $t \leq 230$, 0 otherwise; $Theme_2$ corresponds to the "Milk Life" campaign, and $Theme_2 = 1$ if $231 \leq t \leq 307$, 0 otherwise; and $Theme_3$ corresponds to the "#GOTMILKCHALLENGE" campaign, and $Theme_3 = 1$ if $308 \leq t \leq 336$, 0 otherwise; and v_t is a random error term.

Equation (5) allows goodwill elasticities to differ depending on the campaign theme. Like Kinnucan, Chang, and Venkateswaran (1993), advertising wearout is introduced into the model by specifying γ_t associated with each theme-specific campaign as:

$$\gamma_{lt} = \Omega_{0l} + \Omega_{1l}*T_l + \Omega_{2l}*T_l^2, \quad (6)$$

where l=1,2,3 denotes campaign themes and T_l are trend terms defined as follows:

T_1 = 1, 2,..,,230, and zero otherwise (for the "Got Milk?" Theme),

T_2 = 1,2,…,77., zero otherwise (for the "Milk Life" Theme), and

T_3 = 1,2,….29, zero otherwise (for the "#GOTMILKCHALLENGE" Theme).

Equation (6) is the empirical analogue of equation (4).[1]

Attributed to equation (5), the goodwill promotion elasticity associated with each campaign theme is calculated to be $\gamma_{lt}*GW_t*Theme_l$. To be consistent with advertising wearout, we expect Ω_{1l} to be positive and Ω_{2l} to be negative. If $\Omega_{1l} = \Omega_{2l} = 0$, then $\gamma_{jt} = \Omega_{0l}$, implying that the impact of each promotion campaign is not time sensitive.

Data

Because data pertaining to the retail price of whole milk were only first available in July 1995, the econometric analysis runs from July 1995 to December 2022. The sample size then for the econometric analysis is 330 observations. Promotion expenditures for fluid milk are not available after 2022.

Retail prices for whole milk, the consumer price index for nonalcoholic beverages (a proxy for alternatives to fluid milk), the consumer price index for breakfast cereals, the consumer price index for all items, and the consumer price index for cheese were obtained from the Bureau of Labor Statistics. Data for disposable personal income and population were available from the Federal Reserve Bank of St. Louis. Data pertaining to the proportion of children in various age groups as well as data concerning retail sales for food and beverages (at-home and away-from-home) were obtained from the U.S. Bureau of the Census. The source of the information on demand-enhancing expenditures for fluid milk was the U.S. Department of Agriculture's Agricultural Marketing Service. Finally, information on advertising and promotion expenditures associated with fruit juices and drinks was procured from Competitive Advertising Intelligence, Ad Intel.

Descriptive statistics of the econometric analysis are exhibited in Table 1. Per capita quarterly consumption of fluid milk ranged from 9.89 pounds[2] to 18.20 pounds, averaging 14.37 pounds over the period January 1995 to December 2022. From Figure 1, it is clear that per capita fluid milk consumption not only has been on a steady decline over the past 28 years, but also exhibits a seasonal pattern. The downward trend likely reflects changes in the frequency of fluid milk intake

[1]Reberte et al. (1996) examined two major generic fluid milk advertising campaigns in New York City over the period 1986 to 1992. Estimates from a time-varying parameter model were consistent with a bell-shaped pattern. In that study, $\gamma_{lt} = \exp(\Omega_{0l} + \Omega_{1l}*T_l + \Omega_{2l}*T_l^2)$.

[2]A gallon of milk is equivalent to 8.6 lbs.

rather than changes in portions (Stewart, Dong, and Carlson, 2013). Most Americans born in the 1990s tend to consume fluid milk less often than those born in the 1970s, who in turn consume fluid milk less often than those born in the 1950s. U.S. per capita milk consumption has declined roughly 36% since 1995, largely due to changing consumption habits as well as increased competition from other beverages. Moreover, according to Stewart et al. (2021), U.S. consumers of all ages are drinking less milk and milk drinks.

Table 1. Descriptive Statistics of the Continuous Variables in the Econometric Analysis, July 1995 to December 2022

Variable Name	Mean	Variable Name	Mean
Disappearance of fluid milk per capita (pounds)			
CDFLUIDMILK/POP	14.32		
Advertising/promotion campaigns			
GOT_MILK? (*Reference/Base Category)*	0.6788		
MILK_LIFE	0.2333		
#GOT_MILK_CHALLENGE	0.0879		
Nominal retail price of milk ($/gallon)		Nominal seasonally adjusted advertising/promotion expenditures fluid milk (dollars)	
RETAIL_PRICE_FLUIDMILK	$3.17	DMI_MILKPEP_QP_A_D11	$32,173,048
Real per capita disposable personal income (2017 dollars)		Nominal seasonally adjusted advertising/promotion expenditures fruit juices (1,000 dollars)	
RPCDPI	$40,313	JUICES_AD_D11	$137,930
Consumer price indices (1982-84=100)		Population dynamics (proportion of the U.S. population)	
CPI_NON_ALCOHOLIC_BEV	156.0422	PERCENT_CHILDREN_0TO5	7.8425
CPI_ALLITEMS_SA	212.8428	PERCENT_CHILDREN_6TO11	8.0566
		PERCENT_CHILDREN_12TO17	8.2488
		Food away from home expenditures (% of the dollar spent on food away from home)	
		FAFH_PERCENT	44.2801

Source: Calculations made by the authors using the EViews 11.0 (2020) econometrics software package.

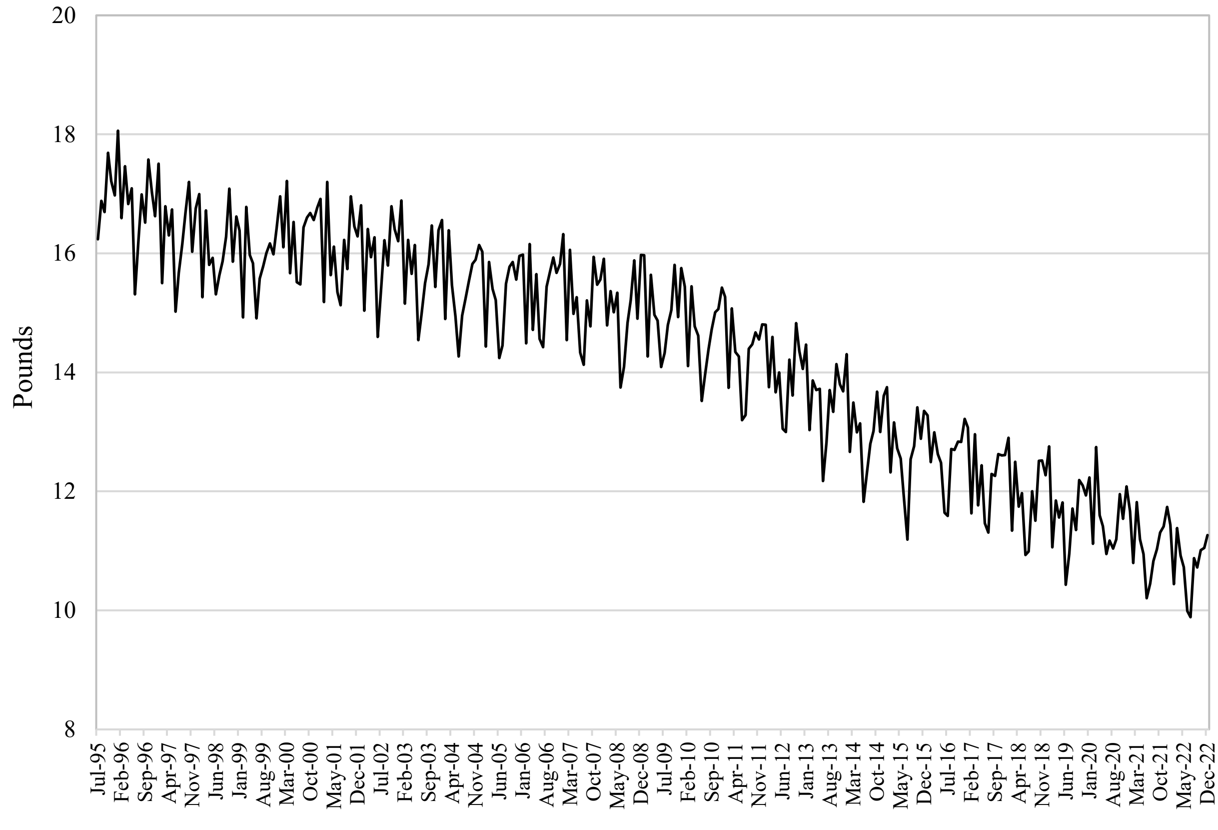


Figure 1. Per Capita Fluid Milk Consumption in Pounds, July 1995 to December 2022

Source: United States Department of Agriculture

The retail price of whole milk is used to measure own price on a dollar per gallon basis. Holding all else constant, fluid milk consumption is expected to be inversely related to price in accord with economic theory. The nominal retail price of whole milk ranged from \$2.46/gallon to \$4.22/gallon over the sample period, \$3.17/gallon on average.

We use the ratio of the retail price of fluid milk to the consumer price index for nonalcoholic beverages in the model specification. This price ratio then accounts not only for inflation, but also for prices of alternative beverages to milk. Consequently, interest lies with the impact of the retail price of whole milk relative to the price of nonalcoholic beverages.

Real per capita disposable personal income serves to account for income, population, and inflation. Holding all other factors constant, fluid milk is expected to be a normal good, and as such we hypothesize that fluid milk consumption is positively related to income. Over the sample period, real per capita disposable personal income measured in 2017 dollars varied from \$30,686 to \$62,509, averaging \$40,314.

Seasonally adjusted consumer price indices of cheese, nonalcoholic beverages, breakfast cereals, and all items serve to isolate the effects of other prices and inflation. Over the sample period, the share of the U.S. population of children 0 to 5 years of age averaged 7.84%; the share of the U.S. population of children 6 to 11 years of age averaged 8.06%; and the share of the U.S. population of children 12 to 17 years of age averaged 8.25%. Because these measures of population dynamics were only available annually, interpolations were done to place these figures on a monthly basis.

Sales from food service and drinking establishments as a percent of the sum of spending at food and beverage stores and food-service and drinking establishments are used as a measure of food-away-from home spending. Since 1995 food-away-from-home expenditures have risen consistently, climbing from roughly 30% to 52% over the sample period, averaging slightly more than 44%. Food-away-from-home expenditures plummeted from 51% to 37% in March 2020, 31% in April 2020, and 36% in May, respectively, due to the COVID-19 pandemic and stay-at-home orders. Since June 2020, sales from food service and drinking establishments as a percent of the sum of spending at food and beverage stores and food-service and drinking establishments have risen monotonically. Fluid milk often is not consumed or on the menu at food-service or drinking places. Hence, milk consumption is expected to decrease with increases in the shares of food-away-from-home expenditures.

The COVID-19 pandemic accounts for using several dummy variables. The first set of dummy variables corresponds to April 2020 only and May 2020 only, designed to capture the impact of the initial onset of the pandemic. A second dummy variable represents the pandemic from June 2020 to December 2020. In this way, we ascertain the impacts of the pandemic, initially and subsequently, in 2020. The final set of dummy variables corresponds to calendar year 2021 and calendar year 2022. We hypothesized that the COVID-19 pandemic negatively affects per capita consumption of fluid milk. The base or the reference period is the pre-pandemic period.

On average, nominal seasonally adjusted advertising and promotion expenditures for fluid milk ranged from \$12.27 million to \$62.75 million, averaging \$32.17 million over the period July 1995 to December 2022. On average, nominal seasonally adjusted advertising and promotion expenditures for fruit juices and drinks ranged from \$52.09 million to \$310.73 million, averaging \$137.93 million over the sample period. The advertising/promotion demand-enhancing expenditure variables were seasonally adjusted using the X13 procedure developed by the Census Bureau.

To measure the impact of three previously mentioned advertising/promotion campaigns, we created three dummy variables. The “Got Milk?” campaign corresponds to the period July 1995 to February 2014. The “Milk Life” campaign corresponds to the period March 2014 to July 2020. The “#GOTMILKCHALLENGE” campaign corresponds to the period August 2020 to December 2022. Nominal advertising expenditures for fluid milk from Dairy Management, Inc. (DMI), MilkPEP, and Qualified Programs (QPs) amounted to \$35.71 million per quarter on average for the “Got Milk?” campaign; \$26.32 million per quarter on average for the “Milk Life” campaign; and \$20.81 million per quarter on average for the “#GOTMILKCHALLENGE” campaign. Consequently, the amount of advertising and promotion expenditures was not constant across the respective campaigns.

Empirical Results

Because of the term $\sum_{l=1}^{3} \gamma_{lt}$*$GW_t$*$THEME_l$, equation (5) corresponds to a nonlinear model. Consequently, the method of estimation is nonlinear least squares. In the search for the optimal lag lengths, second- and third-degree polynomials with lags up to 12 months were considered along with alternative choices of head and tail (endpoint) restrictions for GW_t as well as for promotion expenditures associated with fruit juices and drinks. Based on the AIC, SIC, and HQC, a second-order polynomial distributed lag specification was identified as a lag length of three months for real and seasonally adjusted promotion expenditures of fruit juices and drinks and 12 months for real and seasonally adjusted fluid milk promotion expenditures. To arrive at this empirical specification, a plethora of different combinations of lag structures were considered. For estimation purposes, we adopted the use of logarithmic transformation for all continuous variables in the model.

To mitigate irreconcilable degrading collinearity issues**,** we restricted Ω_{0l} to be 0 for l = 1,2,3, and we dropped the consumer price index for breakfast cereal, the consumer price index for cheese, the percent of the population associated with children 6 to 11 years of age, and the percent of the population associated with children 12 to 17 years of age from the model.[3] Additionally, based on R-student statistics and hat diagonal elements (Belsley, Kuh, and Welsch, 1980), one observation, namely July 2015, was deemed to be an influential data point (outlier and leverage point). To mitigate this issue, we created a dummy variable associated with this observation (1 for July 2015, and 0 otherwise).

The parameter estimates, standard errors, and *p*-values for the explanatory variables of the econometric model obtained from the use of the software package EViews 11.0 (EViews, 2020) are exhibited in Table 2. The R^2 metric was 0.9863 and the adjusted R^2 metric was 0.9850. The standard error of of variability in the per capita consumption of fluid milk, with a negligible variability in the regression.

Table 2. Parameter Estimates, Standard Errors, and *p*-values for the Explanatory Variables of the Econometric Model for Fluid Milk (Dependent Variable: LOG(PERCAPITA_FLUIDMILK)

Variable	Coefficient	Std. Error	*t*-statistic	*p*-value
C	2.3150	0.7629	3.03	0.0026
LOG(RETAIL_PRICE_FLUID_MILK*100/CPI_NON_ALCOHOLIC_BEV)	-0.0700	0.0175	-4.00	0.0001
LOG(REAL_PERCAPITA_DPI)	0.0532	0.0618	0.86	0.3896
LOG(PERCENT_CHILDREN_0TO5)	0.4706	0.1493	3.15	0.0018
@SEAS(1)	0.0121	0.0047	2.57	0.0107
@SEAS(2)	-0.0887	0.0047	-18.86	0.0000
@SEAS(3)	-0.0047	0.0047	-0.99	0.3245

[3]These collinearity issues were revealed based on examination of variance inflation factors, condition indices, and variance decomposition proportions (Belsley, Kuh, and Welsch, 1980).

Table 2 (cont.)

Variable	Coefficient	Std. Error	*t*-statistic	*p*-value
@SEAS(4)	-0.0540	0.0048	-11.32	0.0000
@SEAS(5)	-0.0417	0.0048	-8.77	0.0000
@SEAS(6)	-0.1187	0.0048	-25.20	0.0000
@SEAS(7)	-0.0971	0.0047	-20.65	0.0000
@SEAS(8)	-0.0411	0.0047	-8.84	0.0000
@SEAS(9)	-0.0388	0.0047	-8.34	0.0000
@SEAS(10)	0.0047	0.0046	1.01	0.3157
@SEAS(11)	-0.0192	0.0046	-4.13	0.0000
LOG(FAFH_PERCENT)	-0.2279	0.0455	-5.01	0.0000
D2020M04	-0.0895	0.0281	-3.19	0.0016
D2020M05	-0.0747	0.0223	-3.36	0.0009
GW*GOT_MILK_TREND	0.0012	0.0005	2.29	0.0230
GW*MILK_LIFE_TREND	-0.0026	0.0009	-3.00	0.0029
GW*GOT_MILK_CHALLENGE_TREND	0.0094	0.0040	2.35	0.0192
GW*GOT_MILK_TSQ	-9.25E-06	1.23E-06	-7.51	0.0000
GW*MILK_LIFE_TSQ	-4.53E-06	1.06E-05	-0.43	0.6698
GW*GOT_MILK_CHALLENGE_TSQ	-0.0004	0.0001	-2.99	0.0031
D2015M07	-0.0748	0.0176	-4.26	0.0000
GOT_MILK_CHALLENGE	-0.2650	0.0459	-5.77	0.0000
MILK_LIFE	-0.1136	0.0353	-3.21	0.0015
R-squared	0.9863			
Adjusted R-squared	0.9850			
Standard error. of regression	0.0170			
F-statistic	770.84	Durbin-Watson statistic 2.08		
p-value (F-statistic)	0.0000			

Lag Distribution of LOG(JUICE_AD_EXPENDITURES_SA/CPI_ALLITEMS_SA)	i	**Coefficient**	**Std. Error**	***t*-statistic**	***p*-value**
	0	-0.0054	0.0017	-3.26	0.0012
	1	-0.0081	0.0020	-3.26	0.0012
	2	-0.0081	0.0020	-3.26	0.0012
	3	-0.0054	0.0017	-3.26	0.0012
	Sum of lags	-0.0270	0.0083	-3.26	0.0012

Within sample, the mean absolute error (MAE) was 0.18 pounds, and the mean absolute percent error (MAPE) was 1.30%. These measures corroborate the exceptional goodness-of-fit statistics. Based on the Durbin-Watson statistic, no autocorrelation in the residuals was evident.

Importantly, at the 0.05 level of significance, all estimated coefficients associated with the explanatory variables were statistically significant except for real disposable personal income and the interaction of GW with the square of the "Milk Life" trend term. With the use of logarithmic transformations, the estimated coefficients associated with all retail price of fluid milk, real per capita disposable income, percent of the population associated with children 0 to 5, and percent of food-away-from-home expenditures are elasticities.

The own-price elasticity for fluid milk was estimated to be -0.07, meaning that for every 10% change in the price of fluid milk relative to the price of nonalcoholic beverages, per capita fluid milk consumption changes by 0.70% in the opposite direction. The demand for fluid milk then is inelastic, that is, relatively unresponsive to price changes. This result is consistent with economic theory and with the extant literature (Kaiser, 2010; Dong and Stewart, 2013).

The percentage of the population associated with children from 0 to 5 years of age was a key determinant affecting per capita fluid milk consumption. A 1% rise in the proportion of children under 5 years of age resulted in a 0.47% increase in per capita fluid milk consumption. Clearly, econometric evidence exists to demonstrate that very young children are important drivers of fluid milk consumption. As this segment of the U.S. population declines, per capita fluid milk consumption will follow suit, all other factors being invariant.

The elasticity with respect to the percent of food-away-from-home expenditures was estimated to be -0.23. For every 1% rise in this percentage, per capita fluid milk consumption would fall by 0.23%, *ceteris paribus*. As mentioned previously, real per capita disposable income was not a statistically significant factor associated with per capita fluid milk consumption.

Because the model specification involves the logarithmic transformation of the per capita fluid milk consumption, we invoked the use of the Halvorsen and Palmquist (1980) convention to interpret all estimated coefficients associated with dummy variables.[4] Regarding seasonality, per capita fluid milk consumption was highest in January by 1.21% relative to December. On the other hand, per capita consumption of fluid milk was lower in all remaining months relative to December. In particular, per capita consumption of fluid milk was lower by 8.49% in February, 5.26% in April, 4.09% in May, 11.19% in June, 9.25% in July, 4.03% in August, and 3.80% in September relative to December.

Per capita consumption of fluid milk was lower by 8.56% in April 2020 and by 7.20% in May 2020, immediately following the onset of the pandemic. In subsequent months of 2020, calendar months of 2021, and calendar months of 2002, no statistically significant differences in per capita consumption of fluid milk were evident relative to the pre-pandemic period. Consequently, these explanatory variables were dropped from the econometric analysis.

The impacts of advertising for fruit juices were negative on per capita consumption of fluid milk, as expected. The short-run elasticity of advertising for fruit juices and drinks was estimated to be

[4]With this convention, the percentage change associated with any included dummy variable with respect to its base or reference category is given as (exp(the estimated coefficient) minus 1)*100.

-0.0054, whereas the cumulative or long-run elasticity was estimated to be -0.0270. The optimal cumulative effects of advertising on fruit juices and drinks were over a period of three months.

The estimated coefficients of lag distribution of weights associated with the GW variable exhibited in Table 3 support the hypothesis that the efforts of MilkPEP, DMI, and the QPs to enhance the demand for fluid milk were successful across campaigns. Based on these estimated coefficients, the impacts of the check-off expenditures from milk processors, dairy producers, and the QPs indeed boosted per capita consumption of fluid milk, holding all other factors constant. The optimal cumulative effect of these demand-enhancing promotion activities associated with the GW variable occurred over a period of 12 months. This distribution corresponds to a polynomial distributed lag process of degree 2 with endpoint constraints (both head and tail constraints).[5] The cumulative or long-run elasticity for fluid milk with respect to marketing, advertising, and promotion activities on the part of MilkPEP, DMI, and QPs across campaigns was estimated to be 0.043. Our estimate of the magnitude of the impact of advertising and promotion for fluid milk is in accord with previous studies. Schmit and Kaiser (2004) estimated the average promotion elasticity of fluid milk to be 0.040 over the period 1975 to 2001, using national quarterly data. Kaiser (2010) estimated the advertising and promotion elasticity for fluid milk for the United States to be 0.037 over the period 1997 and 2009.

Table 3. Parameter Estimates, Standard Errors, and *p*-values for the Promotion Expenditures Associated with Fluid Milk in the Econometric Model

Lag Distribution of LOG(PROM_EXPENDITURES_D11* 100/CPI_ALLITEMS_SA)	**i**	**Coefficient**	**Std. Error**	***t*-statistic**	***p*-value**
	0	0.0012	0.0006	2.00	0.0467
	1	0.0023	0.0011	2.00	0.0467
	2	0.0031	0.0016	2.00	0.0467
	3	0.0038	0.0019	2.00	0.0467
	4	0.0043	0.0021	2.00	0.0467
	5	0.0045	0.0023	2.00	0.0467
	6	0.0046	0.0023	2.00	0.0467
	7	0.0045	0.0023	2.00	0.0467
	8	0.0043	0.0021	2.00	0.0467
	9	0.0038	0.0019	2.00	0.0467
	10	0.0031	0.0016	2.00	0.0467
	11	0.0023	0.0011	2.00	0.0467
	12	0.0012	0.0006	2.00	0.0467
	Sum of lags	0.0429	0.0215	2.00	0.0467

Source: Calculations by the authors using EViews 11.0.

[5]Because of the lag distribution associated with GW, we lose 12 observations in the estimation of the model. Hence the period for this analysis runs from July 1996 to December 2022.

However, despite the positive and statistically significant impact of the generic advertising and promotion expenditures for fluid milk relative to the original "Got Milk?" campaign, per capita consumption of fluid milk was lower by 10.74% for the "Milk Life" campaign. As well, relative to the original "Got Milk?" campaign, per capita consumption of fluid milk was lower by 23.28% for the "#GOTMILKCHALLENGE" campaign.

We reject the hypothesis that $\Omega_{1l} = \Omega_{2l} = 0$, implying that the impact of each promotion campaign is time invariant. The "Got Milk?" campaign and the "#GOTMILKCHALLENGE" campaign were consistent with the hypothesis of advertising wearout because Ω_{1l} was estimated to be positive and Ω_{2l} was estimated to be negative. However, the "Milk Life" campaign was not consistent with the wearout hypothesis. Indeed, the time-varying parameter associated with the "Milk Life" campaign declined monotonically during this campaign.

The goodwill promotion elasticity associated with each campaign theme is calculated as γ_{lt}*GW_t*$Theme_l$. The magnitudes of the goodwill promotion elasticities for each of the three campaigns are exhibited in Figures 2, 3, and 4. For the "Got Milk?" campaign, the goodwill promotion elasticity was estimated to be 0.0054 at the beginning of the campaign, peaking at 0.0116 50 months later, and then declining to -0.0531 at the end of the campaign. The goodwill promotion elasticity associated with the "Got Milk?" campaign turned negative after 114 months.

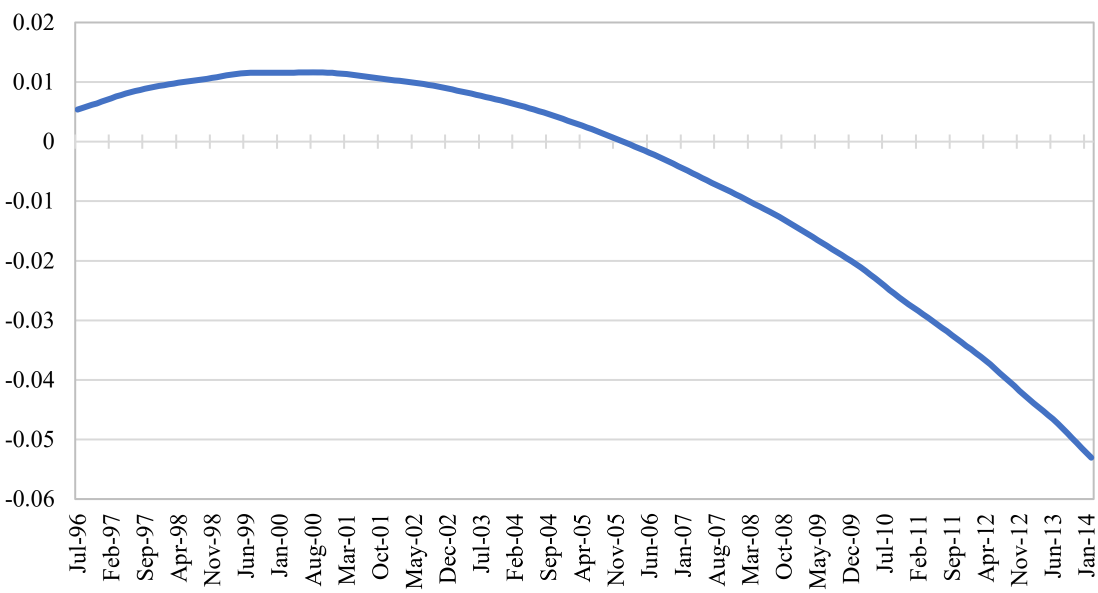


Figure 2. The Goodwill Elasticities Associated with the "Got Milk?" Campaign, July 1996 to February 2014

Source: Calculations by the authors.

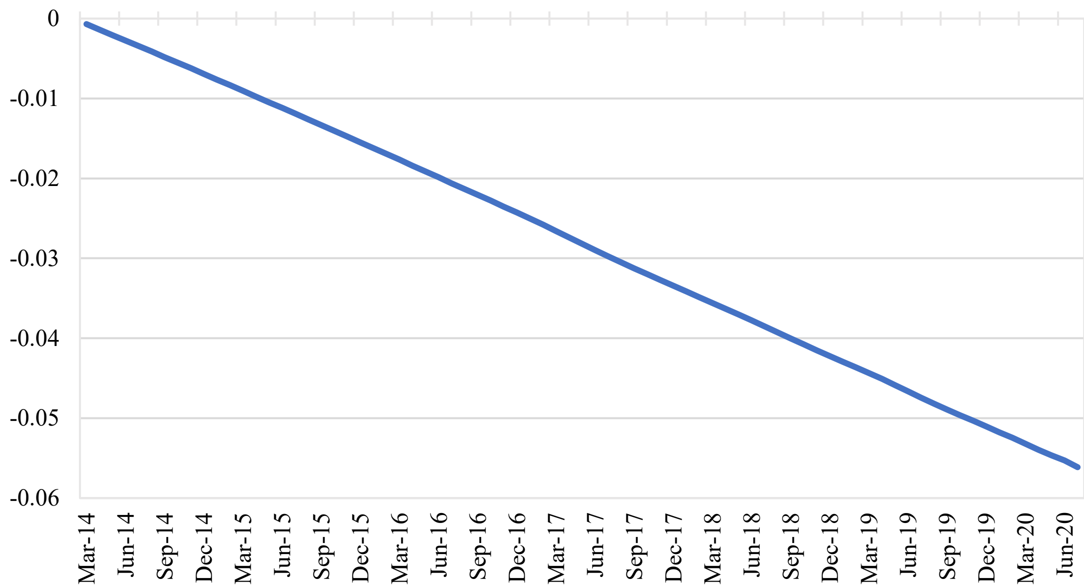


Figure 3. The Goodwill Elasticities Associated with the "Milk Life" Campaign, March 2014 to July 2020

Source: Calculations by the authors.

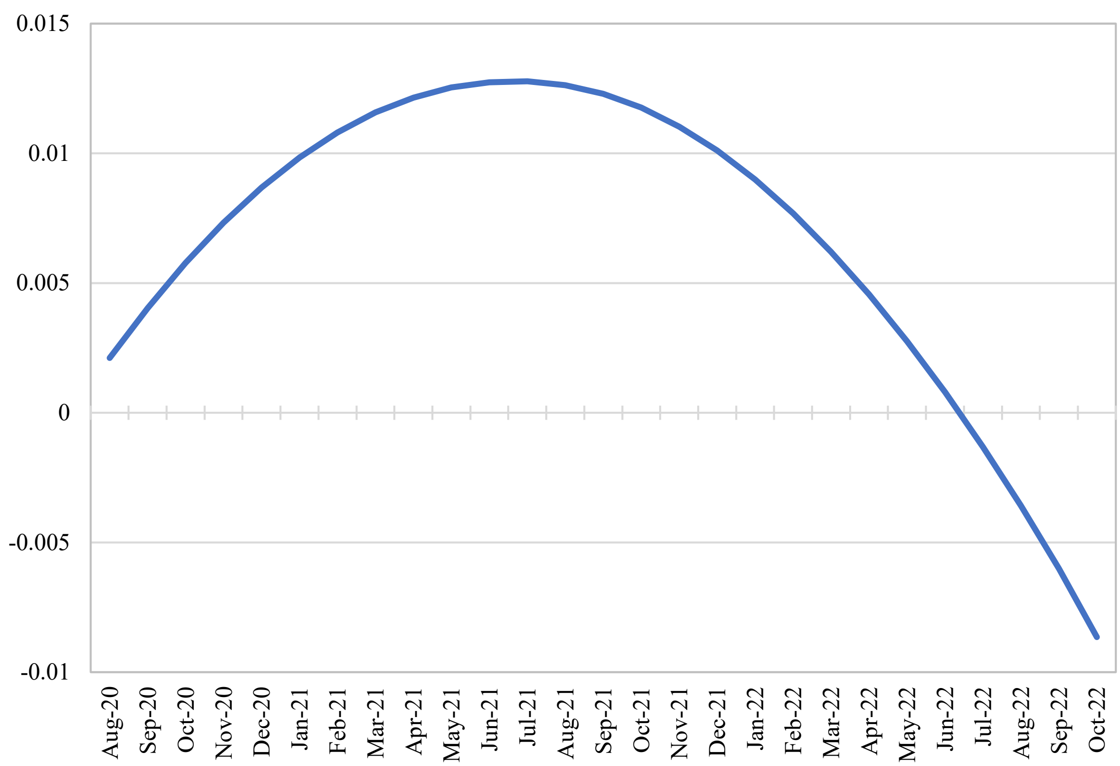


Figure 4. The Goodwill Elasticities Associated with the "#GOTMILKCHALLENGE" Campaign, August 2020 to December 2022

Source: Calculations by the authors.

For the “Milk Life” campaign, the goodwill promotion elasticity was estimated to be -0.0007 at the beginning of the campaign. This impact was also the peak of this campaign. Subsequently, the goodwill promotion elasticity associated with the “Milk Life” campaign declined to -0.0560 at the end of this campaign. For the “#GOTMILKCHALLENGE” campaign, the goodwill promotion elasticity was estimated to be 0.0021 at the beginning of the campaign, peaking at 0.0128 12 months later, then declining to -0.0086 at the end of the campaign. The goodwill promotion elasticity associated with the “#GOTMILKCHALLENGE” campaign turned negative after 23 months.

Without question, advertising impacts are quite dynamic, changing within thematic periods. Additionally, the advertising impacts are not uniform across themes. The peak impacts for the “#GOTMILKCHALLENGE” campaign and for the “Got Milk?” campaign were estimated to be 0.0128 and 0.0116, respectively. Both campaigns were instrumental in positively affecting per capita consumption of fluid milk up to a point in time. On the other hand, the “Milk Life” campaign negatively affected per capita consumption of fluid milk throughout.

Concluding Remarks

The impacts of the “Got Milk?”, “Milk Life,” and “#GOTMILKCHALLENGE” campaigns on per capita fluid milk consumption were analyzed using econometric analysis over the period July 1995 to December 2022. Accounting for a myriad of statistically significant factors, relative to the original “Got Milk?” campaign, per capita consumption of fluid milk was lower by 10.74% for the “Milk Life” campaign and lower by 23.28% for the “#GOTMILKCHALLENGE” campaign.

The long-run elasticity for fluid milk with respect to marketing, advertising, and promotion activities on the part of MilkPEP, DMI, and QPs without consideration of individual campaigns was estimated to be 0.043. This finding implies that the downward trend in per capita fluid milk consumption would have been exacerbated but for the advertising/promotion expenditures made by DMI, MilkPEP, and QPs. This finding also suggests that consumer interest in the generic message of drinking more milk can be maintained even with varying themes.

However, differences in advertising impacts were evident across themes. We reject the hypothesis that the impact of each promotion campaign was time invariant. The “Got Milk?” campaign and the “#GOTMILKCHALLENGE” campaign were consistent with the hypothesis of advertising wear out. Once consumers were familiar with the gist of the respective themes, repeated exposures were eventually tuned out. On the other hand, the “Milk Life” campaign was not consistent with this hypothesis. Indeed, the time-varying parameter associated with the “Milk Life” campaign declined monotonically throughout this campaign.

The respective advertising impacts were quite dynamic, changing within thematic periods. Additionally, the advertising impacts were not uniform across themes. The “Got Milk?” and the “#GOTMILKCHALLENGE” campaigns were instrumental in positively affecting per capita consumption of fluid milk up to a point in time. On the other hand, the “Milk Life” campaign negatively affected per capita consumption of fluid milk throughout.

Going forward, time-varying parameter models in assessing effectiveness of advertising campaigns should be implemented. The models should evaluate not only the effectiveness of the overall generic message (drink more milk in this analysis) but also the effectiveness of the messages linked to the respective campaigns.

As climate change and environmental concerns continue to grow, consumers are moving toward decreased consumption of animal products. Further, concerns over animal welfare and the safety of the milk supply (e.g., the issue of recombinant bovine somatotropin [rBST]) also could be responsible for changes in milk consumption. For future work, to minimize any confounding of impacts of various factors, it may be worthwhile to consider not only the environmental effects associated with dairy cows and the related greenhouse gases from their manure, but also concerns over animal welfare as possible determinants of the decline in per capita consumption of fluid milk.

To further study the impacts of the respective promotion campaigns for fluid milk, neuroeconomics can be utilized. Neuroeconomics is a relatively new discipline that merges concepts from economics, psychology, and neuroscience. Neuroeconomics uses a wide range of neurophysiological measures to study the connection between the nervous system, the body, and decision making (Palma, 2021). Neurophysiological equipment, including eye-tracking and facial expression analysis, can assess emotions to analyze the effectiveness of the three promotional campaigns for fluid milk in generating visual attention, recall, and propensity to purchase fluid milk. With the use of neuroeconomics, we would be able to compare and to contrast the findings gleaned from the use of econometric analysis.

Competing Interest Statement

Both authors have contributed equally to this article.

Acknowledgment

Funding for this research was provided by the Southwest/Southland Dairy Farmers. We wish to acknowledge the efforts of Jim Hill, CEO of the Southwest/Southland Dairy Farmers, in this regard.

Data Availability Statement

All data were obtained from secondary sources cited in the manuscript and are also available from the authors upon request.

References

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

Capps, Jr., O., D.A. Bessler, and G.W. Williams. 2016. "The Ramifications of Nearly Going Dark: A Natural Experiment in the Case of U.S. Generic Orange Juice Advertising." *Agricultural and Resource Economics Review* 45(1):68–97.

Capps, Jr., O., and D.S. Brown. 2023. *Quantitative Evaluation of the Effectiveness of Marketing and Promotion Activities by the National Dairy Promotion and Research Program and the Fluid Milk Processor Promotion Program—1995 to 2021*. Prepared for the U.S. Department of Agriculture, Agricultural Marking Service (AMS) Report to Congress, August.

Competitive Advertising Intelligence. *Ad Intel*. Available online: https://www.nielsen.com/solutions/media-planning/ad-intelligence/.

Daddona, M. 2018. "Got Milk? How the Iconic Campaign Came To Be, 25 Years Ago." *Fast Company*. Available online: https://www.fastcompany.com/40556502/got-milk-how-the-iconic-campaign-came-to-be-25-years-ago.

Davis, C.G., S.T. Yen, D. Dong, and D.P. Blayney. 2011. "Assessing Economic and Demographic Factors that Influence United States Dairy Demand." *Journal of Dairy Science* 94:3715–3723.

Davis, C.G., D. Dong, D.P. Blayney, S.T. Yen, and R. Stillman. 2012. "U.S. Fluid Milk Demand: A Disaggregated Approach." *International Food and Agribusiness Management Review* 15:25–50.

Dong, D., and H. Stewart. 2013. *Households' Choices among Fluid Milk Products: What Happens When Income and Prices Change?* Washington, DC: U.S. Department of Agriculture, Economic Research Report Number 146.

Durbin, D. 2020. "'Got Milk?' Ads Make a Comeback." *Morning Ag Clips*. Available online: https://www.morningagclips.com/got-milk-ads-make-a-comeback/.

EViews. 2020. *EViews 11*. Irvine, CA: IHS Global Inc.

Forker, O.D., and R.W. Ward. 1993. *Commodity Advertising: The Economics and Measurement of Generic Programs*. New York, NY: Lexington Books.

Halvorsen, R., and R. Palmquist. 1980. "The Interpretation of Dummy Variables in Semi-logarithmic Equations." *American Economic Review* 70:474–475.

Kaiser, H.M., J.M. Alston, J.M. Crespi, and R.J. Sexton. 2005. *The Economics of Commodity Promotion Programs*. New York, NY: Peter Lang Publishing.

Kaiser, H.M. 2010. *Measuring the Impacts of Generic Fluid Milk and Dairy Advertising.* Ithaca, NY: National Institute for Commodity Promotion Research and Evaluation, Research Bulletin 2010-01.

Kinnucan, H.W., H. Chang, and M. Venkateswaran, 1993. "Generic Advertising Wearout." *Review of Marketing and Agricultural Economics* 61(3):1–15.

Nerlove, M., and F.V. Waugh. 1961. "Advertising without Supply Control: Some Implications of a Study of the Advertising of Oranges." *Journal of Farm Economics* 43:813–837.

Palma, M.A. 2021. *Neurophysiological Measures and Consumer Food Choice*. College Station, TX: Texas A&M University, Department of Agricultural Economics.

Reberte, C., H.M. Kaiser, J.E. Lenz, and O. Forker. 1996. "Generic Advertising Wearout: The Case of New York City Fluid Milk Demand." *Journal of Agricultural and Resource Economics* 21(2):199–209.

Schultz, E.J. 2014. "'Got Milk' Dropped as National Milk Industry Changes Tactics." *AdAge*. Available online: https://adage.com/article/news/milk-dropped-national-milk-industry-tactics/291819.

Stewart, H., D. Doug, and A. Carlson. 2013. *Why Are Americans Consuming Less Fluid Milk? A Look at Generational Differences in Intake Frequency*. Washington, DC: U.S. Department of Agriculture, Economic Research Service, Economic Research Report Number 149.

Stewart, H., F. Kuchler, D. Dong, and J. Cessna. 2021. *Examining the Decline in U.S. Per Capita Consumption of Fluid Cow's Milk, 2003-18*. Washington, DC: U.S. Department of Agriculture, Economic Research Service, Economic Research Report Number 300.

U.S. Department of Agriculture. n.d. *Food Expenditure Series*. Washington, DC: U.S. Department of Agriculture, Economic Research Service. Available online: https://www.ers.usda.gov/data-products/food-expenditure-series/.

U.S. Department of Agriculture. 2010. *Fluid Milk Promotion Act of 1990*, as amended through May 7, 2010, United States Code, Title 7—Agriculture, Chapter 93—Processor Funded Milk Promotion Program, Chapter 101—Agricultural Promotion. Washington, DC: U.S. Department of Agriculture, Agricultural Marketing Service.

U.S. Department of Labor, Bureau of Labor Statistics, Available online at: www.bls.gov.

U.S. Federal Reserve Bank. *Economic Data (FRED)*. St. Louis, MO: U.S. Federal Reserve Bank. Available online: https://fred.stlouisfed.org.

Williams, G.W., O. Capps, Jr., and M.A. Palma. 2008. "Effectiveness of Marketing Promotion Programs: The Case of Texas Citrus." *HortScience* 43(2):385–392.

Williams, G.W., O. Capps, Jr., and T. Dang. 2010. "Does Lamb Promotion Work?" *Agribusiness: An International Journal* 26(4):536–556.

Williams, G.W., and O. Capps, Jr. 2019. "Generic Promotion of Sorghum for Food and Industrial Uses." *Journal of International Food and Agribusiness Marketing* 31(4):300–320.

Williams, G.W., and O. Capps, Jr. 2020. "Generic Promotion of Norwegian Whitefish Exports." *Journal of International Food and Agribusiness Marketing* 33(2):123–151.

Journal of Food Distribution Research
Volume 55, Issue 2, pp. 23–45

Consumption of Fruits and Vegetables and the Role of the Food Environment

Shannon Giscombe[a]ⓘ, Elizabeth Canales[b], Will Davis[c], Ayoung Kim[d] , and David Buys[e]

[a]*Agribusiness Research and Extension Associate, Sherman Lewis School of Agriculture and Applied Sciences, 100 Success Ave., P.O. Box 1730, Langston, University, Langston, OK 73050, USA*

[b]*Associate Professor, Department of Agricultural Economics, 255 Tracey Dr., P.O. Box 5187, Mississippi State University, Mississippi State, MS 39762*

[c]*Assistant Professor, Department of Agricultural Economics, 255 Tracey Dr., P.O. Box 5187, Mississippi State University, Mississippi State, MS 39762, USA*

[d]*Assistant Professor, Department of Agricultural Economics, 255 Tracey Dr., P.O. Box 5187, Mississippi State University, Mississippi State, MS 39762, USA*

[e]*Associate Professor and State Health Specialist, Department of Food Science, Nutrition, and Health Promotion, P.O. Box 9805, Mississippi State University, Mississippi State, MS 39762*

Abstract

We use survey data to examine the relationship between food environment variables and the frequency of fruit and vegetable consumption in counties with high obesity rates in the Mississippi Delta. Results indicate that a lower vegetable, salad, and fruit consumption frequency is associated with longer distances traveled to a full-service grocery store, whereas access to public transportation is associated with a higher frequency of vegetable, fruit, and fruit juice consumption. The findings of this study can inform the development of localized interventions seeking to improve the food environment and increase fruit and vegetable consumption in rural communities.

Keywords: food access, food environment, fruit and vegetable consumption, obesity

ⓘCorresponding author: Tel: (405) 466-6167
Email: shannon.giscombe@langston.edu

Introduction

Policy makers and health officials have long been concerned with poor-quality diets due to their association with diet-related chronic diseases, including obesity, cardiovascular disease, high blood pressure, and type 2 diabetes. The promotion of healthy diets is especially important given the growing trend in obesity rates in the United States over the past few decades. The rising prevalence of obesity is recognized as a national health epidemic with an immediate need for effective and sustainable interventions (Wang et al., 2020). A higher prevalence of obesity is often observed among African Americans, females, older adults, and individuals with a high school education or less (Hales et al., 2020).

In 2022, Mississippi had an obesity rate of 39.5% and was ranked fourth highest obesity prevalence in the United States (Americashealthrankings.org, 2024). Some counties in the Delta region of Mississippi have obesity rates of 40% and higher, representing some of the highest rates in the nation (countyhealthrankings.org, 2023). Additionally, Mississippi's poverty rate of 19.6% in 2023 was the highest in the country (Americashealthrankings.org, 2024), which suggests many households in this region lack the resources required to purchase the foods they need to live a healthy lifestyle. This relationship is supported by previous studies in Mississippi, which found evidence of poor dietary quality, a lower intake of key nutrients, and a higher intake of unhealthy foods, particularly among disadvantaged sociodemographic groups (Connell et al., 2007; McCabe-Sellers et al., 2007). Environmental disparities, and particularly the built food environment, may contribute to observed poor health outcomes, including higher rates of obesity in the Delta region.

The CDC (2021) defines the food environment as "the physical presence of food that affects a person's diet, a person's proximity to food store locations, a connected system that allows access to food, or the distribution of food stores, food service, and any physical entity by which food may be obtained." Previous studies examining the food environment and food availability found significantly higher levels of deprivation in grocery store access within low-income or otherwise socioeconomically disadvantaged communities (Morland et al., 2002; Connell et al., 2007; Powel et al., 2007; Larson, Stort, and Nelson, 2009; Ko et al., 2018). Similarly, studies found limited assortments of healthy food options in areas with low food access (Cheranides and Jeanicke, 2019). However, studies researching the influence of the food environment on obesity outcomes have predominantly found null and inverse associations between store availability and negative health outcomes, contrary to some expectations (Cobb et al., 2015).

Limited access to healthy and affordable food at the local level increases the likelihood that individuals must travel greater distances to access healthy food options (Kaiser, Carr, and Fontabella, 2017). Some studies suggest that lower levels of local food access directly affect the quality of residents' diets (Caspi et al., 2012), potentially increasing health risks associated with poor diet and nutrition (Hill-Briggs et al., 2021). However, there are unobservable or unmeasurable factors (e.g., preferences, perceptions) that affect diet quality and health outcomes but are not always controlled for in empirical studies due to the lack of individual data and measures. These factors create a potential for bias due to the use of causal inference methods or due to their unobservable and immeasurable nature (Cobb et al., 2015).

Multiple local, state, and federal initiatives have sought to improve the food environment, food choices, and associated health outcomes. The CDC's High Obesity Program (HOP) (CDC, 2023b) is one such federal initiative consisting of cooperative agreements with Cooperative Extension Services in counties with the nation's highest obesity rates (obesity rates of 40% or higher). Based on the premise that improvements to local food environments can lead to healthier consumer choices and health outcomes—including obesity rate reductions (Steeves, Martins, and Gittelsohn, 2014)—HOP's primary goal is to combat obesity through improved consumption of healthy foods and increased levels of physical activity.

Our study is part of the HOP-funded project titled, "Advancing, Inspiring, Motivating for Community Health through Extension" (AIM for CHangE), led by Mississippi State University. To identify potential strategies for improving the local food environment, the AIM for CHangE team conducted a community survey to assess the food environment of Mississippi Delta counties with the state's highest obesity rates. Specifically, the goal of this study is to examine the link between fruit and vegetable consumption frequency and food environment variables, such as distance traveled to the nearest full-service grocery store and access to healthy foods as measured by access to public transportation and whether respondents shop for food at convenience and/or dollar store formats. We found that, on average, residents in the target counties travel 13 miles to the nearest full-service grocery store. For comparison, the average U.S. household travels 2.19 miles to the nearest supermarket or large grocery store (Ver Ploeg et al., 2015). According to the USDA (2022), rural areas are considered low access if residents are within 10 to 20 miles of a grocery store or supermarket. Residents with limited access to grocery stores often resort to shopping at convenience and dollar store formats to meet their food needs. Our results suggest a lower frequency of vegetable and salad consumption associated with longer grocery store distance, but a higher consumption frequency when residents have access to public transportation. While our results are not novel in that they agree with previous findings in the literature, our study provides localized information that could be shared with community coalitions to commence discussions regarding the community and initiatives that could be implemented. Our assessment aims to provide insights for communities in these Delta counties and inform local strategies to address obesity from a food environment, food systems, and policy perspective.

Data and Methods

Survey Data

The data used in this study are from a community survey of seven Mississippi Delta counties (see Figure 1) with the state's highest obesity rates (obesity > 40.0%). The survey was administered by the AIM for CHangE team with the help of community coalitions using paper and online formats. Study participants were randomly recruited throughout the target counties using flyers advertising the survey and containing QR codes with links to the online survey. Flyers were posted in frequently visited locations in each of the counties. The team also administered in-person surveys to help solicit responses from individuals without internet access. Participants who completed the survey were entered into a raffle for a $25 gift card to a local retailer. The data collection took place between January 2020 and March 2020, with a total of 352 completed survey responses.

Given the approach used to recruit participants, it is difficult to calculate a response rate. Because we excluded observations with missing data, we only used 222 observations in our analysis. The survey included questions pertaining to respondents' demographic characteristics, diet and nutrition, and physical activity.

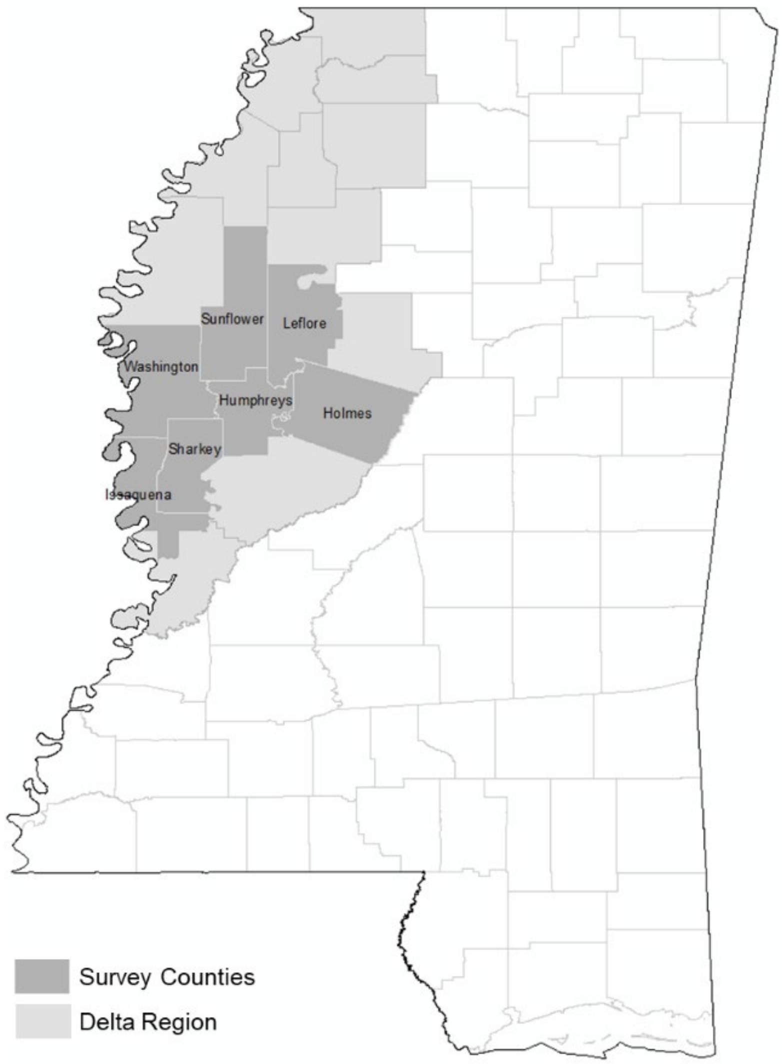


Figure 1. Map of the Mississippi Delta Region Highlighting the Targeted Counties Used for HOP Survey. Targeted Counties Include Holmes, Humphreys, Issaquena, Leflore, Sharkey, Sunflower, and Washington.

To measure the frequency of fruit and vegetable consumption, the survey included a simplified version of the National Cancer Institute's (NCI) Eating at America's Table Study (EATS) food frequency questionnaire, which is based on a 30-day dietary recall period (Thompson et al., 2011). We assess the consumption of specific food categories by asking respondents how often they ate or drank specific foods or beverages within the last 30 days. Foods and beverages include 100% fruit juice, fruits (fresh, frozen, and canned), lettuce salad consumed with or without other vegetables, and all other vegetables (raw, cooked, canned, and frozen, excluding lettuce salads and potatoes). The original survey responses had seven frequency categories, including "never," "1–3 times last month," "1–2 times a week," "3–4 times a week," "5–6 times a week," "once a day," and "more than once a day." These categories were then grouped into three categories, including *Monthly* (respondents reporting "never" or "1–3 times last month"), *Weekly* ("1–2 times a week" or "3–4 times a week"), and *Daily* ("5–6 times a week," "once a day," or "more than once a day"). While five to six times a week does not perfectly equate daily consumption, it was classified as *Daily* for the purpose of our analysis given its proximity. Changes in the original variable categories were made because of the limited number of responses within the "never" and "more than once a day" response categories for some of the fruit and vegetable groups.

The key independent variables of interest are measures of the respondent's local food environment. The survey included questions to gauge self-reported food accessibility in terms of public transportation and distance traveled to the nearest full-service grocery store. The variables included are the distance traveled to the nearest full-service grocery store (*Store Distance*) and an indicator of whether the community where the respondents live has any form of public transportation, such as bus routes (*Transportation*). To assess *Store Distance,* we specifically asked respondents, "How many miles do you have to travel to the nearest full-service grocery store, like Walmart or Sunflower, where you can get most of your groceries." Access to public transportation is included to control for accessibility of full-service grocery stores. A longer distance traveled to a grocery store is expected to be negatively correlated with fruit and vegetable consumption frequency (Connell et al., 2007; Michimi and Wimberly, 2010). While availability of public transportation is not a direct measure of transportation access, it helps control for individuals' ability to access grocery stores in cases where they may not have access to a personal vehicle. On average, Mississippi residents have limited access to personal vehicles, particularly among individuals with low grocery store access (USDA-ERS, 2020). We also include zip-code level population from the American Community Survey (U.S. Census Bureau, 2022) as a measure of rurality to account for the size of the location where residents live and control for food environment differences not captured by the variables in our survey.

Other food environment variables include indicators of where individuals report shopping for food. We include indicators for whether respondents shop for food at convenience stores (*Shop Conv Store*) or dollar store formats (*Shop Doll Store*). To obtain this information, we asked respondents, "Where do you get food in your county?" Respondents were able to select multiple responses from a list of options, including different store formats, farmers' markets, food banks and other assistance programs, home gardens, full-service restaurants, fast food restaurants, and other. Shopping at either a convenience store or a dollar store format is expected to be correlated with lower consumption frequencies of fruits and vegetables, as these store formats generally offer a less healthy and less varied assortment of food options (Larson, Stort, and Nelson, 2009; Canales et al., 2021). In addition to food environment variables, we included variables capturing what respondents believed to be barriers preventing higher fruit and vegetable consumption. The survey asked respondents if they would consume more fruits and vegetables if the prices were cheaper (*Price*), or if they tasted better (*Taste*). These variables control for respondent preferences as well as affordability.

We report summary statistics in Table 2. The original sample in our study had a larger proportion of women, older respondents, and respondents with a college degree when compared to the population in the study area (see Table 2), which is consistent with the profile of individuals who are generally more likely to respond to surveys (Curtin, Presser, and Singer, 2000). African Americans make up 71.2% of the Delta region but represented 81.8% of the survey sample. Given that our sample was not representative of the overall characteristics of the Delta region, we applied poststratification weighting using iterative proportional fitting or ranking in STATA v.18 (Bergmann, 2011). Poststratification weights were estimated based on the following distribution of demographic (U.S. Census) variables in our target region: age (18–60 years 71%, above 60 years 29%), gender (male 48%, female 52%), race (African American 71%, other races 29%), college degree (college 14%, no college 86%), and employment (employed 42%, other 58%). In Tables 1 and 2 we report both weighted and unweighted sample summary statistics. Because there is no means of verifying the representativeness of the findings against the general population of the area, any general extrapolations of the findings should be done with this caveat in mind.

Table 1. Summary Statistics for Categorical Dependent Variables

	Unweighted Sample				**Weighted Sample**			
Consumption Frequency	**Vegetables**	**Salad**	**Fruit**	**Fruit Juice**	**Vegetables**	**Salad**	**Fruit**	**Fruit Juice**
Monthly	28.8%	36.9%	18.0%	32.9%	29.4%	46.2%	18.3%	34.7%
Weekly	53.2%	50.9%	55.9%	45.5%	52.6%	39.0%	52.8%	43.2%
Daily	18.0%	12.2%	26.1%	21.6%	18.1%	14.8%	29.0%	22.2%

Table 2. Summary Statistics for Independent Variables

		Unweighted Sample		**Weighted Sample**	
Variable	**Description**	**Mean**	**St. Dev.**	**Mean**	**St. Dev.**
Age	Respondent's age in years	51.77	15.42	49.91	1.75
Gender	= 1 if Male	0.20	0.40	0.48	0.50
Race	= 1 if African American	0.82	0.38	0.71	0.45
College degree	= 1 if respondent has college education	0.42	0.49	0.14	0.35
Employed full time	= 1 if respondent is employed full time	0.58	0.49	0.42	0.49
Taste	= 1 if respondent would eat more vegetables if they tasted better	0.35	0.48	0.38	0.48
Price	= 1 if respondent would eat more vegetables if they were cheaper	0.51	0.50	0.44	0.50
Zip code population	Total population in zip code of residence	6,077	6,386	5,229	460
Store distance	Distance to a full-service grocery store from residence location	12.94	13.16	13.04	1.28

Table 2 (cont.)

Variable	Description	Unweighted Sample Mean	Unweighted Sample St. Dev.	Weighted Sample Mean	Weighted Sample St. Dev.
Shop conv store	= 1 if respondent shops at a convenience store	0.41	0.49	0.47	0.50
Shop dollar store	= 1 if respondent shops at a dollar store format	0.74	0.44	0.77	0.42
Transportation	= 1 public transportation is available in community	0.16	0.37	0.17	0.38
No. observations	222				

Note: The standard deviation of binary variables was calculated as $\sqrt{\rho(1-\rho)}$, where ρ is the mean value of the binary variable.

Regression Analysis

We used regression analysis to examine the associations among respondents' fruit and vegetable consumption patterns and their local food environment, demographic characteristics, socioeconomic status, and perceived barriers to fruit and vegetable consumption. Specifically, we used an ordered logit regression model to account for the discrete and ordered nature of the dependent variable. The dependent variable of interest, *y*, is the consumption frequency of vegetables, fruits, salad, and fruit juice. The frequency of consumption is a discrete categorical variable, with ordered potential responses of *Monthly*, *Weekly*, and *Daily*. We focus on fruit and vegetable consumption as a proxy for overall dietary quality, as the existing literature often finds a positive correlation between fruit and vegetable consumption and a healthier diet (Thompson et al., 2011; Aune et al., 2017; Schlesinger et al., 2019; Wallace et al., 2019).

In the order logit model, the unobserved latent dependent variable, y^*, is related to the observed dependent variable y (frequency of consumption) as follows:

$$y_i = \begin{cases} = Monthly & if \quad y_i^* \leq 0 \\ = Weekly & if \quad 0 < y_i^* \leq \tau_1 \\ = Daily & if \quad \tau_1 < y_i^* \leq \tau_2 \end{cases} \tag{1}$$

where τ_1 and τ_2 are unknown threshold parameters to be estimated. The regression model of y^* is specified such that:

$$y_i^* = \boldsymbol{\beta}'\boldsymbol{x_i} + \varepsilon_i \tag{2}$$

where $\boldsymbol{x_i}$ is a set of explanatory variables for individual *i* that includes food environment measures, demographic variables (race, age, gender, employment, and education), barriers to frequent fruit and vegetable consumption, and whether respondents shop at a dollar store or convenience store.

The error term ε has a standard logistic distribution. The model was estimated via maximum likelihood estimation in STATA V.18 (StataCorp, 2021).

Results and Discussion

Table 1 and Table 2 present summary statistics for the dependent and independent variables in our study, respectively. Our analysis will focus on the weighted sample. Based on survey responses, the average distance respondents travel to the nearest full-service grocery store where they can meet all their grocery needs is 13.0 miles (see Table 2). However, some respondents reported traveling significantly longer distances than the average, as depicted in Figure 2. For example, several respondents reported traveling between 20 miles and 55 miles to reach a grocery store. Extended distances to a full-service grocery store may partially account for the high percentage of survey respondents who report shopping at convenience (46.5%) and dollar stores (76.6%) for their grocery needs. Limited proximity to a full-service grocery store may prompt some individuals to supplement their grocery purchases with purchases at dollar store formats or convenience stores, which are often more accessible (i.e., higher store density) than supermarkets and full-service grocery stores in the Mississippi Delta region (Canales et al., 2021).

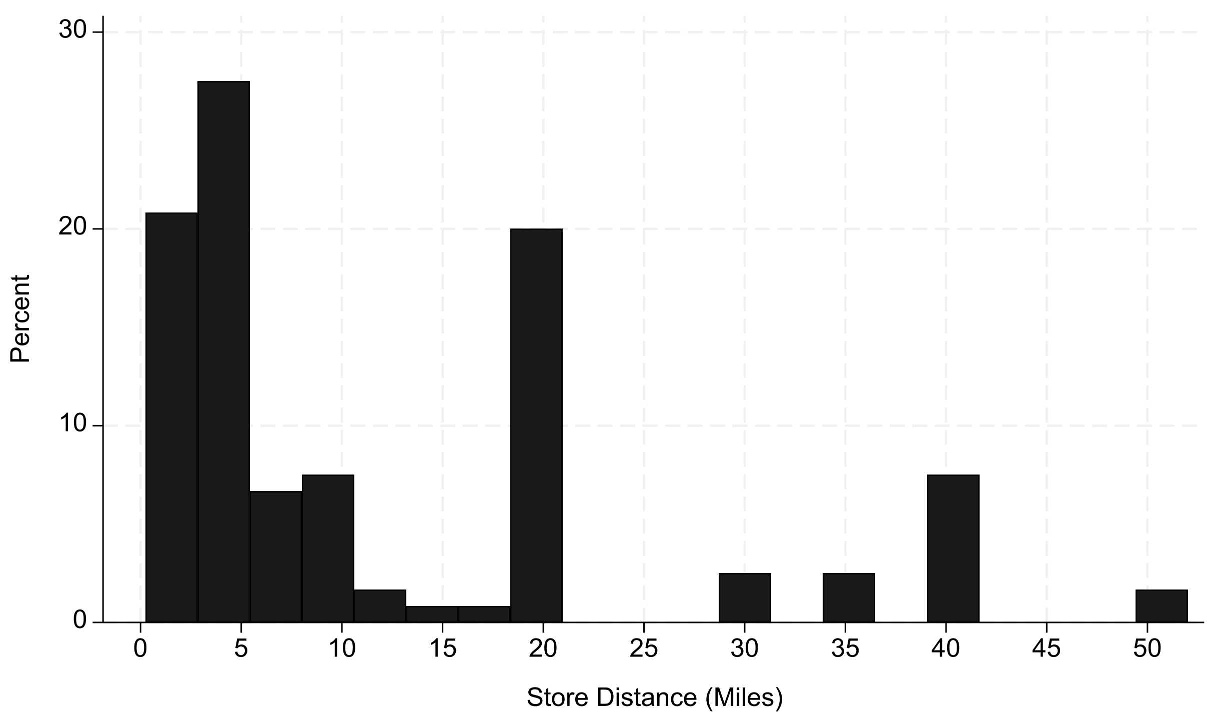


Figure 2. Distribution of Distances Traveled to the Nearest Full-Service Grocery Store by Survey Respondent, Weighted Sample

After being asked what respondents believed would help them eat healthier, we found that less than half of the sample view the taste of the food (37.7%) as a potential factor that would help them improve their vegetable intake (see Table 2). Our results also suggest that 44.2% of respondents saw price as a barrier to consuming more vegetables (i.e., they indicate they would eat more vegetables if they were cheaper), creating an area of concern surrounding the choice to eat less healthy options due to food prices. A previous study by Sharkey at al. (2010) found price to be a recurring barrier to healthy food consumption in rural areas.

Using a 30-day food recall method, we found that many respondents in our sample are not consuming fruits and vegetables daily as recommended by the Dietary Guideline of America (Dietaryguideline.gov). Figure 3 shows the percentage of respondents reporting Daily, Weekly, and Monthly for our fruit and vegetable consumption categories. Most individuals report weekly consumption of fruits (52.8%) and vegetables (52.6%). Only 18.1% of individuals in our sample consume vegetables daily, and only 14.8% consume salad daily. Although these respondents report consuming fruit and vegetables daily, they do not necessarily consume the recommended nutritional intake of 1.5 cups per day for fruits and 2−3 cups per day for vegetables (USDA and USDHHS, 2020). These findings are consistent with those of a previous study conducted in the Mississippi Delta. McCabe-Seller et al. (2007) found an overall lower diet quality in the lower Mississippi Delta area, when compared to white and African American adults in the National Health and Nutrition Examination Survey (NHANES) of 1999–2000. Based on a 24-hour recall method, the authors found that less than 25% and 16% of adults meet the vegetable and fruit intake recommendation, respectively.

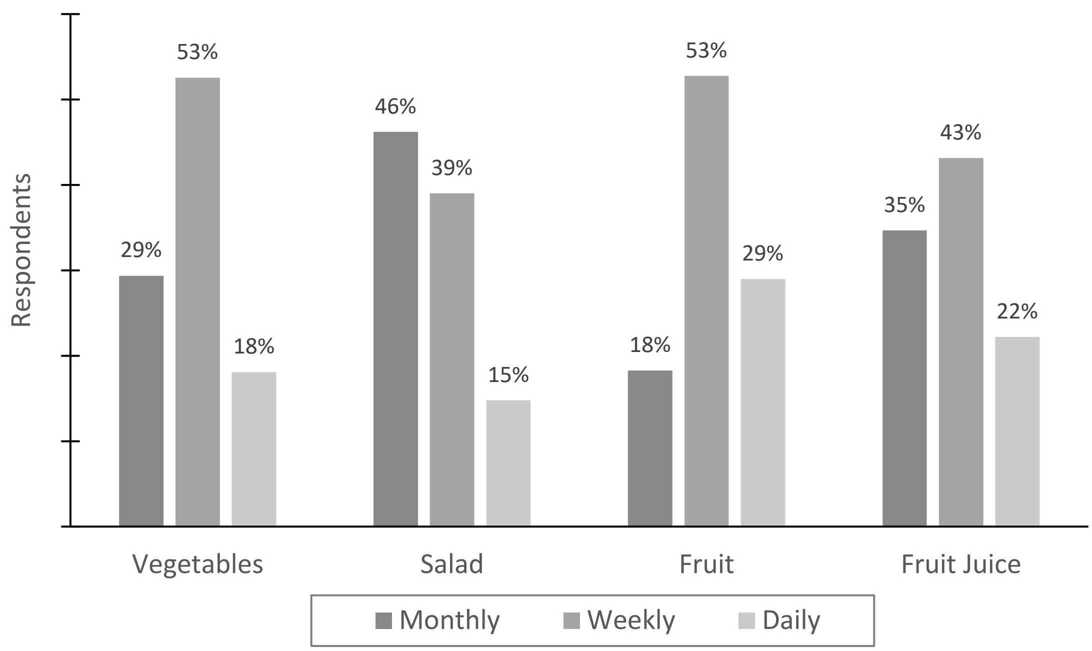


Figure 3. Frequency of Consumption of Fruits and Vegetables from Survey Respondents for Weighted Sample

The results of the ordered logit model are reported in Table 3. We present results for the weighted sample in the main text, whereas results for the unweighted sample are reported in Appendix A1.

When applying weights, the results are similar to results using the unweighted sample. To increase the number of observations, we also imputed missing responses while applying weights. Imputation resulted in a sample of 250, and the results of the ordered logit for this sample are reported in Appendix A2. The results on the imputed and weighted sample are similar to the results reported in Table 3. Given the nonlinear functional form of the ordered logit model, the magnitudes of the coefficients are not directly interpretable, and the signs of the coefficients only show whether the dependent variable (frequency of consumption) increases or decreases given a change in each explanatory variable. To aid in effective interpretation, we report the average marginal effects from the ordered logit estimates (see Table 3) in Table 4. The marginal effects can be interpreted as the average change in the probability of each consumption frequency (i.e., daily, weekly, monthly consumption), given a 1-unit change in the explanatory variables.

We find a significant negative association between store distance and frequent consumption of vegetables, salads, and fruits. Average marginal effects for store distance indicate that for each additional mile that an individual must travel to a full-service grocery store, they are 0.3 percentage points less likely to consume vegetables (p-value < 0.10) and salads daily (p-value < 0.05), and 0.6 percentage points less likely to consume fruits daily (p-value < 0.10). For each additional mile, respondents were also 0.5, 0.7, and 0.4 percentage points more likely to consume vegetables (p-value < 0.05), salads (p-value < 0.05), and fruit (p-value < 0.10) less frequently only on a monthly basis, respectively. In previous studies, farther commute distances to a full-service grocery store are associated with decreased fruit and vegetable intake (Rose and Richards, 2004; Michimi and Wimberly, 2010; Sharkey, Johnson, and Dean, 2010). For example, Sharkey, Johnson, and Dean (2010) found a 1.2 percentage point decrease in fruit consumption for each additional mile to a store with a good selection of food. Other studies in the literature also found that longer distances are correlated with eating less healthy food options, which has a disproportionate negative effect on disadvantaged groups (Connell et al., 2007; Jilcott et al., 2010; Michimi and Wimberly, 2010).

Table 3. Regression Results for Logit Model on the Frequency of Consumption for Vegetables, Salads, Fruits, and Fruit Juice, Weighted Sample

	Vegetables		Salad		Fruit		Fruit Juice	
	Coeff.	**Std. Error**	**Coeff.**	**Std. Error**	**Coeff.**	**Std. Error**	**Coeff.**	**Std. Error**
Age	-0.003	(0.013)	0.018	(0.017)	0.002	(0.013)	-0.020	(0.017)
Gender (male = 1)	-0.135	(0.426)	-0.356	(0.441)	0.202	(0.452)	0.044	(0.432)
Race (African American = 1)	-0.890**	(0.365)	0.724	(0.475)	0.087	(0.450)	1.231**	(0.573)
College degree	0.150	(0.325)	0.217	(0.345)	-0.116	(0.449)	0.370	(0.387)
Employed full time	-0.329	(0.429)	0.657	(0.502)	-0.726*	(0.375)	-1.434***	(0.463)
Taste	0.500	(0.404)	0.360	(0.445)	-0.183	(0.506)	0.013	(0.520)
Price	0.744	(0.457)	0.357	(0.430)	0.141	(0.434)	-0.360	(0.460)
Zip code population	-0.000	(0.000)	-0.000	(0.000)	-0.000	(0.000)	0.000	(0.000)
Store distance	-0.025*	(0.013)	-0.032**	(0.013)	-0.032*	(0.019)	-0.017	(0.012)
Shop conv store	0.039	(0.443)	-0.043	(0.480)	0.279	(0.464)	-0.582	(0.563)
Shop dollar store	-0.237	(0.486)	0.337	(0.448)	0.229	(0.424)	0.439	(0.647)
Transportation	0.819*	(0.484)	1.019**	(0.401)	0.431	(0.569)	0.876**	(0.407)
τ_1	-1.852**	(0.754)	1.653	(1.047)	-2.024**	(0.796)	-1.450	(1.165)
τ_2	0.836	(0.803)	3.803***	(1.184)	0.589	(0.838)	0.829	(1.120)
No. observations	222		222		222		222	

Note: ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively

Table 4. Estimated Marginal Effects of the Independent Variables on the Frequency of Consumption

	Vegetables		Salad		Fruit		Fruit juice	
	Marginal Effect	**St. Error**	**Marginal Effect**	**St. Error**	**Marginal Effect**	**St. Error**	**Marginal Effect**	**St. Error**
Age								
Monthly	0.001	(0.002)	-0.004	(0.004)	0.000	(0.002)	0.004	(0.003)
Weekly	0.000	(0.001)	0.002	(0.002)	0.000	(0.001)	-0.001	(0.001)
Daily	0.000	(0.002)	0.002	(0.002)	0.000	(0.002)	-0.003	(0.002)
Gender								
Monthly	0.025	(0.080)	0.076	(0.095)	-0.028	(0.061)	-0.008	(0.080)
Weekly	-0.007	(0.023)	-0.036	(0.050)	-0.011	(0.028)	0.002	(0.016)
Daily	-0.018	(0.057)	-0.040	(0.047)	0.039	(0.089)	0.007	(0.064)
Race								
Monthly	0.154**	(0.062)	-0.157	(0.101)	-0.012	(0.064)	-0.251**	(0.117)
Weekly	-0.021	(0.029)	0.082	(0.065)	-0.004	(0.021)	0.091	(0.073)
Daily	-0.134**	(0.056)	0.075*	(0.040)	0.017	(0.085)	0.160***	(0.057)
College								
Monthly	-0.028	(0.061)	-0.046	(0.074)	0.016	(0.062)	-0.069	(0.070)
Weekly	0.007	(0.018)	0.022	(0.036)	0.006	(0.025)	0.014	(0.019)
Daily	0.020	(0.044)	0.025	(0.038)	-0.022	(0.087)	0.055	(0.055)
Employed								
Monthly	0.061	(0.079)	-0.142	(0.107)	0.104*	(0.055)	0.277***	(0.079)
Weekly	-0.017	(0.022)	0.065	(0.056)	0.033	(0.030)	-0.073*	(0.037)
Daily	-0.044	(0.059)	0.077	(0.056)	-0.137*	(0.073)	-0.204***	(0.072)
Taste								
Monthly	-0.092	(0.074)	-0.077	(0.095)	0.025	(0.071)	-0.002	(0.097)
Weekly	0.024	(0.026)	0.036	(0.043)	0.010	(0.025)	0.000	(0.019)
Daily	0.068	(0.054)	0.041	(0.053)	-0.035	(0.095)	0.002	(0.077)
Price								
Monthly	-0.137*	(0.078)	-0.076	(0.092)	-0.020	(0.060)	0.067	(0.084)
Weekly	0.036	(0.023)	0.036	(0.040)	-0.007	(0.023)	-0.013	(0.019)
Daily	0.102	(0.069)	0.041	(0.053)	0.027	(0.083)	-0.053	(0.068)
Zip code population								
Monthly	0.000	(0.000)	0.000	(0.000)	0.000	(0.000)	0.000	(0.000)
Weekly	0.000	(0.000)	0.000	(0.000)	0.000	(0.000)	0.000	(0.000)
Daily	0.000	(0.000)	0.000	(0.000)	0.000	(0.000)	0.000	(0.000)
Store distance								
Monthly	0.005**	(0.002)	0.007**	(0.003)	0.004*	(0.003)	0.003	(0.002)
Weekly	-0.001	(0.001)	-0.003***	(0.001)	0.002	(0.001)	-0.001	(0.001)
Daily	-0.003*	(0.002)	-0.004**	(0.002)	-0.006*	(0.003)	-0.002	(0.002)
Shop conv store								
Monthly	-0.007	(0.082)	0.009	(0.103)	-0.039	(0.064)	0.108	(0.104)
Weekly	0.002	(0.022)	-0.004	(0.048)	-0.015	(0.026)	-0.022	(0.028)
Daily	0.005	(0.060)	-0.005	(0.055)	0.053	(0.088)	-0.086	(0.082)

Table 4 (cont.)

	Vegetables		Salad		Fruit		Fruit juice	
	Marginal Effect	St. Error	Marginal Effect	St. Error	Marginal Effect	St. Error	Marginal Effect	St. Error
Shop dollar store								
Monthly	0.044	(0.090)	-0.072	(0.095)	-0.032	(0.059)	-0.082	(0.121)
Weekly	-0.011	(0.027)	0.034	(0.042)	-0.012	(0.024)	0.016	(0.027)
Daily	-0.032	(0.065)	0.039	(0.054)	0.044	(0.082)	0.065	(0.097)
Transportation								
Monthly	-0.151*	(0.087)	-0.218***	(0.081)	-0.060	(0.078)	-0.163**	(0.073)
Weekly	0.040	(0.035)	0.102**	(0.048)	-0.023	(0.032)	0.033	(0.031)
Daily	0.112*	(0.066)	0.117**	(0.046)	0.083	(0.108)	0.130**	(0.054)

Note: ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

A low density of full-service grocery stores and supermarkets has been shown to decrease healthy food access among rural, low-income, and elderly residents (Morland et al., 2002; Hendrickson, Smith, and Eikenberry, 2006; McGee et al., 2011). The lack of local access to full-service grocery stores forces residents to travel longer distances to meet their food needs. Consequently, individuals with limited access resort to purchasing foods at alternative store formats, including convenience and dollar stores. While we expected the store format would influence frequency of fruit and vegetable consumption due to differences in the assortment of fresh fruit and vegetables, we did not find any statistically significant difference in fruit and vegetable consumption between individuals who shop at convenience and dollar stores and those who do not. In a similar vein, a systematic review of studies examining the effect of the food environment on obesity found limited statistical evidence that store availability affects obesity (Cobb et al., 2015).

Limited access to transportation can further exacerbate the negative association between longer store distances and the consumption of healthy foods, as it may constrain individuals' ability to reach well-assorted stores, such as supermarkets and grocery stores. Our results suggest that respondents living in areas with access to some form of public transportation are 11.2 percentage points more likely to consume vegetables (p-value < 0.10), 11.7 percentage points more likely to consume salad (p-value < 0.05), and 13.0 percentage points more likely to consume fruit juice daily (p-value < 0.05). Similarly, in areas where respondents indicated having access to transportation, they were 15.1, 21.8, and 16.3 percentage points less likely to consume vegetables, salad, and fruit juice only monthly, respectively. These findings are consistent with previous studies that demonstrate the adverse impact of unreliable transportation on food access, particularly among low-income households (Connell et al., 2007; Kaiser et al., 2017). The observed marginal effects of public transportation underscore the potential need for interventions that facilitate ease of access to healthy food in disadvantaged communities through improved public transportation systems. There is a need to implement more accessible public transportation to bridge the gap between the lack of personal transportation and spatial access to supermarkets in the Mississippi Delta region. Facilitating access to bus routes or rideshare programs in rural areas could be one strategy worth examining, as increasing the presence of large food retail stores in rural areas is challenging due to the high entry and operation costs, supply chain issues, and limited demand, which make the retail market in low-income and rural areas unattractive to larger food retail outlets (Paddison and Calderwood, 2007; Cheranides and Jeanicke, 2019).

With regard to perceived barriers, individuals who perceive price as a barrier to consuming more vegetables were 13.7 percentage points (p-value < 0.05) less likely to consume vegetables monthly compared to individuals who do not perceive price as a barrier. Price is commonly identified as a barrier to the consumption of healthier and more expensive food options. Prior studies found that price barriers decrease the probability of eating healthier (French, 2003; Kaiser et al., 2017), which is sometimes explained by the higher cost per serving of fruits and vegetables in rural areas. One study conducted in the Mississippi Delta found that the price per serving of fruits and vegetables is higher in the Delta relative to the national average price per serving (Connell et al., 2012). Another study also found that prices of healthier foods—such as fruits and vegetables–are higher in counties in Mississippi with high obesity rates when compared to prices in counties with lower obesity rates (Fan et al., 2021). Overall, however, findings by Carlson and Frazao (2012) indicate that healthy foods are not always more expensive than less healthy foods. The decrease in infrequent consumption despite price being perceived as a barrier in our study may indicate that individuals who acknowledge price as a barrier may also want to eat vegetables and consume them more frequently. It can be inferred that individuals allocate spending toward different food items based on factors other than healthy eating and meeting dietary recommendations.

According to McGee et al. (2011), while residents may perceive price barriers to purchasing healthy foods, personal preferences and individual family members' preferences tend to have a greater influence on food purchasing behaviors. For individuals with lower fruit and vegetable consumption in our sample, the consumption of healthier food options might be due to preferences and other behavioral components rather than factors that are generally expected to prevent more frequent fruit and vegetable consumption like price. A limitation of our study is that we were not able to capture the effects of preferences. To do this, we would need to collect data on respondents' preferences over different types of food, as well as data on how prices, availability, and accessibility affect their choices to consume one food item compared to other food items.

An implication of our findings regarding respondents' perceived barriers can be the implementation of behavioral interventions in the Mississippi Delta region to address high obesity rates, as direct solutions targeting price barriers may not prove effective among individuals who consume fruits and vegetables infrequently. Several existing policies have focused on decreasing healthier food prices in efforts to increase healthier food consumption. While this approach may prove effective, it may miss the target audience of those consuming fruit and vegetables less frequently and whether they do not perceive price as a barrier to consumption. Initiatives should identify strategies that target individuals who do not view food price as a barrier to consuming more fruits and vegetables and whose low consumption may be due instead to dietary preferences. This alternative approach could improve the potential effectiveness of policies designed to improve consumption frequency and reduce the occurrence of obesity and noncommunicative weight-related health risks. In our study, we did not find a statistically significant association between the taste of food as a barrier and the frequency of fruit and vegetable consumption.

We included full-time employment and college education in our regression as a proxy measure of respondents' socioeconomic status. Socioeconomic status partially helps to shape individual food consumption choices as well as the consumption frequency of certain foods based on cost, accessibility, and other related factors. While we did not find that attending college has a significant effect on consumption frequency, we found that individuals with full employment were less likely to consume fruit and fruit juice daily but more likely to consume them monthly. While we had expected fruit consumption to be more frequent among the employed, it is important to note that the fruit category in the survey included the consumption of canned products, which are affordable and more widely available at various store formats compared to fresh fruits.

With regard to differences in the frequency of consumption across demographic groups, we found that African Americans were 13.4 percentage points (p-value < 0.05) less likely to consume vegetables daily and were 15.4 percentage points (p-value < 0.05) more likely to consume vegetables on a less frequent monthly basis. African Americans, on the other hand, were more likely (16.0 percentages points) to consume fruit juice daily (p-value < 0.01). As differences in health outcomes are observed, it is important to understand differences in consumption frequency across demographic groups.

Conclusion

The goal of our study was to examine the relationship between local food environment factors and the consumption of fruits and vegetables among individuals living in the Mississippi Delta, a region with one of the nation's highest obesity rates. Specifically, we studied how consumption patterns of healthy foods are affected by proximity to full-service grocery stores, healthy food accessibility as measured by access to public transportation, and whether respondents shop for food at convenience or dollar store formats. We also examined differences in fruit and vegetable consumption frequencies across groups based on demographic characteristics, such as reported age, gender, race, employment, and educational attainment. Results from our study provide insights for communities in the Mississippi Delta and may inform local strategies to address obesity from a food environment, food systems, and policy perspective. These findings are particularly important for policy makers seeking to address issues within food systems in the Mississippi Delta region.

The food environment measures were statistically significant across the various food groups considered. We found that individuals who travel longer distances to the nearest full-service grocery store were less likely to consume vegetables frequently (i.e., daily). This finding is informative, particularly when considering the effects of proximity and store density within rural Mississippi Delta communities. On average, individuals in our sample reported traveling 13 miles from their residence to a full-service grocery store, with several survey respondents traveling between 20 miles and 55 miles. The longer travel distances required to access full-service stores could explain why many respondents (76.6%) shop for groceries at dollar store formats, which are more accessible (i.e., higher store density) when compared to supermarkets and full-service grocery stores. Understanding this aspect of the local food environment is particularly insightful for initiatives geared toward improving healthy consumption via increasing access to the different food shopping options that are available to individuals. In many cases, supermarket or grocery store operators do not find it economically viable to locate in certain areas. In such cases, it is important to identify strategies that promote healthier food assortments in existing stores and improve physical access to healthy food through channels like transportation infrastructure. Such strategies could include increasing access to public transportation. The availability of public transportation, as a measure of accessibility, is another statistically significant food environment variable in our study. Our results suggest that public transportation access increases the likelihood that individuals consume vegetables and salad more frequently, highlighting the potential importance of transportation service availability, particularly in areas with low store access, like the Mississippi Delta.

The data used in our study have some important limitations. The first limitation worth acknowledging is the relatively small sample size and representativeness of the sample. To address this issue, we used poststratification weights and imputation of missing data and found consistent results. Because we cannot verify the representativeness of the findings against the general population of the area, any general extrapolations of the findings should be done with this data limitation in mind. Second, because the data collection took place between January and March of 2020, our reported frequency of fruit and vegetable consumption would likely be lower when compared to the average annual consumption, due to lower availability of fresh produce during the winter months. Because the availability of fresh produce during the winter is lower across all store formats and that other explanatory factors are not likely to change seasonally, we do not expect the association between the explanatory variables and consumption frequency to be affected. It is also important to note that the consumption frequency questions in the survey asked for consumption in all forms (fresh, canned, and frozen). In the case that overall fruit and vegetable availability was more restrictive during the winter in convenience and dollar stores relative to supermarkets, we would expect to see lower consumption frequency associated with shopping at these store formats. However, we did not find any statistical differences in consumption associated with store format. Another possible issue with the timing of the survey is COVID-19, which was officially declared a pandemic on March 11, 2020, with the implementation of shutdowns beginning March 15, 2020 (CDC, 2023a). While most of the data had been collected at that point, it was foreseeable that the pandemic altered shopping and consumption patterns (e.g., less frequent visits to crowded stores and greater consumption of canned and frozen products). Given all of these potential dynamics, it is difficult to assess how the pandemic could have affected the direction of the effect of the explanatory variables in our study on consumption frequency. Third, we did not explore the role of food away from home and access to retail food service establishments on the frequency of fruit and vegetable consumption, highlighting an area for possible future research. If consumption away from home is negatively correlated to fruit and vegetable consumption and negatively (positively) correlated to the likelihood of purchasing foods at dollar or convenience stores, it is possible that we are overestimating (underestimating) the effect of shopping at convenience or dollar stores. As seen in our results, we did not find any statistical difference in consumption when respondents shopped at dollar or convenience stores. Lastly, the HOP Community Survey did not include some key variables of interest, such as income, a major determinant of socioeconomic status that could play a role in an individual's ability to afford a healthy diet. It is also possible that unobserved factors, such as preferences, play a large role in consumption decisions. Understanding consumer preferences for fruit and vegetables could inform behavior-based interventions related to the food environment of the Mississippi Delta.

Notwithstanding the limitations of the data used in our study, we believe the results provide useful insights regarding the food environment in the Mississippi Delta region and how food environment factors may play a role in fruit and vegetable consumption frequency. These insights can be used to inform further research and outreach and provide a starting point for conversations about initiatives to improve the food environment based on the unique conditions and characteristics of the population examined.

References

Aune, D., E. Giovannucci, P. Boffetta, L.T. Fadnes, N. Keum, T. Norat, D.C. Greenwood, E. Riboli, L.J. Vatten, and S. Tonstad. 2017. "Fruit and Vegetable Intake and the Risk of Cardiovascular Disease, Total Cancer and All-Cause Mortality—A Systematic Review and Dose-Response Meta-Analysis of Prospective Studies." *International Journal of Epidemiology* 46(3):1029–1056.

Bergmann, M. 2011. *IPFWEIGHT: Stata Module to Create Adjustment Weights for Surveys*. Statistical Software Components S457353. Chestnut Hill, MA: Boston College Department of Economics.

Canales, E., L. Fan, D.R. Buys, and M.D. Cantave. 2021. "A Market Basket Assessment: Prices and Availability of Healthy Foods across SNAP-Authorized Food Outlets in Counties with High Obesity Rates in Mississippi." *Preventing Chronic Disease* 18:210173.

Carlson, A., and E. Frazao. 2012. *Are Healthy Foods Really More Expensive? It depends on How You Measure the Price*. Washington, DC: U.S. Department of Agriculture, Economic Research Service, Economic Information Bulletin EIB-96.

Caspi, C.E., G. Sorensen, S.V. Subramanian, and I. Kawachi. 2012. "The Local Food Environment and Diet: A Systematic Review." *Health & Place* 18(5):1172–1187.

Centers for Disease Control and Prevention. 2021. *Healthy Food Environments*. Washington, DC: Centers for Disease Control and Prevention. Available online: https://www.cdc.gov/nutrition/healthy-food-environments/index.html.

Centers for Disease Control and Prevention. 2023a. *CDC Museum COVID-19 Timeline*. Washington, DC: Centers for Disease Control and Prevention. Available online: https://www.cdc.gov/museum/timeline/covid19.html.

Centers for Disease Control and Prevention. 2023b. *D. High Obesity Program (HOP), 2018-2023*. Washington, DC: Centers for Disease Control and Prevention. Available online: https://archive.cdc.gov/www_cdc_gov/nccdphp/dnpao/state-local-programs/hop/high-obesity-program-1809.html [Accessed November 23].

Cheranides, L., and E.C. Jaenicke. 2019. "Documenting the Link between Poor Food Access and Less Healthy Product Assortment across the U.S. Applied Economic Perspectives and Policy." *Applied Economics and Perspectives* 41(3):434–474.

Cobb, L.K., L.J. Appel, M. Franco, J.C. Jones-Smith, A. Nur, and C.A. Anderson. 2015. "The Relationship of the Local Food Environment with Obesity: A Systematic Review of Methods, Study Quality, and Results." *Obesity* 23(7):1331–1344.

Connell, C.L., M.K. Yadrick, P. Simpson, J. Gossett, B.B. McGee, and M.L. Bogle. 2007. "Food Supply Adequacy in the Lower Mississippi Delta." *Journal of Nutrition Education and Behavior* 39:77–83.

Connell, C.L., M.K. Yadrick, J.M. Zoellner, S.C. Chekuri, L.B. Crook, and M.L. Bogle. 2012. "Energy Density, Nutrient Adequacy, and Cost per Serving Can Provide Insight into Food Choices in the Lower Mississippi Delta." *Journal of Nutrition Education and Behavior* 44:148–153.

Curtin, R., S. Presser, and E. Singer. 2000. "The Effects of Response Rate Changes on the Index of Consumer Sentiment." *Public Opinion Quarterly* 64(4):413–428.

Fan, L., E. Canales, B. Fountain, D. Buys. 2021. "An Assessment of the Food Retail Environment in High Obesity Counties in Mississippi." *Journal of Hunger and Environmental Nutrition* 16(4):571–593.

French, S.A. 2003. "Pricing Effect on Store Choice." *Journal of Nutrition* 133(3):841S–843S.

Hales, C.M., M.D. Carroll, C.D. Fryar, and C.L. Ogden. 2020. *Prevalence of Obesity and Severe Obesity among Adults: United States, 2017–2018*. Hyattsville, MD: National Center for Health Statistics, NCHS Data Brief, No 360. Available online: https://www.cdc.gov/nchs/products/databriefs/db360.htm

Hendrickson, D., C. Smith, and N. Eikenberry. 2006. "Fruit and Vegetable Access in Four Low-Income Food Desert Communities in Minnesota." *Agriculture and Human Values* 23:371–383.

Hill-Briggs, F., N.E. Adler, S.A. Berkowitz, M.H. Chin, T.L. Gary-Webb, A. Navas-Acien, P.L. Thornton, and D. Haire-Joshu. 2021. "Social Determinants of Health and Diabetes: A Scientific Review." *Diabetes Care* 44(1):258–279.

Jilcott, S.B., H. Liu, J.B. Moore, J.W. Bethel, J. Wilson, and A.S. Ammerman. 2010. "Commute Times, Food Retail Gaps, and Body Mass Index in North Carolina Counties." *Preventing Chronic Disease* 7(5):A107.

Kaiser, M.L., J.C. Carr, and S. Fontabella. 2017. "A Tale of Two Food Environments: Differences in Food Availability and Food Shopping Behaviors between Food Insecure and Food Secure Households." *Journal of Hunger and Environmental Nutrition* 14:297–317.

Ko, L.K., C. Enzler, C.K. Perry, E. Rodriguez, N. Mariscal, S. Linde, and C. Duggan. 2018. "Food Availability and Food Access in Rural Agricultural Communities: Use of Mixed Methods." *BMC Public Health* 18:634.

Larson, N.I., M.T. Stort, and M.C. Nelson. 2009. "Neighborhood Environments: Disparities in Access to Healthy Foods in the U.S." *American Journal of Preventive Medicine* 36(1):74–81.e10.

McCabe-Seller, B.J., S. Bowman, J.E. Stuff, C.M. Champagne, P.M. Simpson, and M.L. Bogle. 2007. "Assessment of the Diet Quality of US Adults in the Lower Mississippi Delta." *American Journal of Clinical Nutrition* 86:697–706.

McGee, B.B., G.S. Johnson, M.K. Yadrick, V. Richardson, P.M. Simpson, J.M. Gossett, A. Thornton, C. Johnson, and M.L. Bogle. 2011. "Food Shopping Perceptions, Behaviors, and Ability to Purchase Healthful Food Items in the Lower Mississippi Delta." *Journal of Nutrition Education and Behavior* 43(5):339–348.

Michimi, A., and M.C. Wimberly. 2010. "Associations of Supermarket Accessibility with Obesity and Fruit and Vegetable Consumption in the Conterminous United States." *International Journal of Health Geographics* 9:49.

Morland, K., S. Wing, A.D. Roux, and C. Poole. 2002. "Neighborhood Characteristics Associated with the Location of Food Stores and Food Service Places." *American Journal of Preventive Medicine* 2002:23–29.

Paddison, A., and E. Calderwood. 2007. "Rural Retailing: A Sector in Decline?" *International Journal of Retail and Distribution Management* 35(2):136–155.

Powell, L.M., S. Slater, D. Mirtcheva, Y. Bao, and F.J. Chaloupka. 2007. "Food Store Availability and Neighborhood Characteristics in the United States." *Preventive Medicine* 44(3):189–195.

Rose, D., and R. Richards. 2004. "Food Store Access and Household Fruit and Vegetable Use among Participants in the US Food Stamp Program." *Public Health Nutrition* 7(8):1081–1088.

Schlesinger, S., M. Neuenschwander, C. Schwedhelm, G. Hoffmann, A. Bechthold, H. Boeing, and L. Schwingshackl. 2019. “Food Groups and Risk of Overweight, Obesity, and Weight Gain: A Systematic Review and Dose-Response Meta-Analysis of Prospective Studies.” *Advances in Nutrition* 10(2):205–218.

Sharkey, J.R., C.M. Johnson, and W.R. Dean. 2010. “Food Access and Perceptions in the Community and Household Food Environment as Correlates of Fruits and Vegetable Intake among Rural Seniors.” *BMC Geriatrics* 10(32),

StataCorp. 2021. *Stata Statistical Software: Release 17*. College Station, TX: StataCorp LLC.

Steeves, E.A., P.A. Martins, and J. Gittelsohn. 2014. “Changing the Food Environment for Obesity Prevention: Key Gaps and Future Directions.” *Current Obesity Reports* 3(4):451–458.

Thompson, J.L., S.J. Onufrak, C.L. Connell, J.M. Zoellner, L.M. Tussing-Humphreys, M.L. Bogle, and K. Yadrick. 2011. “Food and Beverage Choices Contribution to Dietary Guidelines Adherence in the Lower Mississippi Delta.” *Public Health Nutrition* 14(12):2099–2109.

United Health Foundation. 2024. *America's Health Rankings, Summary of MS*. Available online: https://www.americashealthrankings.org/explore/states/MS.

U.S. Census Bureau. 2022. *Age and Sex. American Community Survey, ACS 5-Year Estimates Subject Tables, Table S0101.* Washington, DC: U.S. Department of Commerce, Bureau of the Census. Available online: https://data.census.gov/table/ACSST5Y2022.S0101?g=040XX00US28$8600000_860XX00US54855.

University of Wisconsin Population Health Institute. *County Health Rankings. Adult Obesity in Mississippi*. Madison, WI: University of Wisconsin Population Health Institute. Available online: https://www.countyhealthrankings.org/explore-health-rankings/county-health-rankings-model/health-factors/health-behaviors/diet-and-exercise/adult-obesity?year=2023&state=28&tab=1.

Ver Ploeg, M., L. Mancino, J.E. Todd, D.M. Clay, and B. Scharadin. 2015. “Where Do Americans Usually Shop for Food and How Do They Travel to Get There? Initial Findings from the National Household Food Acquisition and Purchase Survey.” *AgEcon Search.* 10.22004/ag.econ.262116. Available online: https://ageconsearch.umn.edu/record/262116.

Wallace, T.C., R.L. Bailey, J.B. Blumberg, B. Burton-Freeman, C.O. Chen, K.M. Crowe-White, A. Drewnowski, S. Hooshman, E. Johnson, R. Lewis, R. Murray, S.A. Shapses, and D.D. Wand. 2019. “Fruits, Vegetables, and Health: A Comprehensive Narrative, Umbrella Review of the Science and Recommendations for Enhanced Public Policy to Improve Intake.” *Critical Reviews and Food Science and Nutrition* 60(13):2174–2211.

Wang, Y., M.A. Beydoun, J. Min, H. Xue, L.A. Kaminsky, and L.J. Cheskin. 2020. "Has the Prevalence of Overweight and Obesity Levelled Off in the United States? Trends, Patterns, Disparities, and Future Projections for the Obesity Epidemic." *International Journal of Epidemiology* 49(3):810–823.

Appendix A1. Regression Results for Logit Model on the Frequency of Consumption for Vegetables. Salads, Fruits, and Fruit Juice, Unweighted Sample

	Vegetables		**Salad**		**Fruit**		**Fruit Juice**	
	Coeff.	**Std. Error**	**Coeff.**	**Std. Error**	**Coeff.**	**Std. Error**	**Coeff.**	**Std. Error**
Age	0.007	(0.010)	0.017	(0.010)	0.008	(0.010)	-0.003	(0.010)
Gender (male = 1)	-0.053	(0.345)	-0.418	(0.351)	-0.064	(0.344)	0.311	(0.332)
Race (African American = 1)	-0.609*	(0.352)	0.298	(0.362)	0.541	(0.351)	1.544***	(0.372)
College degree	0.091	(0.274)	0.046	(0.274)	-0.014	(0.275)	0.155	(0.272)
Employed full time	-0.238	(0.309)	0.708**	(0.317)	-0.361	(0.308)	-0.597*	(0.315)
Taste	-0.056	(0.287)	0.059	(0.291)	-0.281	(0.291)	0.247	(0.288)
Price	1.053***	(0.300)	0.088	(0.289)	0.278	(0.287)	0.068	(0.286)
Zip code population	-0.000	(0.000)	-0.000	(0.000)	-0.000	(0.000)	0.000	(0.000)
Store distance	-0.023**	(0.011)	-0.028**	(0.011)	-0.020*	(0.011)	-0.011	(0.010)
Shop conv store	-0.128	(0.292)	0.037	(0.295)	0.090	(0.295)	-0.047	(0.290)
Shop dollar store	-0.584*	(0.331)	-0.028	(0.334)	-0.325	(0.328)	-0.324	(0.330)
Transportation	0.715*	(0.369)	0.605	(0.381)	0.227	(0.368)	0.625*	(0.374)
τ_1	-1.540*	(0.831)	0.500	(0.840)	-1.440*	(0.813)	0.014	(0.823)
τ_2	1.117	(0.826)	3.151***	(0.868)	1.238	(0.810)	2.245***	(0.835)
No. observations	222		222		222		222	
AIC	446.9		440.5		452.5		463.2	
Log likelihood	-209.4		-206.2		-212.2		-217.6	

Note: ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Appendix A2. Regression Results for Logit Model on the Frequency of Consumption for Vegetables. Salads, Fruits, and Fruit Juice, Imputed and Sample Using Weights

	Vegetables			**Salad**			**Fruit**			**Fruit juice**		
	Coeff.		**Std. Error**	**Coeff.**		**Std. Error**	**Coeff.**		**Std. Error**	**Coeff.**		**Std. Error**
Age	0.003		(0.013)	0.023		(0.016)	0.004		(0.012)	-0.018		(0.016)
Gender (male =1)	-0.021		(0.412)	-0.390		(0.420)	0.321		(0.446)	0.163		(0.419)
Race (African American =1)	-0.632	*	(0.339)	0.695		(0.483)	0.196		(0.454)	1.307	**	(0.531)
College degree	0.105		(0.319)	0.087		(0.344)	-0.092		(0.448)	0.334		(0.387)
Employed full time	-0.404		(0.400)	0.636		(0.502)	-0.701	*	(0.371)	-1.530	***	(0.451)
Taste	0.105		(0.390)	0.356		(0.436)	-0.339		(0.482)	-0.118		(0.493)
Price	0.934	**	(0.441)	0.317		(0.413)	0.307		(0.412)	-0.310		(0.452)
Zip code population	0.000		(0.000)	0.000		(0.000)	0.000		(0.000)	0.000		(0.000)
Store distance	-0.026	**	(0.013)	-0.032	**	(0.013)	-0.034	*	(0.019)	-0.019		(0.012)
Shop conv store	0.064		(0.431)	-0.063		(0.476)	0.404		(0.444)	-0.421		(0.528)
Shop dollar store	-0.165		(0.437)	0.328		(0.435)	0.178		(0.384)	0.393		(0.627)
Transportation	0.641		(0.446)	0.986	**	(0.383)	0.334		(0.546)	0.936	**	(0.400)
τ_1	-1.491		(0.717)	1.646		(1.048)	-1.799		(0.740)	-1.388		(1.117)
τ_2	1.061		(0.739)	3.890		(1.189)	0.756		(0.774)	0.870		(1.054)
No. observations	250			250			250			250		

Note: ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Journal of Food Distribution Research
Volume 55, Issue 2, pp. 46–64

Consumer Preference Regarding a New Corn Variety: A Willingness to Pay Study

Trevor D. Johnson[a]Ⓘ, Donna Mitchell-McCallister[b,c], and Wenwei Xu[d]

[a]*Graduate Research Assistant, Department of Agricultural and Applied Economics,
2500 Broadway, Texas Tech University,
Lubbock, TX 79409, USA*

[b]*Assistant Professor, Department of Agricultural and Applied Economics,
2500 Broadway, Texas Tech University,
Lubbock, TX 79409, USA*

[c]*Assistant Professor,
1102 East FM 1294, Texas A&M AgriLife Research
Lubbock, TX 79403, USA*

[d]*Regents Fellow and Professor of Corn Breeding and Genetics,
1102 East FM 1294, Texas A&M AgriLife Research,
Lubbock, TX 79403, USA*

Abstract

The purpose of this study is to analyze willingness to pay (WTP) and consumer preference for a red, Hi-A™ (high-antioxidant) corn variety. This paper used the double-bounded contingent valuation method and a binary logit model to analyze the responses of an online survey conducted in the fall of 2021. Survey results indicated that nearly 69% of respondents were willing to pay a premium for the new variety with an overall average WTP value of 81.40 cents per ear. This research highlights the economic implications of introducing nutrient-dense agricultural products to meet emerging consumer demand for healthier food alternatives..

Keywords: consumer preference; contingent valuation; corn; double-bounded; local food; willingness to pay

ⒾCorresponding author: Tel: (806) 742-2821
Email: trevor.d.johnson@ttu.edu

Introduction

Consumers have steadily exhibited an inclination toward healthier food alternatives in the last decade (Goetzke and Spiller, 2014; Martinez et al., 2018; Karpyn et al., 2020). This change can be attributed in part to economic and industrial disturbances in society and the food processing sector, causing disruptions in the food supply chain, leading companies to focus more on products that satisfy consumer demand for healthy alternatives (Bigliardi and Galati, 2013). Objective standards for what constitutes healthy foods are still unclear, but they are often categorized as foods with higher nutritional quality compared to alternatives (e.g., low sugar/calorie/saturated fat/sodium) (Motoki et al., 2021). A diet comprised of healthier foods is generally associated with a decreased risk of disease and an increase in overall well-being with consumers (Swinburn et al., 2015; Wahl et al., 2017). These risk factors have shifted individual preferences associated with food alternatives and provided an opportunity for new product entry. Heightened consumer awareness and a new focus on sustainability have also increased demand for healthy food alternatives (Grunert, 2006). These trends have helped decrease the intake of many negative nutrients but have not yielded a significant improvement in the overall diet of the American consumer (Miller et al., 2009)

Previous studies have indicated consumers with higher incomes have better access to healthy foods with relatively inelastic demand regarding changes in price, whereas lower income individuals resort to highly processed, cheaper alternatives (Andreyeva, Long, and Brownell, 2010; Chau, Zoellner, and Hill, 2013; Talukdar and Lindsey, 2013). Results from Feng and Chern (2000) reveal higher price elasticity for fresh fruits and vegetables, showing the importance of competitive pricing and understanding the average consumer's willingness to pay (WTP) for healthier alternatives. Price has shown to be a significant barrier to healthy food access, resulting in low-income individuals restricting their consumption (Jetter and Cassady, 2006; Steenhuis, Waterlander, and de Mul, 2011).

Among the primary drivers of consumer food choices has been product taste for the last several decades, prioritized far above healthiness (Verbeke, 2006; Aggarwal et al., 2016). Though taste is still a primary influence, health-focused labeling is devoid of this important attribute. Instead, health-focused labeling concentrates on nutritional benefits and verbal descriptions that mislead consumers to believe healthy alternatives taste worse and are less filling (Raghunathan, Naylor, and Hoyer, 2006; Suher, Raghunathan, and Hoyer, 2016). Often used as a signal for taste, the color of fresh produce has become an increasingly important factor in consumer decisions and consumption patterns. Many consumers associate divergent produce colors with nutritional benefits and the visual appeal of vibrant colors to good taste (Hein, 2023). For example, red, purple, and blue fruits may have high levels of antioxidants because they possess a subgroup of polyphenols called flavonoids, which includes anthocyanins (antioxidants). These factors have contributed to an increased demand for novel color selections among fruit and vegetable breeding firms and retailers looking to differentiate their product selection (Hein, 2023).

The increased demand for healthy food alternatives coincides with an increased consumer demand for locally sourced food products. There are a variety of reasons for consumers to have an increased desire to buy locally sourced products, including environmental concern, local economic support, land preservation, perceived nutritional benefit, etc. (Zepeda and Leviten-Reid, 2004; Groves, 2005). These factors proved not enough to change the underlying trends present in the market. Local food production and consumption have been reduced over time due to the consolidation in the U.S. agriculture market, reducing the prospects available to small farms (Stephenson and Lev, 2004). This trend has begun to subside in recent years as consumers convey a growing demand and preference for locally grown, fresh food to highly processed and traveled alternatives. Recent marketing studies have also explored these trends in support of local food (Jekanowski et al., 2000; Darby et al., 2006).

Many Americans associate sweet corn with fresh and local food because it is routinely sold in roadside stands or farmers' markets, is widely available as seeds for home gardeners to produce, and many consumers prefer to consume it uncooked and fresh rather than frozen or canned. Sweet corn is also routinely voted as the most popular vegetable in the United States and is one of the top 10 vegetables in terms of per capita consumption and market value (USDA-ERS, 2016). Sweet corn also possesses a range of minerals, vitamins, and resistant starches that can contribute to positive health-related outcomes (Sheng, Tong, and Liu, 2018). Despite these factors, the consumption of sweet corn is decreasing, as Americans are eating fewer vegetables overall, according to the USDA (Bentley, 2017). However, the introduction of innovative varieties aims to redefine the perception and consumption patterns of sweet corn.

Hi-A™ (high-antioxidant) sweet corn is a new variety currently being developed by selective breeding and field trials at the Texas A&M AgriLife Research facility in Lubbock, TX. The Hi-A™ has a pronounced red coloration, elevated levels of antioxidants similar to that of a blackberry, and is less sweet with a slightly tougher texture than a generic, yellow variety of sweet corn. As consumers shift their preferences toward healthier food alternatives and the consumption of generic sweet corn declines among Americans, there is an opportunity for the Hi-A™ variety to address these trends and potentially renew interest in sweet corn consumption. While many prior works have looked at consumer demand for healthy food alternatives, there is a lack of research related to consumer preference and WTP for specific enhanced nutritional attributes (e.g., elevated levels of antioxidants). Markosyan et al. (2009) found that consumers were willing to pay a premium for apples enriched with an antioxidant coating, especially when the health benefits of antioxidants were noted. While the research found small premiums for the average consumer, the additional antioxidants were from the wax coating rather than the produce itself. Additionally, Colson and Huffman (2011) found that consumers have positive valuations of enhanced levels of antioxidants and vitamin C, gained through genetic modification; however, the study focused solely on broccoli, tomatoes, and potatoes.

As consumers convey an increasing demand for healthy food alternatives and fresh produce with relatively elastic demand, there is an opportunity for novel food products to satisfy elevated demand and a need to evaluate the WTP of average consumers regarding new alternatives. Moreover, as taste is the primary driver of food choice, understanding the tradeoffs consumers make regarding product taste and enhanced nutritional benefits is essential in introducing new alternatives to satisfy changing consumer preferences (Aggarwal et al., 2016). Therefore, the objective of this study is to evaluate WTP and consumer preference for a new variety of sweet corn that exhibits a red color and has higher levels of antioxidants compared to other varieties currently in the marketplace. This research aims to provide justification for further development of the variety to make it more competitive and desirable among consumers for its potential entrance into the marketplace.

Data

The data for this research were collected through a nationwide, online survey distributed by Qualtrics. Screening questions included a minimum age requirement of 18, and the respondent had to be the main shopper for their household. The Texas Tech University Human Research Protection Program Institutional Review Board and Qualtrics both approved the survey before it was distributed to participants. The survey was first released in September of 2021, with a soft launch (n = 95) to confirm the effectiveness of the questions and survey flow regarding the different blocks of the double-bounded contingent valuation questions. Additional responses were collected through October 2021. In total, 1,052 responses were collected, and 1,037 were used in the study after omitting partial responses.

The survey was designed so that respondents would engage in two rounds of bidding based on the double-bounded contingent valuation method (CVM). The bid amounts used in the survey were based upon regional fresh sweet corn prices in the United States. The price from each region (Northeast, South, Midwest, West) was determined by averaging the price from the top five grocery stores in each region to give an average price of ≈ \$0.50 per ear nationwide. Bid amounts of \$0.40, \$0.60, \$0.80, \$1.00, and \$1.20 were constructed from this average price, and the lowest initial bid of \$0.40 was used to capture the lower bound of WTP estimates from consumers wary of new, novel-colored produce.

Respondents were randomly presented with one of four blocks for the double-bounded contingent valuation questions while completing the survey. The blocks were the same taste/texture (\$0.40 starting price), different taste/texture (\$0.40 starting price), same taste/texture (\$0.60 starting price), and different taste/texture (\$0.60 starting price). The starting price refers to the base price of the generic, yellow sweet corn used for the comparison. The different starting prices were used to better model variability in produce prices and to estimate the entire distribution of WTP values more accurately. This method also helped control for inflated WTP estimates by providing different values that respondents could use to gauge their choices. Taste and texture were also stated in the description to determine if the added nutritional benefits of the red Hi-A™ variety were enough to overcome the less sweet and tougher texture. For example, half of the respondents were presented with a description of the Hi-A™ variety as "less sweet and slightly tougher," and the other half received a description of "identical sweetness and texture" as the generic yellow variety. All respondents were informed that the Hi-A™ corn had elevated levels of antioxidants similar to those of a blackberry. Respondents were also presented with a cheap-talk script before the WTP questions to help reduce the hypothetical bias often observed in CVM studies, given that no currency is actually exchanged.[1] Figure 1 presents a graphical interpretation of the survey flow.

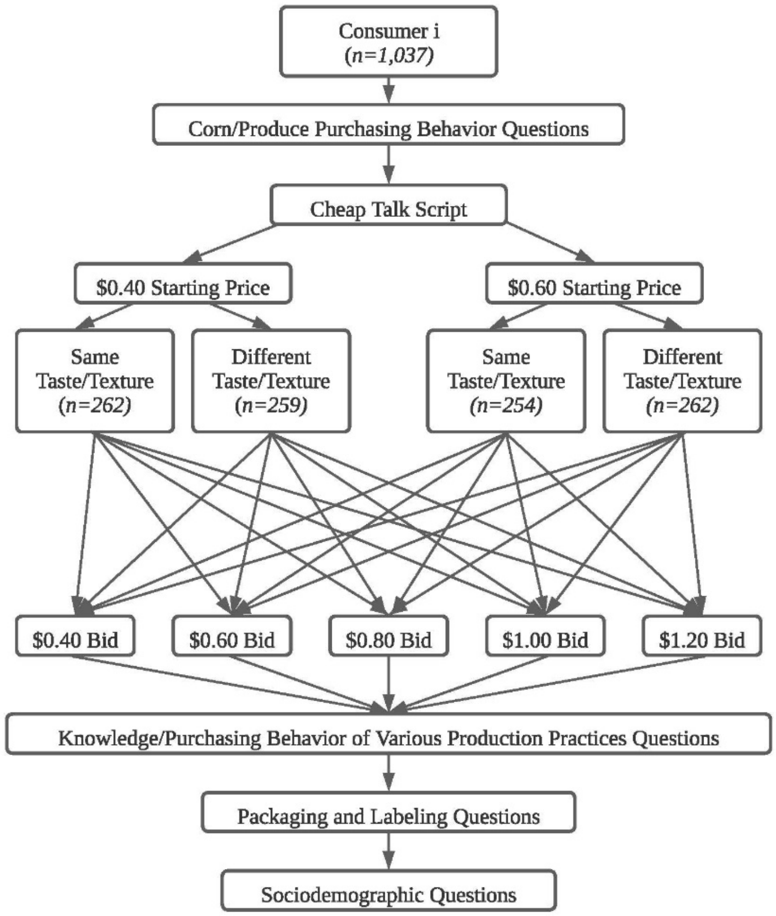


Figure 1. Survey Flow

In addition to the double-bounded contingent valuation questions, respondents were asked about their purchasing habits regarding fresh produce, preferences for packaging and labeling fresh corn, and sociodemographic questions.

Economic Framework/Methods

Contingent valuation is a method that uses nonmarket valuation to evaluate deviations from what is generally perceived to be "common." Respondents were asked to state their preference (i.e.,

[1]The cheap-talk script reminded respondents about their budget constraints and to make choices based upon their own preferences, asked them to make selections as if the choices were faced in an actual purchasing venue, and explained how previous research often found inflated consumer WTP values.

whether or not they will purchase a good if it costs x amount of dollars) regarding the alternatives that were presented. The double-bounded method for analyzing WTP in contingent valuation surveys has routinely been used to produce more accurate estimates than the single-bounded method (Hanemann, Loomis, and Kanninen, 1991). The efficiency of WTP estimates is improved by asking respondents to engage in two rounds of bidding (Hanemann, Loomis, and Kanninen; 1991; Kanninen, 1993; Riddel and Loomis, 1998). The follow-up bid, which is dependent on the response to the first bid, leads to asymptotically more efficient gains, improving upon the single-bounded approach and providing considerably improved statistical evidence from the response data (Hanemann, Loomis, and Kanninen, 1991). Also, the double-bounded approach allows each respondent's WTP to be placed in one of four choice categories with reduced, more statistically valuable intervals: "yes/yes," "yes/no," "no/yes," or "no/no" (Kanninen and Khawaja, 1995). For example, the WTP of participants who respond "yes" to an initial bid of \$0.60 and "no" to a follow-up bid of \$0.80 is narrowed down to the interval comprised of both the first and second bid amounts.

The following econometric interpretation was derived by López-Feldman (2012). y_i^1 and y_i^2 can be defined as the dichotomous variables that report the answers to the two close-ended questions (e.g., $y_i^1 = 1$ and $y_i^2 = 0$ if the responses to the first and second closed questions are "yes" and "no," respectively), where the probability that an individual responds yes to the initial question and no to the subsequent question can be expressed $Pr(y_i^1 = 1, y_i^2 = 0 \middle| z_i) = Pr(s, n)$, where *s* represents "yes" and *n* represents "no" (the conditionality of the probability on explanatory variables is removed for simplification). Respondent *i*'s WTP can be written as follows:

$$WTP_i(z_i, u_i) = z_i'\beta + u_i \text{ and } u_i \sim N(0, \sigma^2), \tag{1}$$

where z_i is a vector of explanatory variables, β is a vector of coefficients to be estimated, and u_i is the error term (López-Feldman, 2012). In this case, the z_i vector contains sociodemographic variables and additional control variables related to the purchasing habits of fresh produce and fresh corn, specifically.[2] Additionally, it is assumed that an individual will answer "yes" when their respective WTP exceeds some bid value (i.e., $WTP_i > t^n$). Using the previous assumptions, we have the probability for the first of the four cases given by:

$$\begin{aligned}
&y_i^1 = 1 \text{ and } y_i^2 = 0. \\
&Pr(s, n) = Pr(t^1 \leq WTP \leq t^2) \qquad (2) \\
&= Pr(t^1 \leq z_i'\beta + u_i < t^2) \\
&= Pr\left(\frac{t^1 - z_i'\beta}{\sigma} \leq \frac{u_i}{\sigma} < \frac{t^2 - z_i'\beta}{\sigma}\right) \\
&= \phi\left(\frac{t^2 - z_i'\beta}{\sigma}\right) - \phi\left(\frac{t^1 - z_i'\beta}{\sigma}\right),
\end{aligned}$$

[2]For a comprehensive list of explanatory variables included in z_i, please refer to the parameters included in Table 1.

where the last expression follows from $Pr(a \leq X < b) = F(b) - F(a)$. Therefore, using symmetry of the normal distribution we have that:

$$Pr(s,n) = \phi\left(z_i' \frac{\beta}{\sigma} - \frac{t^1}{\sigma}\right) - \phi\left(z_i' \frac{\beta}{\sigma} - \frac{t^2}{\sigma}\right). \tag{3}$$

The two outcomes when the respondent gives the same answer to both dichotomous choice questions (e.g., "yes/yes" or "no/no") do not correspond to a pre-existent model. Therefore, a likelihood function is constructed to directly estimate β and σ using maximum likelihood estimation (López-Feldman, 2012). The following likelihood function should be maximized to estimate the parameters for the model:

$$\sum_{i=1}^{N} [d_i^{sn} \ln\left(\phi\left(z_i' \frac{\beta}{\sigma} - \frac{t^1}{\sigma}\right) - \phi\left(z_i' \frac{\beta}{\sigma} - \frac{t^2}{\sigma}\right)\right) + d_i^{ss} \ln\left(\phi\left(z_i' \frac{\beta}{\sigma} - \frac{t^2}{\sigma}\right)\right) +$$
$$d_i^{ns} \ln\left(\phi\left(z_i{}' \frac{\beta}{\sigma} - \frac{t^2}{\sigma}\right) - \left(z_i{}' \frac{\beta}{\sigma} - \frac{t^1}{\sigma}\right)\right) + d_i^{nn} \ln\left(1 - \phi\left(z_i' \frac{\beta}{\sigma} - \frac{t^2}{\sigma}\right)\right)], \tag{4}$$

where $d_i^{sn}, d_i^{ss}, d_i^{ns}, d_i^{nn}$ are indicator variables equal to 1 or 0 for each individual case, which means that a unique individual contributes to the logarithm of the likelihood function in only one of the four parts (López-Feldman, 2012). This approach directly estimates $\widehat{\beta}$ and $\hat{\sigma}$, which is contrary to the single-bounded approach. Using STATA, the *doubleb* command directly estimates these parameters and allows for accurate WTP measures with or without control variables using the *nlcom* command.

A binary logit model was used to estimate consumer preference because the dependent variable has a finite number of possible outcomes that is equal to 2 (i.e., choosing either the Hi-A™ variety or the generic variety). Using the assumption that the error terms of the model are independently and identically distributed (iid) allows for simplification in estimation. Therefore, following Train's (2009) formulation, the probability of the person choosing the Hi-A™ variety is:

$$\begin{aligned} Pr &= \int I[\beta' x + \varepsilon > 0] f(\varepsilon) d\varepsilon \\ &= \int I[\varepsilon > -\beta' x] f(\varepsilon) d\varepsilon \\ &= \int_{\varepsilon=-\beta' x}^{\infty} f(\varepsilon) d\varepsilon \\ &= 1 - F(-\beta' x) = 1 - \frac{1}{1 + e^{\beta' x}} \\ &= \frac{e^{\beta' x}}{1 + e^{\beta' x}}, \end{aligned} \tag{5}$$

where $f(\cdot)$ is the density of ε, and assuming ε is distributed logistically where its density is $f(\varepsilon) = e^{-\varepsilon}/(1 + e^{-\varepsilon})^2$ and the cumulative distribution is $F(\varepsilon) = 1/(1 + e^{-\varepsilon})$. Using the above estimation, for any x, the probability can be calculated as $P = \exp(\beta' x)/(1 + \exp(\beta' x))$. The *logit* and *logistic* commands in STATA were used to estimate the coefficients and odds-ratios of the logistic regression, respectively. The difference between the model used to estimate WTP and the binary logit model is in the dependent variable, which is changed to *Stated_Red* for the logistic regression. This is a dummy variable where the respondents were asked to state their preference for purchasing either the Hi-A^{TM} variety or a generic sweet corn variety if they were both equally priced.

Results

In the current study, we are interested in consumer preference and average WTP regarding the Hi-A^{TM} corn variety. Different methods are available for the estimation of WTP values. For example, WTP can be estimated for certain portions of the overall sample, for certain respondents in the sample possessing specific characteristics, or using average values of control variables to construct an overall mean WTP value. For the purposes of this research, the primary focus will be on the latter.

The summary of statistics of the survey respondents' sociodemographic characteristics and additional variables used in the analysis are presented in Table 1. The average age of the respondents was 46.72 years old with an average household size of between two and three people. Average household income was found to be $56,558, with 36.74% of the respondents having a college-level education, and 76.28% of the respondents were female. Our sample is older, has a lower income, is slightly less educated, and has a higher number females as a percentage compared to the general population (U.S. Census Bureau, 2020). The data are skewed toward female respondents, which is consistent with prior research looking at WTP for healthier food products where main household shoppers were most commonly found to be female (Alsubhi et al., 2023).

Table 1. Summary Statistics for the Survey Respondents

Variable	Description	Percentage of Occurrence	Mean	Standard Deviation
Age	Age of the consumer:		46.7195	16.7122
	1 = 18–30	20.73%		
	2 = 31–45	30.67%		
	3 = 46–60	22.47%		
	4 = > 60	26.13%		
Gender	Dummy variable:		0.7628	0.4256
	0 = Male	23.72%		
	1 = Female	76.28%		
Household size	Number of people living in the household		2.54	1.2309
	1 = 1	21.31%		
	2 = 2	35.58%		
	3 = 3	20.64%		
	4 = 4	12.73%		
	5 = > 4	9.74%		

Table 1 (cont.)

Variable	Description	Percentage of Occurrence	Mean	Standard Deviation
College educated	Dummy variable:		0.3674	0.4823
	0 = No	63.26%		
	1 = Yes	36.74%		
Income	Pre-tax level of household income:		$56,558	39,410
	1 = < $25,000	22.71%		
	2 = $25,000–$50,000	29.86%		
	3 = $50,001–$75,000	21.84%		
	4 = $75,001–$100,000	11.11%		
	5 = $100,001–$125,000	6.67%		
	6 = $125,000–$150,000	4.25%		
	7 = > $150,000	3.57%		
Venue	Where the consumer most frequently purchases fresh produce:		2.2324	0.7976
	1 = Farmers' market	9.93%		
	2 = Large grocery chain	66.54%		
	3 = Small, local grocery store	17.16%		
	4 = Health food store	3.09%		
	5 = Wholesale club store	3.28%		
Color	If the consumer would purchase novel-colored corn for additional health benefits:		0.81	0.3925
	0 = No	19.00%		
	1 = Yes	81.00%		
Local label effect	If a locally produced label would increase the likelihood of purchasing the Hi-A™ variety:		1.0415	0.6672
	Dummy variables:			
	1 = Yes	55.35%		
	1 = No	20.25%		
	1 = No change	24.40%		
Taste	If the taste was described as similar to a generic variety:		0.4976	0.5002
	0 = No	50.24%		
	1 = Yes	49.76%		

Table 1 (cont.)

Variable	Description	Percentage of Occurrence	Mean	Standard Deviation
Nutrition/health	The level of importance consumers place on nutrition/health benefits when purchasing fresh produce:		2.4291	0.5960
	1 = Low	5.50%		
	2 = Medium	46.09%		
	3 = High	48.41%		
Red/black	If a red or black color indicates higher levels of antioxidants, how likely would this affect purchasing habits regarding corn:		1.9826	0.6861
	1 = Not Likely	24.40%		
	2 = Somewhat Likely	52.94%		
	3 = Very Likely	22.66%		
Local purchasing	How often the consumer seeks out locally-produced products:		3.0087	0.8342
	0 = Never	3.38%		
	1 = Not sure	1.93%		
	2 = Not very often	18.42%		
	3 = Somewhat often	46.38%		
	4 = Very often	29.89%		
Social responsibility	If the consumer feels a responsibility to seek out locally-produced products to support local producers and their community:		0.7338	0.4422
	0 = No	26.62%		
	1 = Yes	73.38%		

Additional variables that were incorporated in the subsequent models include the following: *Nutrition/Health*—importance the consumer places on nutritional/health benefits when purchasing fresh produce; *Red/Black*—the effect on purchasing behavior when a red or black color indicates higher levels of antioxidants; *Local Purchasing*—how often the consumer seeks out locally produced products; and *Social Responsibility*—perception of social responsibility regarding local economic support.

Table 2 presents the overall responses to each combination of bid levels (i.e., how the participants responded to each combination of bid amounts). Results show that 68.85% of respondents were willing to pay a premium value for the Hi-A^{TM} corn variety, whereas 31.15% were not willing to pay any level of premium. The percentage of respondents who said yes to the first bid and no to the second bid was 15.53%, whereas 17.07% of respondents said no to the first and yes to the second. It is important to note that the percentage of Yes–Yes responses generally decreased as the bid amounts increased. This choice combination had the highest percentage of respondents at 32.26%.

Table 2. WTP Premium Distributions

First Bid	Yes–Yes	Yes–No	No–Yes	No–No
$0.40	10.32%	3.95%	1.35%	4.15%
$0.60	8.20%	4.73%	3.76%	3.09%
$0.80	6.65%	3.86%	4.82%	5.50%
$1.00	4.63%	2.03%	4.34%	8.58%
$1.20	6.46%	0.96%	2.80%	9.84%
% Total	36.26%	15.53%	17.07%	31.15%

Willingness to Pay

Table 3 presents the results of the double-bounded WTP model. Because z_i from equation 1 is simply a vector of explanatory variables, the coefficients of each variable can be interpreted as the direct impact on WTP for each control variable on a per ear basis. The constant in the regression can represent a base price per ear (35.18 cents) that consumers are willing to pay. It is important to note that all reported WTP values are on a per ear basis. A mean WTP for the Hi-A™ variety was calculated using the results of the regression and is equal to 81.40 cents with upper and lower bounds of 77.52 cents and 85.27 cents based on the 95% confidence interval, respectively. Using the base prices of the normal variety for comparison, this average value is equal to a 41.40 cent premium when the normal variety is priced at $0.40, or a 21.40 cent premium when the price of the normal variety is $0.60.

Table 3. WTP Estimates for the Red, Hi-A™ Sweet Corn

Parameter	Coefficient	Std. Err.	Pr > z
Constant	0.3518***	0.1249	0.0050
Age	-0.1131***	0.0174	0.0000
Gender	-0.0648	0.0409	0.1140
Household size	0.0176	0.0150	0.2430
College education	-0.0152	0.0391	0.6980
Income	0.0015	0.0125	0.9040
Venue			
Farmers' market	0.0950	0.0692	0.1700
Large grocery chain	0.0369	0.0458	0.4210
Health food store	0.2503**	0.1139	0.0280
Wholesale club store	0.1679*	0.1017	0.0990
Nutritional/health			
Benefits	0.0011	0.0306	0.9720
Color	0.3758***	0.0513	0.0000
Red/black	0.0339	0.0421	0.4210

Table 3 (cont.)

Parameter	Coefficient	Std. Err.	Pr > z
Local label effect:			
Yes	0.2849***	0.0437	0.0000
No	-0.1032*	0.0538	0.0550
Local purchasing	0.0463**	0.0233	0.0470
Social responsibility	0.1016**	0.0438	0.0200
Taste	0.0594*	0.0345	0.0850
Log likelihood	-1186.07		

Note: The variables are described in Table 1 and in above discussions.
*, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

At the 1% level, the constant and variables *Age, Color,* and *Local Label* are statistically significant. The age of the consumer decreases WTP by 11.31 cents for each increase in the age category. Regarding the color of the corn, the estimated results show that consumers who are willing to purchase sweet corn that is not yellow for additional health benefits are willing to pay an additional 37.58 cents. Additionally, if the consumer indicated that a locally produced label would increase their likelihood of purchasing the red, Hi-A^{TM} (*Local Label Effect: Yes),* their WTP increases by 28.49 cents. This amount is compared to the base value of a locally produced label having no effect on likelihood of purchase. On the other hand, if a locally produced label would not increase likelihood of purchase, then consumer WTP decreases by 10.31 cents, as indicated by *Local Label Effect: No,* which is statistically significant at the 10% level. Further support for locally produced foods is shown by *Local Purchasing,* which is statistically significant at the 5% level. That is, the more often a consumer seeks out products labeled as "locally produced," the more often they are willing to pay an additional 4.63 cents for each increase on the Likert scale. *Social Responsibility* is also statistically significant at the 5% level. Consumers who felt it is their social responsibility to seek out locally produced foods in order to support their local producers and economy are willing to pay 10.16 cents more per ear for the Hi-A^{TM} variety.

Because the new variety is not yet available for purchase, it is important to determine at which purchasing venue consumers are willing to pay the highest level of premium. Five different venues were considered in the model, and a small local grocery store was used as the base for comparison. Of the venues considered, *Health Food Store* and *Wholesale Club Store* were statistically significant at the 5% and 10% levels, respectively. Consumers who most frequently purchase their fresh produce at a health food store or wholesale club store, as compared to a small local grocery store, were willing to pay an additional 25.03 cents and 16.79 cents for the new variety, respectively.

Taste is the primary driver of consumer food choices, so it was important to determine if the less sweet and tougher texture of the Hi-A^{TM} altered WTP estimates. In order to do so, the variable *Taste* was considered in the analysis. It is a constructed dummy variable that is equal to 1 if the respondent received a description of the new variety, stating that it was similar to a generic sweet corn variety in both taste and texture, and equal to 0 otherwise.[3] *Taste* is statistically significant at the 10% level, showing that consumers are willing to pay 5.94 cents more for the Hi-A™ if the taste and texture are similar to a generic variety of sweet corn. This result is consistent with previous studies showing that consumers are driven by product taste, which is generally prioritized above healthiness (Verbeke, 2006; Aggarwal et al., 2016).

Consumer Preference

In order to analyze consumer preference toward the Hi-A™ corn variety, a binary logit model was utilized with the same components as the WTP model. The difference between the two models is in the dependent variable, which is changed to *Stated_Red* for the logistic regression. This is a dummy variable that indicates whether the respondent prefers the Hi-A™ variety to the normal variety when they are equally priced. The logit model results show which characteristics and preferences of the consumer increase the likelihood of purchasing the Hi-A™ variety. The logistic regression is specifically focused on consumers' stated preference toward a new variety of sweet corn with a unique color and additional health benefits.

Table 4 shows the results of the logistic regression. The estimated logit model results show that five variables are statistically significant, with *Color* having the largest effect on stated preference. Consumers who are willing to purchase sweet corn varieties that are not yellow for additional health benefits are 6.21 times more likely to purchase the Hi-A™ variety. Similarly, consumers who place a high importance (*Nutrition/Health* = 3*)* on the nutritional or health benefits of their fresh produce are 4.69 (3 × 1.5633) times more likely to prefer the new variety. Furthermore, if the Hi-A™ had the same level of sweetness as a generic sweet corn variety, the consumer was 1.74 times more likely to have stated that he or she prefers the Hi-A™ variety. Compared to the base of no effect, consumers who would be positively affected by a locally produced label are 1.75 times more likely to have a stated preference for the new variety.

Table 4. Stated Preference Regression for Red, Hi-A™ Sweet Corn

Parameter	Coefficient	Odds Ratio	Std. Err.	Pr > z
Constant	-4.1382***	0.0159	0.0099	0.0000
Age	-0.1359*	0.8729	0.0646	0.0660
Gender	-0.0638	0.9382	0.1630	0.7140
Household size	-0.0793	0.9237	0.0607	0.2270
College education	0.1278	1.1363	0.1904	0.4460
Income	0.0209	1.0211	0.0537	0.6920
Venue				
Farmers' market	-0.3362	0.7145	0.2154	0.2650
Large grocery chain	0.2047	1.2272	0.2446	0.3040
Health food store	0.0353	1.0359	0.4443	0.9340

[3]As pointed out by a reviewer, the Hi-A™ differs in both taste and texture, but texture is not explicitly controlled for in the regression. However, the variable *Taste* also controls for texture given the descriptions provided to respondents where taste and texture changed concurrently.

Table 4 (cont.)

Parameter	Coefficient	Odds Ratio	Std. Err.	Pr > z
Wholesale club store	-0.3087	0.7344	0.3371	0.5010
Nutritional/health	0.4468***	1.5633	0.2148	0.0010
Color	1.8257***	6.2073	2.0039	0.0000
Red/black	0.0577	1.0594	0.2048	0.7650
Local label effect:				
Yes	0.5609**	1.7522	0.3230	0.0020
No	-0.8989**	0.4070	0.1197	0.0020
Local purchasing	0.1446	1.1556	0.1249	0.1810
Social responsibility	0.1345	1.1440	0.2320	0.5070
Taste	0.5514***	1.7356	0.2572	0.0000
Log likelihood	-551.3249			

Note: The variables are described in Table 1 and in above discussions.
*, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

In contrast, for each increase in their age category the likelihood of the consumer having a stated preference for the Hi-A^{TM} variety decreases by a factor of 0.87, and those who would not change their purchasing habits based on a locally produced label are less likely to prefer the new variety.

Discussion

A significant result from this study is that consumers who are willing to purchase novel colors of sweet corn for added health benefits are willing to pay 37.58 cents more per ear for the new variety. Additionally, consumers who would respond positively to a locally produced label on the new variety are willing to pay an additional 28.49 cents. These same characteristics are also the two most important factors in predicting stated preference for the Hi-A^{TM} variety. Moreover, consumers who regularly seek out locally produced foods are more willing to pay a premium and have a stated preference for the new variety. These findings align with previous research indicating that consumers value both local sourcing and enhanced health attributes in their food choices (Zepeda and Leviten-Reid, 2004; Colson and Huffman, 2011).

It is also interesting to note that when using a local grocery store as a baseline comparison, consumers who most frequently purchase their fresh produce at health food stores and wholesale club stores are willing to pay an additional 25.03 and 16.79 cents for the Hi-A^{TM} corn, respectively. This result underscores the importance of understanding distribution channels in influencing consumer behavior, especially as the number of large grocery stores and chain stores that can handle more product variety has increased in the United States (Jekanowski and Binkley, 2000; Cho and Volpe, 2017). The results used to calculate mean WTP in this study are generated by using this entire specific sample, and the results can be generalized to a degree to estimate WTP values for specific segments of consumers. For example, the specific results above show that the older the consumer is, the less likely they are willing to pay a premium or have a stated preference for the Hi-A™ variety, *ceteris paribus.* Therefore, different venues that may hypothetically sell this product could target younger consumers and use other data from the results to determine specific price levels to market toward various consumers.

Conclusions

Results indicate that, on average, consumers are willing to pay 81.40 cents for the Hi-A™ corn variety. This result is notable considering there are currently no other varieties of sweet corn in the marketplace that possess this particular combination of color and health benefits. Given that consumers may use color as a proxy for both taste and baseline levels of nutrition, the introduction of a red variety of sweet corn could have high economic value (Hein, 2023). This possibility is especially notable considering that nearly 70% of respondents were willing to pay some level of premium for the Hi-A™ variety.

The models used in this research help identify which factors have the largest influence on both consumers' WTP and stated preference for the Hi-A™ variety. In the double-bounded model, results showed that WTP was affected by the location where the consumer most frequently purchases fresh produce, how often they seek out products labeled as locally produced, and whether or not they feel a social responsibility to support local economies and producers. For the logistic regression, stated preference for the new variety was positively affected by the level of importance consumers place on the nutritional/health benefits of fresh produce. The age of the consumer had a significant negative effect on both WTP and stated preference for the new variety. If consumers were willing to purchase sweet corn that is not yellow for added health benefits, they had a higher WTP and were more likely to prefer the Hi-A™ variety. Similarly, those who would be positively affected by a locally produced label were willing to pay more and were more likely to prefer the Hi-A™ variety.

Taste was also an important factor in determining WTP and stated preference. When the Hi-A™ variety was described as having the same level of sweetness and an identical texture as a generic sweet corn variety while simultaneously having high levels of antioxidants, consumers were more likely to prefer the new variety and were willing to pay more for it. This finding is consistent with previous research showing that taste is a primary driver of food choices, and consumers value taste above healthiness (Verbeke, 2006; Aggarwal et al., 2016). This result provides additional justification for further development of the Hi-A™ variety to improve taste/texture and make it more competitive for hypothetical, future marketing purposes.

The research successfully identified that many consumers are willing to pay for and purchase a new, nutrient dense, and uniquely colored sweet corn variety and identified how they differ among specific characteristics. Future research should focus on the tradeoffs consumers make between product taste and the nutritional qualities they possess. Specifically, at what point are consumers not willing to sacrifice taste any longer for additional health benefits. Other research should focus on analyzing WTP and consumer preference for other types of novel-colored produce to determine whether the results of this study are reproducible for other fruits, vegetables, grains, etc.

The main limitation of the study is the use of contingent valuation in an online survey to collect the data for analysis. There are issues with hypothetical bias when using the CVM because there is not actually any money being transacted. Additionally, the sample data are not well balanced considering that around 75% of the respondents were female, and there are numerous opportunities to improve estimation in consumer-based experiments to help reduce bias in WTP studies. Future research could aim to have a more equal proportion of both male and female respondents to better reflect the population as a whole. Finally, although there may be increased demand for novel produce colors, it is still part of a niche market with lower market value compared to overall total market value.

Competing Interests

Authors declare no competing interest.

Data Availability Statement

For replication purposes, the data that support the findings of this study are available from the corresponding author, T.D.J, upon reasonable request.

Funding Statement

This work was supported by the High Plains Underground Water Conservation District #1.

References

Aggarwal, A., C.D. Rehm, P. Monsivais, and A. Drewnowski. 2016. "Importance of Taste, Nutrition, Cost and Convenience in Relation to Diet Quality: Evidence of Nutrition Resilience among US Adults Using National Health and Nutrition Examination Survey (NHANES) 2007–2010." *Preventive Medicine* 90:184–192.

Alsubhi, M., M. Blake, T. Nguyen, I. Majmudar, M. Moodie, and J. Ananthapavan. 2023. "Consumer Willingness to Pay for Healthier Food Products: A Systematic Review." *Obesity Reviews* 24(1):p.e13525.

Andreyeva, T., M.W. Long, and K.D. Brownell. 2010. "The Impact of Food Prices on Consumption: A Systematic Review of Research on the Price Elasticity of Demand for Food." *American Journal of Public Health* 100(2):216–222.

Bentley, J. 2017. *U.S. Trends in Food Availability and a Dietary Assessment of Loss-Adjusted Food Availability, 1970–2014*. Washington, DC: U.S. Department of Agriculture, Economic Research Service, Economic Information Bulletin EIB-166.

Bigliardi, B., and F. Galati. 2013. "Innovation Trends in the Food Industry: The Case of Functional Foods." *Trends in Food Science and Technology* 31(2):118–129.

Chau, C.N., J.M. Zoellner, and J.L. Hill. 2013. "Availability of Healthy Food: Does Block Group Race and Income Matter?" *Journal of Hunger and Environmental Nutrition* 8(1):22–38.

Cho, C., and R. Volpe. 2017. *Independent Grocery Stores in the Changing Landscape of the US Food Retail Industry.* Washington, DC: U.S. Department of Agriculture, Economic Research Service, Economic Research Report No. ERR-240.

Darby, K., M.T. Batte, S. Ernst, and B. Roe. 2006. *Willingness to Pay for Locally Produced Foods: A Customer Intercept Study of Direct Market and Grocery Store Shoppers*. Available online: https://ageconsearch.umn.edu/record/21336/ [Accessed December 14, 2022].

Feng, X., and W.S. Chern. 2000. *Demand for Healthy Food in the United States*. Available online: https://ageconsearch.umn.edu/record/21857 [Accessed December 14, 2022].

Goetzke, B.I., and A. Spiller. 2014. "Health-improving Lifestyles of Organic and Functional Food Consumers." *British Food Journal* 116(3):510–526.

Groves, A.M. 2005. *The Local and Regional Food Opportunity*. Watford, England: Institute of Grocery Distribution.

Grunert, K.G. 2006. "How Changes in Consumer Behavior and Retailing Affect Competence Requirements for Food Producers and Processors." *Economía Agraria y Recursos Naturales* 6(11):3.

Hanemann, M., J. Loomis, and B. Kanninen. 1991. "Statistical Efficiency of Double-Bounded Dichotomous Choice Contingent Valuation." *American Journal of Agricultural Economics* 73(4):1255–1263.

Hein, T. 2023. "Introducing the Rainbow to Vegetable Breeding." *SeedWorld* 51–54.

Jekanowski, M.D., and J.K. Binkley. 2000. Food Purchase Diversity across US Markets. *Agribusiness: An International Journal* 16(4):417–433.

Jekanowski, M.D., D.R. Williams, and W.A. Schiek. 2000. "Consumers' Willingness to Purchase Locally Produced Agricultural Products: An Analysis of an Indiana Survey." *Agricultural and Resource Economics Review* 29(1):43–53.

Jetter, K.M., and D.L. Cassady. 2006. "The Availability and Cost of Healthier Food Alternatives." *American Journal of Preventive Medicine* 30(1):38–44.

Kanninen, B.J. 1993. "Optimal Experimental Design for Double-Bounded Dichotomous Choice Contingent Valuation." *Land Economics* 138–146.

Kanninen, B.J., and M.S. Khawaja. 1995. "Measuring Goodness of Fit for the Double-Bounded Logit Model." *American Journal of Agricultural Economics* 77(4):885–890.

Karpyn, A., K. McCallops, H. Wolgast, and K. Glanz. 2020. "Improving Consumption and Purchases of Healthier Foods in Retail Environments: A Systematic Review." *International Journal of Environmental Research and Public Health* 17(20):7524.

López-Feldman, A. 2012. "Introduction to Contingent Valuation using Stata." Munich, Germany: Munich Personal RePEc Archive, Paper No. 41018. Available online: https://mpra.ub.uni-muenchen.de/41018/2/MPRA_paper_41018.pdf.

Martinez, O., N. Rodriguez, A. Mercurio, M. Bragg, and B. Elbel. 2018. "Supermarket Retailers' Perspectives on Healthy Food Retail Strategies: In-Depth Interviews." *BMC Public Health* 18(1).

Miller, G.D., A. Drewnowski, V. Fulgoni, R.P. Heaney, J. King, and E. Kennedy. 2009. "It Is Time for a Positive Approach to Dietary Guidance Using Nutrient Density as a Basic Principle." *The Journal of Nutrition* 139(6):1198–1202.

Motoki, K., J. Park, A. Pathak, and C. Spence. 2021. "Constructing Healthy Food Names: On the Sound Symbolism of Healthy Food." *Food Quality and Preference* 90:104157.

Raghunathan, R., R.W. Naylor, and W.D. Hoyer. 2006. "The Unhealthy = Tasty Intuition and Its Effects on Taste Inferences, Enjoyment, and Choice of Food Products." *Journal of Marketing* 70(4):170–184.

Riddel, M., and J. Loomis. 1998. "Joint Estimation of Multiple CVM Scenarios under a Double Bounded Questioning Format." *Environmental and Resource Economics* 12:77–98.

Sheng, S., L. Tong, and R. Liu. 2018. "Corn Phytochemicals and Their Health Benefits." *Food Science and Human Wellness* 7(3):185–195.

Steenhuis, I.H.M., W.E. Waterlander, and A. de Mul. 2011. "Consumer Food Choices: The Role of Price and Pricing Strategies." *Public Health Nutrition* 14(12):2220–2226.

Stephenson, G., and L. Lev. 2004. "Common Support for Local Agriculture in Two Contrasting Oregon Communities." *Renewable Agriculture and Food Systems* 19(4):210–217.

Suher, J., R. Raghunathan, and W.D. Hoyer. 2016. "Eating Healthy or Feeling Empty? How the 'Healthy 5 Less Filling' Intuition Influences Satiety." *Journal of the Association for Consumer Research* 1(1):26–40.

Swinburn, B., V. Kraak, H. Rutter, S. Vandevijvere, T. Lobstein, G. Sacks, F. Gomes, T. Marsh, and R. Magnusson. 2015. "Strengthening of Accountability Systems to Create Healthy Food Environments and Reduce Global Obesity." *The Lancet* 385(9986):2534–2545.

Talukdar, D., and C. Lindsey. 2013. "To Buy or Not To Buy: Consumers' Demand Response Patterns for Healthy Versus Unhealthy Food." *Journal of Marketing* 77(2):124–138.

Train, K.E. 2009. *Discrete Choice Methods with Simulation*. Cambridge, United Kingdom: Cambridge University Press.

U.S. Census Bureau. 2020. "2020 Census Demographic Profile QuickFacts." Available online: https://www.census.gov/data/tables/2023/dec/2020-census-demographic-profile.html.

U.S. Department of Agriculture. 2016. *Potatoes, Tomatoes, and Sweet Corn Remain Most Popular Vegetables Among U.S. Consumers*. Washington, DC: U.S. Department of Agriculture, Economic Research Service.

Verbeke, W. 2006. "Functional Foods: Consumer Willingness to Compromise on Taste for Health?" *Food Quality and Preference* 17(1–2):126–131.

Wahl, D.R., K. Villinger, L.M. König, K. Ziesemer, H.T. Schupp, and B. Renner. 2017. "Healthy Food Choices Are Happy Food Choices: Evidence From a Real Life Sample Using Smartphone Based Assessments." *Scientific Reports* 7(17609):1–8. Available online: https://www.nature.com/articles/s41598-017-17262-9.

Zepeda, L., and C. Leviten-Reid. 2004. "Consumers' Views on Local Food." *Journal of Food Distribution Research* 35(3):1–6.

Journal of Food Distribution Research
Volume 55, Issue 2, pp. 65–95

Online Food Ordering and Delivery: A Study on the Use of Customer Service Data and Quality Function Deployment

Bo Peng[a]Ⓣ, Murat Erkoc[b], and V.K. Omachonu[c]

[a]*PhD Candidate, Department of Industrial and Systems Engineering, 1251 Memorial Drive, University of Miami, Coral Gables, Florida 33146, USA*

[b]*Associate Professor, Department of Industrial and Systems Engineering, 1251 Memorial Drive, University of Miami, Coral Gables, Florida 33146, USA*

[c]*Professor, Department of Industrial and Systems Engineering, 1251 Memorial Drive, University of Miami, Coral Gables, Florida 33146, USA*

Abstract

This study addresses the challenge of measuring the intricate nature of service design in the context of online food ordering and delivery. Despite a plethora of service industry studies, a comprehensive approach to understanding customer experience and perceptions is lacking. Leveraging e-commerce innovations, we introduce a service blueprint for the online food delivery industry. Through data collection, surveys, and statistical tools, key factors influencing the business are identified. Utilizing machine learning, our methodology aids decision makers in aligning services with customer needs. A Quality Function Deployment table is proposed to translate these insights into service design imperatives for the decision makers.

Keywords: service design, online food delivery services, customer experience, quality function deployment

ⓉCorresponding author: Tel: (786) 246-3644
Email: BPeng@ihv.umaryland.edu

Introduction

The online food ordering and delivery (OFD) industry, a vibrant and rapidly evolving sector, has become a crucial component of the modern service economy. Its growth, driven by technological advancements and changing consumer behaviors, especially in the wake of the COVID-19 pandemic, has brought new challenges and opportunities for service design (Donthu and Gustafsson, 2020). This industry, which intertwines complex logistics, customer interactions, and digital platforms, exemplifies the need for innovative service design approaches.

Service design is a method that orchestrates several service elements (e.g., physical environment, materials, and employees) in order to achieve the desired customer experiences. In the context of OFD, service design involves the orchestration of multiple components: digital interfaces, logistics and delivery, customer support, and the culinary experience itself. These components must be seamlessly integrated to deliver the intended service outcome. The concept of service design, especially in digital and e-commerce platforms, has garnered significant attention in recent years. Contemporary service design methodologies, building upon foundational work in service blueprinting and modelling, are now crucial for creating customer-centric experiences in digital-first businesses (Kunneman, Alves da Motta-Filho, and van der Waa, 2022; Iriarte et al., 2023).

The complex and abstract nature of OFD service has made service design an onerous and challenging task that is usually hard to measure. Recent literature underscores the complexity inherent in designing services that cater to the dynamic needs of this industry. For example, Jun et al. (2021) highlight the critical role of technology in enhancing customer experiences in food delivery, emphasizing the need for user-friendly digital platforms. In terms of logistical efficiency and reliability, a study by Lin et al. (2023) indicates the significant impact of food delivery speed and accuracy on customer satisfaction. Moreover, the integration of customer feedback into service improvement has become increasingly prominent, as noted by Holmlund et al. (2020), where the authors emphasize the use of data analytics for understanding and responding to customer preferences and behaviors. This need for a holistic understanding of customer experiences in the food delivery sector is echoed by Noyes et al. (2019), who argue for a more comprehensive approach, combining qualitative insights with quantitative data analysis.

Our research aims to bridge this gap in the context of the online food ordering and delivery industry. We propose a novel framework that leverages the power of machine learning and Quality Function Deployment (QFD) to dissect and reconstruct the customer experience. Employing this framework is not only innovative but also necessary in the current landscape, where the fast-paced nature of the food delivery industry demands a more agile and data-driven response to service design challenges. Overall, our study addresses two critical research questions: "What are the key customer requirements that ensure their satisfaction with online food delivery services and their propensity to endorse these services to others?" and "How can the concept of service design and machine learning be applied to identify these requirements?" To tackle these questions, we adopt a customer-centric approach by designing and implementing a survey informed by the service blueprint framework. Subsequently, we develop a holistic analysis, powered by advanced machine learning algorithms, which reveals the core elements that shape customer satisfaction.

The proposed methodology ultimately results in a Quality Function Deployment (QFD) table, constructed from our analysis and predictive results. This QFD table is not only a theoretical framework, but is also a practical tool for businesses to align their services with real customer needs. As noted by industry leader Jack Ma, founder of Alibaba, “I'm not a tech guy. I'm looking at the technology with the eyes of my customers, normal people’s eyes.” By implementing these insights into service design, companies are poised to deliver a significantly enhanced customer experience. This study, therefore, not only contributes to academic discourse but also offers tangible strategies for businesses striving to excel in the competitive realm of OFD, a sector where customer satisfaction is paramount and directly linked to business success.

Research Background

The advent of rapid internet and smartphone penetration in shopping practices has catalyzed a transformation in the courier and delivery landscape, ushering in the era of online food delivery—a market with a user base that exceeded 1 billion globally by the end of 2019 (Business Wire, 2020). The projected trajectory suggests a global revenue growth to 1.39 trillion USD by 2025, marking a significant upturn from 0.36 trillion USD in 2019 (Al Amin et al., 2021). This growth reflects a shift in consumer behavior toward convenience-driven services, a trend that has become more pronounced in the wake of the COVID-19 pandemic.

As the pandemic redefined social norms, the food service industry grappled with unprecedented challenges. The CDC’s guidelines recommended takeout and delivery as the safest options for food service, encouraging restaurants to pivot swiftly to these models (Centers for Disease Control and Prevention, 2020). This shift was not only a response to immediate health concerns, but was also a strategic move to align the business practice with evolving consumer expectations. Research indicates that convenience (Rathore and Chaudhary, 2018), transactional ease (Natarajan, Gupta, and Nanda, 2019), and a broad spectrum of choices (Tandon et. al, 2021; Bir et. al, 2023) are the primary motivators for consumers opting for OFD—a service that has seen a sharp rise in engagement post-pandemic. Gunden, Morosan, and DeFranco (2020) examined a wide variety of factors that motivate consumers to use OFD systems in the United States using a conceptual model. The authors conclude that performance expectancy was the strongest predictor of intentions to use OFD systems, followed by congruity with self-image.

A comprehensive survey conducted in the United Kingdom in March 2020 provides valuable insights into this behavioral shift. As illustrated in Figure 1, a significant 60% of respondents aged 18–34 reported an increase in OFD, with a substantial proportion planning further increase. The trend persists across older demographics, indicating a widespread adoption of online food ordering in food delivery services (Statista, 2020). These data underpin the need for adaptive service design in the OFD industry to cater to a diverse customer base with heightened expectations. As the food delivery market continues to expand, the industry must adapt to these patterns to maintain customer satisfaction and business growth.

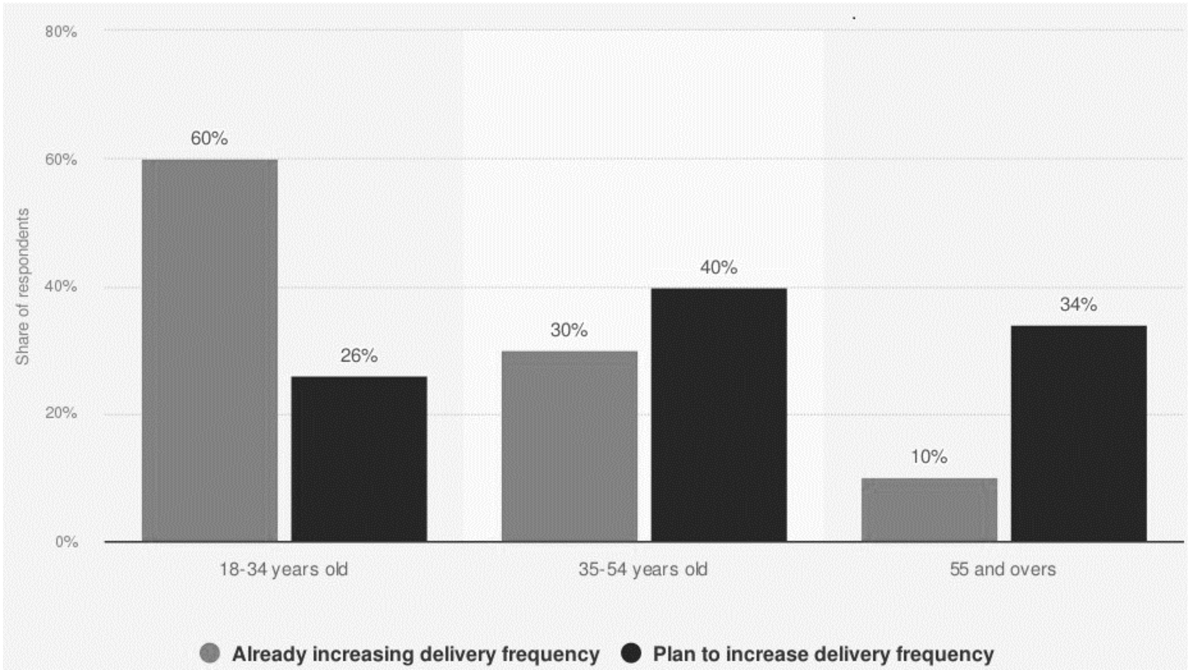


Figure 1. Percentage of the People in Different Age Groups Who Plan to Increase Their Use of OFD Services in the United Kingdom as of March 2020 (Statista, 2020)

The OFD industry's rapid expansion necessitates an agile service design that accommodates the nuances of digital interaction and customer engagement. Contrary to traditional service models, the online platform has enhanced visibility and necessitated a more complex interplay between customer and service provider (Pal et al., 2021). With the increased expectations and reliance on online food ordering, restaurants have been compelled to reassess how best they can adapt to the evolving business models and improve their service operations. Because the customer stands at the center of the service systems, achieving an effective service design in this emerging ecosystem depends critically on understanding the customers' perceptions and preferences (Natarajan et al., 2019). Acknowledging the centrality of customer perception in service design, this study endeavors to map out the service blueprint of online food ordering, assessing the end-to-end customer experience from food browsing to after-sales services.

While existing literature, such as the work of Smith and Heriyati (2023), who examine the impact of service quality on customer loyalty in OFD, and the study by Tandon et al. (2021), which explores the role of customer perceptions in food delivery app usage, offer valuable insights into consumer behavior and service delivery, a holistic analysis encompassing the complete spectrum of OFD services is less explored. In a recent study, Hoang and Le Tan (2023) investigated the effects of user interface design on customer ordering experiences, and Chowdhury (2023) examined the impact of perceived convenience and security on repeat purchase intention. However, these studies often address isolated factors within the service delivery system. In one of the pioneering studies that incorporates multiple factors, Chan and Gao (2021) introduce a comprehensive OFD service quality framework, referred to as DEQUAL. The framework addresses the omni-channel feature of OFD services that encompasses both the digital and physical components. Using a similar approach, Cheng, Chang, and Chen (2021) propose an alternative service quality scale for 20 key service factors with six dimensions, including reliability, maintenance of meal quality and hygiene, assurance, security, system operation, and traceability. In a later study, Koay, Cheah, and Chang (2022) focus on five significant service dimensions comprising assurance, meal quality, reliability, security, and system operation. Despite the valuable insights offered by these initial studies, they either lack a systematic framework that delineates OFD service quality or provide exploratory approaches.

In a more recent study, Ma et al. (2024) identify key service topics (qualities) pertaining to consumers' OFD experiences by utilizing the advanced BERTopic machine learning algorithm. In this regard, they developed a systematic framework that integrates aforementioned traditional methods with data analytics modeling based on user-generated online reviews. Our paper provides an alternative approach that can be utilized to synthesize findings from customer satisfaction surveys and machine learning analysis into a Quality Function Deployment (QFD) framework, thus refining service design in alignment with customer feedback. QFD is an effective tool to translate customer requirements into measurable design targets and drive them from the assembly level down through the sub-assembly, component, and production process levels. It provides a defined set of matrices utilized to facilitate this progression. What makes QFD unique is its primary focus on the customer requirements; in other words, what the customer truly wants rather than the innovation in technology. It has a wide spectrum of application areas in many key sectors, such as hospitality, logistics, healthcare, manufacturing, and education (Bossert 2021). In what follows, we discuss the details of our proposed framework in the context of OFD.

Research Methodology

This study employs a comprehensive methodology to tackle the service quality assessment within the online food ordering and delivery industry. The approach is grounded in the established principles of service design and systems thinking, providing a structured yet flexible framework that can accommodate the complex interplay of factors influencing customer experiences. We initiate our exploration by developing a service blueprint, a tool that has been effectively utilized to map out customer touchpoints and internal processes (Bitner, Ostrom, and Morgan, 2008; Kostopoulos, Gounaris, and Boukis, 2012; Hossain, Enam, and Farhana, 2017). This visual approach enables us to dissect the multifaceted nature of the food ordering and delivery business, aligning with the methodology of Patrício et al. (2011), who demonstrated how service blueprints could articulate the relationships among different service components and customer interactions.

Following the blueprint development, we designed and deployed a customer experience survey. The survey design is primarily informed by the service blueprint and focuses on the key influential factors identified during the blueprint development. To analyze the survey data, we have chosen a combination of statistical tools and machine learning algorithms. Our choice of tools is substantiated by the success of such methods in recent studies, such as one conducted by Markoulidakis et al. (2020), where the authors extract meaningful patterns and insights from complex customer datasets. The machine learning aspect in particular is an extension of the work of Sharma, Kumar, and Chuah (2021), who utilized predictive analytics to identify key drivers of customer satisfaction in e-commerce. Integrating machine learning outcomes into a QFD table represents a novel application, which has a proven track record in aligning service features with customer desires, as demonstrated by Wang, Guo, and Chen (2023) in the context of service enhancements.

The chosen methodology is in line with this study's objective, which is to gain a comprehensive understanding of and enhance the customer experience in the online food ordering and delivery business. This sequential linking of the service blueprint, survey data, machine learning analysis, and QFD creates a robust framework that ensures a thorough investigation of customer satisfaction drivers. It also provides actionable insights for decision makers as they seek to evolve their services in response to customer feedback.

Service Blueprint and Influential Factors

In our exploration of the online food ordering and delivery service, the customer's journey is mapped out through a service blueprint that begins with the digital engagement phase. The service blueprint was initially introduced as a visual representation to map the customer process (customer journey) against the organizational structure (Kostopoulos, Gounaris, and Boukis, 2012). The inclusion of physical evidence and the distinction between frontstage and backstage elements were later incorporated to shed light on the roles of service providers and customers using the service (Hossain, Enam, and Farhana, 2017). Providing a comprehensive view of critical components in a service process, the service blueprint guided the development of the online food ordering experience, as illustrated in Figure 2.

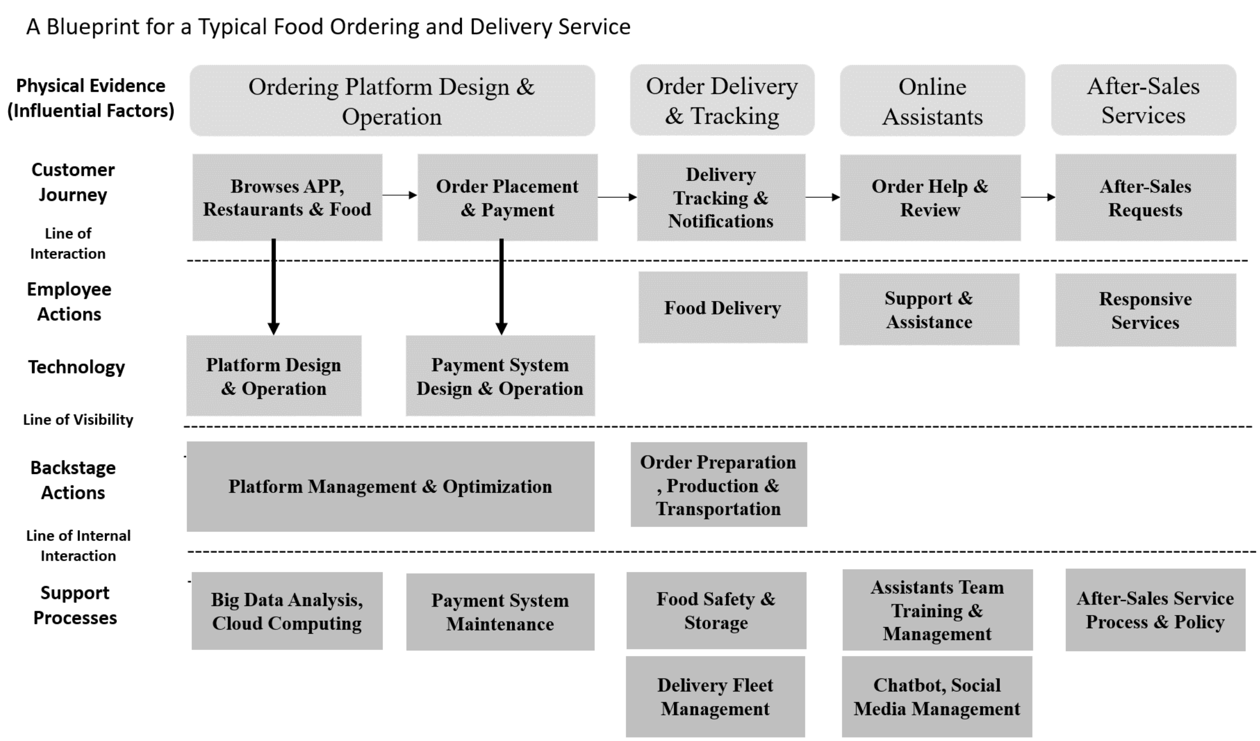


Figure 2. Service Blueprint for a Typical Online Food Ordering and Delivery Service

In the context of OFD, customers initiate their experience by navigating through an array of culinary options via a food ordering app or website, making critical judgments based on the platform's design and information. In a recent relevant study, Pal et al. (2021) investigated university students' satisfaction and loyalty in using online food delivery apps. They capture the customer's experience with the mobile apps via their attributes, including app visual design, navigational design, and information design. Their insights resonate with the observations of Lee et al. (2015) and Peters et al. (2016), who assert the significant influence of app attributes on cognitive and emotional responses—factors yet to be thoroughly investigated in the domain of online food delivery.

Expanding on this customer journey, the blueprint outlines subsequent phases, including order placement and payment, delivery tracking, order help and review, and after-sales requests. Each of these stages is influenced by identified factors, such as order delivery and tracking, online assistants, and after-sales services—areas highlighted by research as crucial for customer satisfaction (Hong et al., 2021; Wang et al., 2021; Roy Dholakia and Zhao, 2010). Our service blueprint serves as a foundational framework, linking these influential factors with the customer journey. It provides a clear visualization of the frontend user experience and the backend processes supporting it, ensuring a holistic understanding of the service's design and operation. This integrated perspective is vital for constructing our customer survey, which delves deeper into how these factors collectively influence the overall customer experience.

Based on this service blueprint, the four influential factors following the customer journey are identified and discussed below.

Ordering Platform Design and Operation

Platform Appearance and Layout: The visual appeal of an online food ordering platform is a critical determinant of customer trust and engagement. Jeannot, Jongmans, and Dampérat (2022) highlighted the direct correlation between a website's aesthetic appeal and user trust, underpinning the significance of design in the digital consumer experience. Kumar, Purani, and Viswanathan (2018) extend this understanding specifically to online food delivery platforms, demonstrating that aesthetic design not only enhances perceived usefulness and ease of use, but also fosters user enjoyment and loyalty. A well-crafted interface can captivate users, making the experience of browsing menus and placing orders more enjoyable (Cheung et al., 2015). Conversely, platforms with subpar design and poor visual appeal face user aversion due to the negative impact on user engagement (El Said, 2015).

The layout aspect of a platform—how its content is organized and presented—is equally important. Users expect a seamless and intuitive navigation experience that aligns well with advanced web technologies. Modern online food ordering apps have embraced a variety of user-centric customization features, enabling customers to tailor their browsing and ordering experience to their personal preferences (Liu and Lin, 2020). Additionally, they offer detailed and vibrant visuals of dishes and interactive elements, such as using various angles to illustrate food items, which help simulate a rich and engaging selection process (Vermeir and Roose, 2020). These elements are crucial in building a connection with the user, ensuring that the initial digital interaction is as compelling as the meal they intend to enjoy.

Information Quality: In the realm of online food ordering, the caliber of information presented on platforms plays a pivotal role in shaping customer satisfaction and trust. Furthering the concepts introduced by Chotigo and Kadono (2021), information quality on food delivery apps is evaluated based on its accuracy, comprehensiveness, relevance, and clarity. These attributes contribute to the perceived effectiveness of the information system, facilitating informed and confident user decisions. The presentation of this information, as emphasized by recent studies, including the arrangement, accessibility, and timeliness of updates, further influences user engagement and satisfaction (Lim and Rasul, 2022).

Contemporary web and app technologies have evolved to offer personalized and multimedia-rich content, enhancing interactivity and understanding for users (Shahbaznezhad, Dolan, and Rashidirad, 2021). High-quality information—characterized by its completeness, detail, precision, and reliability—becomes a cornerstone for user convenience, providing a seamless and enjoyable experience that can significantly impact purchasing decisions. The information quality in online food delivery apps, therefore, is not only a functional aspect, but is a strategic tool that businesses leverage for competitive advantage (Belanche, Flavián, and Pérez-Rueda, 2020).

Web and App System Quality: According to Kwaku and Antwi (2021), the quality of an e-commerce system is measured by the consumers' evaluation of the website's technical characteristics, which include usefulness, functionality, reliability, accessibility, flexibility, portability, integration, and interactivity. The behavioral intention of online shoppers is significantly influenced by the "ease of use" of the app, which is consistently demonstrated to be a crucial factor (Higgins et al., 2015). Additionally, a good online shopping web system should save the customers' transaction efforts and payment time. Otherwise, the customers may hesitate to use the website's payment system (Chen and Chang, 2023).

Security and Privacy: A customer's intention to buy a product from the website is heavily affected by the level of trust. Web system security and customers' privacy have been addressed as primary concerns among online consumers and treated as key elements for generating online trust (Flavián, Guinalíu, and Gurrea, 2006).

Price and Promotions: It is obvious that the price of the product, shipping costs, and discounts play a major role in driving customers to purchase online. Almost 4 out of 5 Americans say finding a great offer or discount is always on their minds throughout the entire buying experience (Roesler, 2018).

AI (Artificial Intelligence) Food Recommendation: Although little research has been done in this new area, web developers are using AI to help consumers choose meals based on their ordering history and preferences. An increasing number of companies claim that this technology could enhance customer ordering experience and boost sales (Haleem et al., 2022).

Order Delivery and Tracking

When restaurants receive an order, they prepare meals according to the stipulations of the order. The production process is usually invisible to customers. After the food is prepared and packaged, it is delivered by the courier to the customer's address. Delivery is particularly important to online retailing where there is a temporal separation between order placement and delivery. In this stage, distributing the right food to the right place at the right time plays a very significant role in overall customer satisfaction and loyalty. With the help of GPS navigation and tracking systems, delivery fleets can identify the most efficient routes and consequently improve the order on-time arrival rate. Similarly, customers can trace their order simultaneously using their smartphones.

Additionally, the visual presentation and temperature of delivered food significantly influence customer perceptions of quality and service excellence. Research by Zhong and Moon (2020) indicates that customers equate the care taken in food presentation with the overall quality of the service provided. Moreover, maintaining the appropriate food temperature from kitchen to consumer is not only a matter of taste, but also a health consideration, reinforcing trust in the service provider (Serhan and Serhan, 2019).

The appearance of delivery personnel also plays a pivotal role in shaping customer impressions. The uniform is a symbol of professionalism and a visual cue of a brand's commitment to quality and safety. Recent studies by Meena and Kumar (2022) have shown that delivery staff attire can significantly enhance the perceived value of the service and foster a sense of security among customers, which are particularly salient in the context of food handling and hygiene protocols. This extension of the service experience to include the conduct and appearance of delivery personnel underscores the need for comprehensive service design that encompasses all aspects of the customer journey, not just the digital interface or the food itself.

Online Assistants

In the OFD industry, online assistants play an indispensable role in enhancing customer experience and satisfaction. These digital interfaces, encompassing a range of technologies from chatbots to sophisticated virtual agents, are integral in providing immediate responses to customer inquiries, offering real-time assistance, and efficiently managing feedback and complaints. The utility of online assistants is rooted in their ability to offer personalized and contextual support, a factor that significantly influences customer loyalty and retention. Jenneboer, Herrando, and Constantinides (2022) highlight the effectiveness of chatbots in increasing user engagement and satisfaction by offering quick and accurate responses to common queries. Moreover, Makarius et al. (2020) underscore the role of virtual agents in handling complex customer service scenarios, thereby reducing wait times and improving overall service quality. The integration of AI-driven online assistants in the food delivery sector not only streamlines customer interaction but also contributes to building a robust customer service framework that is crucial for sustaining competitive advantage in this rapidly evolving industry.

After-Sales Services

As a crucial stage within the customer service life cycle, after-sales service represents the ongoing interaction between the service provider and the customer. The significance of after-sales services has been substantiated as a key predictor of customer satisfaction and retention (Shokouhyar, Shokoohyar, and Safari, 2020). The availability of after-sales services serves as an indispensable criterion in assessing customer satisfaction and driving recommendations. Consequently, e-commerce businesses are expected to deliver the highest level of after-sales customer service experience.

In the context of an online food ordering company, after-sales customer service encompasses various quality aspects, including the response time to customer inquiries, the politeness of staff, the handling of complaints, and the procedures for managing refunds. These elements collectively contribute to the overall after-sales experience and play a pivotal role in shaping customer satisfaction and loyalty.

Customer Satisfaction Survey Design and Data Collection

Understanding customer satisfaction, a crucial indicator of consumer contentment post-purchase, is essential for fostering loyalty, remedying service shortcomings, and attracting new patrons. To gauge this factor effectively, our study employs a comprehensive customer satisfaction survey. The survey's design captures both transactional experiences—individual interactions with the service—and overall satisfaction, a broader reflection of customer attitude toward the entire product/service offering, as conceptualized by Voorhees et al. (2017). The former pertains to discrete encounters, which are identified in our service blueprint as "Influential Factors," while the latter aggregates these experiences into a composite service impression, influencing the customer's propensity to endorse the service to others (Xu, 2021). The recommendation likelihood is another outcome variable, indicative of recommendation intentions and future business potential. Thus, our survey aims to dissect the determinants of customer satisfaction and their interplay with recommendation intent. Through empirical analysis, we seek to establish the key factors that drive consumer contentment and how they correlate with the willingness to recommend the platform, providing actionable insights for service enhancement.

A customer satisfaction questionnaire survey is designed accordingly and conducted online. One pre-screen question—"Have you ever ordered food online?"—was included at the very beginning to filter out those people who have never ordered food online. Demographic data, including gender, age, level of education, employment status, and marital status, and consumer behavior data including ordering platform, ordering frequency, and average expenses are collected to better understand the social background and shopping habits of the respondents (see Survey Dimension 1–2 in Appendix A).

Considering the service blueprint and the previous discussion about influential factors, the questionnaire divides the OFD business process into four survey dimensions (see Survey Dimensions 3–6 in Appendix A, corresponding to the four influential factors of the service blueprint). Each dimension has its related sub-questions based on the discussion of influential factors (9 questions for Dimension 3, 5 questions for Dimension 4, 1 question for Dimension 5, and 2 questions for Dimension 6). The primary goal is to explore consumers' actual ordering app user experience by holistically investigating various service quality parameters, starting from the time users interact with the apps for searching food to the after-sales service. This analysis aims to understand the complicated relationship among perceived service quality, satisfaction, and loyalty in using food ordering services.

Each sub-question is treated as an independent variable and is rated using a 1–5 Likert scale, with 1 indicating the lowest possible customer satisfaction level and 5 representing the highest. The survey questions and the summary outcome are presented in Appendix A. At the end of the survey, respondents are asked to rate their overall satisfaction (Q21) and the likelihood of recommending the food ordering platform to others (Q22). Both questions are set as two dependent variables (see Survey Dimension 7 in Appendix A).

The survey was conducted via a Qualtrics survey research suite—a popular cloud-based web survey tool enabled by a globally recognized survey technology enterprise. The survey targeted respondents who are familiar with or have used online food ordering services and delivery options in the recent past. The responses of participants who have not used online food ordering services and delivery were excluded from the analysis. After the survey was published, the survey platform notified those individuals who belong to the demographic group, including participants currently residing in the United States and those who ordered food online via email, in-app, and SMS notifications. Each respondent's address, demographic information, and email address had been verified by Qualtrics before participation. These individuals were then able to take the survey after passing the qualifying screeners to move forward to being counted as acceptable "completes." The respondents who finished in less than one half of the median completion time were disregarded, because they were viewed as answering the survey in a perfunctory manner. Potential biases were addressed through the survey design and administration process. Selection bias was mitigated by using a random sampling technique, and response bias was minimized by ensuring anonymity. Additionally, speeders and straight-liners were filtered out to maintain data quality. A total of 379 qualified survey samples over a 2-month period were successfully populated. This sample size falls within the recommended range of 200 to 500 responses, as recommended by the guidelines provided by Iacobucci and Churchill (2018) and satisfies the minimum requirement of a sample size of 322, as recommended by Zikmund et al. (2013).

Customer Survey Data Analysis by Machine Learning Algorithm

Machine learning algorithms usually employ computational methods to "learn" information directly from data for making predictions or decision supports. With its growing popularity in a wide variety of industries, machine learning methods are increasingly used for various aspects of survey research, which include data processing, responsive/adaptive designs, nonresponse adjustments and weighting, classification, and making predictions (Buskirk et al., 2018). In our study, after gathering the basic statistics of the survey results, we employed and compared three machine learning algorithms, namely, decision tree, random forest, and support vector machines, to identify the independent variables that represent the key drivers of customer value and extract useful data insights.

In the realm of data analysis, Decision Tree Methods (DTMs), including their derivative Random Forest Method (RFM), stand out for their interpretability and robustness. These methods, which have been applied across engineering, medicine, finance, and marketing, have proven particularly effective in analyzing customer behavior and survey data. For instance, decision trees have offered valuable insights into the main factors affecting customer satisfaction by revealing priority areas for service improvement (Xie and Zhao, 2010). The RFM, built on the decision tree foundation, enhances prediction accuracy by aggregating multiple trees to form a more potent model ensemble, thus offering a nuanced understanding of customer survey data (Tsami et al., 2018).

Support Vector Machines (SVMs) complement these methods by classifying data with a high-dimensional approach that maximizes the margin between data points, making it suitable for complex classification tasks often encountered in survey analysis (Kirchner and Signorino, 2018). Together, these machine learning algorithms form a comprehensive toolkit for deriving actionable insights from customer feedback, essential for service design and development in today's data-driven decision-making environments. The three models were all built by 80% of the whole survey data entries (randomly selected), and their prediction accuracies were tested on the remaining 20% of the data. We selected the best model with the highest testing data prediction accuracy and built the QFD table based on the survey data insights provided by the selected machine learning model.

Quality Function Deployment (QFD)

Introduced in the 1970s, the Quality Function Deployment is a crossfunctional method that can be utilized to translate customer requirements of product and service design specifications (Jin et al., 2009). It can be quite instrumental in guiding businesses in designing their products or services to meet the requirements and expectations of customers (Erdil and Arani, 2019). The QFD table is a basic tool of the QFD method. The structure of a QFD table can be divided into nine parts: voice of customers (wants), importance of wants, relationships between customer wants and technical specifications, competitive analysis, correlation between technical specifications, technical specifications, technical specification priorities, technical comparisons, and technical targets. In this research, a QFD table (see Appendix B) was developed based on the survey data analysis and prediction results to integrate the voice of customers into service design.

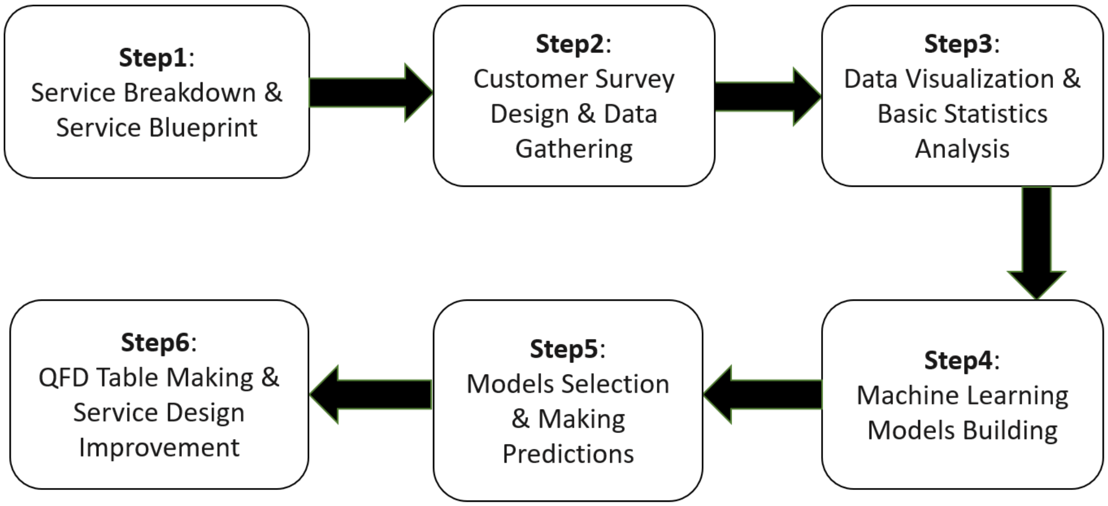


Figure 3. Flowchart of Research Methodology

Research Findings

Survey Sample Characteristics

The survey was completed in June 2020, with 379 qualified respondents, 60% of whom identified as female and 40% as male. The characteristics of the survey samples are summarized in **Error! Reference source not found.**. Most respondents were under the age of 45 (80%) and hold at least a bachelor's degree (66.5%). Most were currently married (57.3%) and held a full- time job (62.3%). More than half of the respondents have ordered food from both a restaurant app and a third-party app (e.g., Uber Eats, Grub Hub, etc.) in the recent past, whereas the number of respondents who have ordered only from a third-party platform (19%) was less than the number who only used a restaurant app (29%). It is also worth noticing that most of the respondents indicated that they order fewer than five times per week on average (62.3%) and typically spend \$15 to \$35 each time (51.7%).

The survey data were imported into R-studio for statistical analysis, and all of the survey responses were converted to categorical variables. We first calculated the mean for each survey item and Spearman Correlation Coefficient of independent variables Q4 through Q20 to the two dependent variables (targets): Q21 and Q22 (see Appendix A). Some values were missing (skipped questions by the respondents), and the two targets were unbalanced with more selections of "4" and "5" than the others (see Figure 4). Thus, some data preprocessing steps were needed before building the complete machine learning models. To address the missing values, we performed missing value imputations in predictor data using the proximity matrix. After imputation, all of the missing values were backfilled. Later, the oversampling technique was applied to balance the proportion of classes in the targets.

Table 1. Characteristics of Survey Samples

Characteristics	**Category**	**% of Respondents**
Gender	Male	40.0%
	Female	60.0%
Age	18–24 years old	22.7%
	25–34 years old	28.5%
	35–44 years old	29.6%
	45–54 years old	8.4%
	Over 55	10.8%
Level of education	High school degree or equivalent	17.4%
	Bachelor's degree (e.g., BA, BS)	35.9%
	Master's degree (e.g., MA, MS, MEd)	30.6%
	Doctorate (e.g., PhD, EdD)	12.7%
	Other	3.4%
Current employment status	Employed full time	62.3%
	Employed part time	9.5%
	Self-employed	5.0%
	Unemployed	4.7%
	Student	10.6%
	Retired	6.3%
	Other	1.6%
Marital status	Single	33.8%
	Married	57.3%
	In a domestic partnership	4.2%
	Divorced	3.4%
	Widowed	1.3%
Platform used to order food online	Order directly from restaurant app	29.0%
	Order through a third-party platform (e.g., Uber eats, Grub Hub)	19.0%
	Both	52.0%
Order frequency per week	More than 5–7 times / week	18.2%
	5–7 times/ week	19.5%
	2–4 times/week	32.7%
	1–2 times/ week or less	29.6%
Amount spent per order	Less than $15	7.7%
	$15–$35	51.7%
	$35–$50	26.9%
	More than $50	13.7%

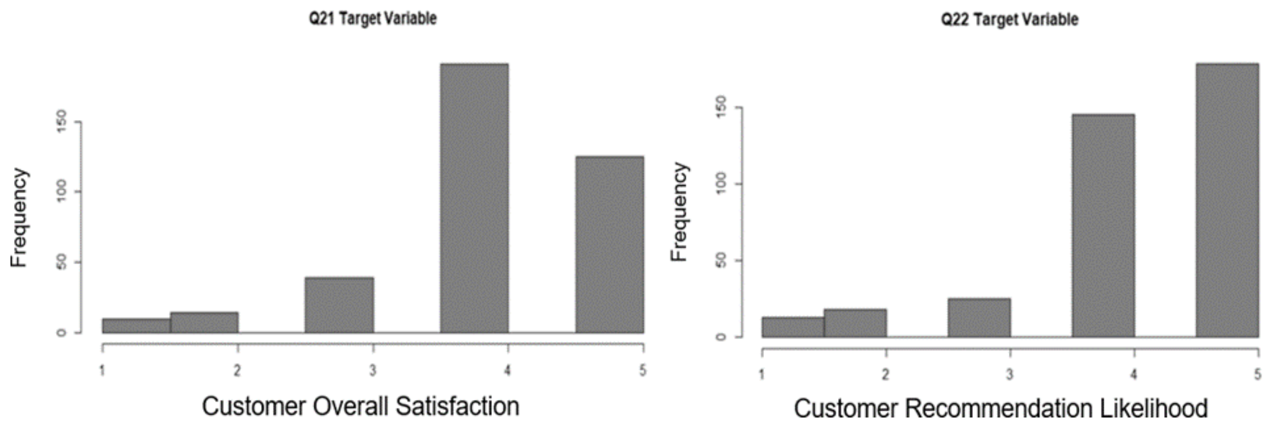


Figure 4. Histogram of the Target Variable Q21 (Customer Overall Satisfaction) and Q22 (Customer Recommendation Likelihood)

Machine Learning Models Building and Selection

Two decision tree models were built for Q21 and Q22. After cross-validation and tree pruning, the best decision tree models resulted in the testing data classification accuracy rates of 70.87% and 61.82% for Q21 and Q22, respectively. On the other hand, after parameters tuning, the best random forest model resulted in 88.19% testing set classification accuracy for Q21, and 79.09% testing set classification accuracy for Q22. Moreover, the SVM model yielded 67.72% and 60.90% testing classification accuracy levels for Q21 and Q22, respectively. Our results for this stage are summarized in Table 2.

Table 2. Training and Testing Classification Accuracy for Q21 and Q22 by Each Machine Learning Model

Accuracy / Model	Training Classification Accuracy (Q21)	Testing Classification Accuracy (Q21)	Training Classification Accuracy (Q22)	Testing Classification Accuracy (Q22)
Decision tree	82.74%	70.87%	78.08%	61.82%
RFM	87.50%	88.19%	82.19%	79.09%
SVM	77.58%	67.72%	66.21%	60.90%

According to these results, RFM has provided the highest classification accuracy for both target variables; therefore, this method is selected for the remainder of the analysis and prediction tasks. Figure 5 shows the random forest classification model confusion matrix and statistics. The accuracies are much higher than the No Information Rates (p-value $< 2 \times 10^{-16}$), and the Cohen's Kappa values are all above 0.7, which indicates the prediction model is "substantial and reliable" (McHugh, 2012). Figure 6 and Figure 7 demonstrate the independent variables' significance for Q21 and Q22.

(a) Q21

```
Confusion Matrix and Statistics

          Reference
Prediction  1  2  3  4  5
         1 19  0  0  0  0
         2  0 22  0  0  0
         3  0  0 24  0  0
         4  0  1  1 31  6
         5  0  0  0  7 16

Overall Statistics

               Accuracy : 0.8819
                 95% CI : (0.8127, 0.9324)
    No Information Rate : 0.2992
    P-Value [Acc > NIR] : < 2.2e-16

                  Kappa : 0.8497
```

(b) Q22

```
Confusion Matrix and Statistics

          Reference
Prediction  1  2  3  4  5
         1 15  0  0  0  0
         2  0 10  0  0  0
         3  0  0 10  3  0
         4  0  0  0 23  7
         5  0  0  1 12 29

Overall Statistics

               Accuracy : 0.7909
                 95% CI : (0.703, 0.8626)
    No Information Rate : 0.3455
    P-Value [Acc > NIR] : < 2.2e-16

                  Kappa : 0.7183
```

Figure 5. Random Forest Classification Model Confusion Matrix and Statistics for Testing Set

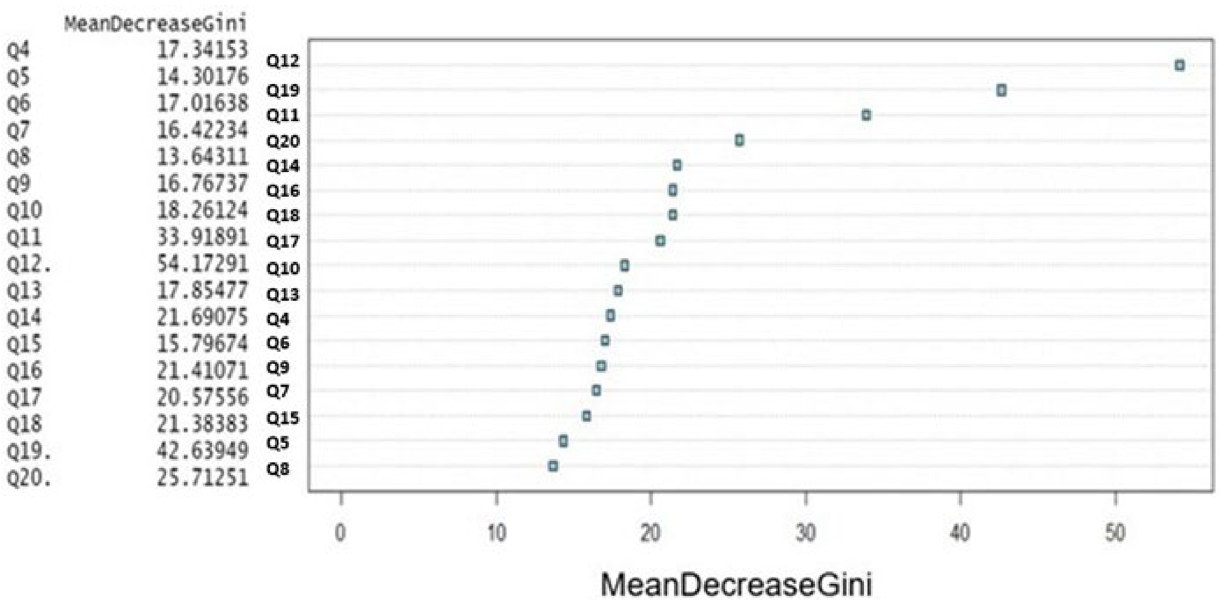


Figure 6. Variable Importance for Target Q21

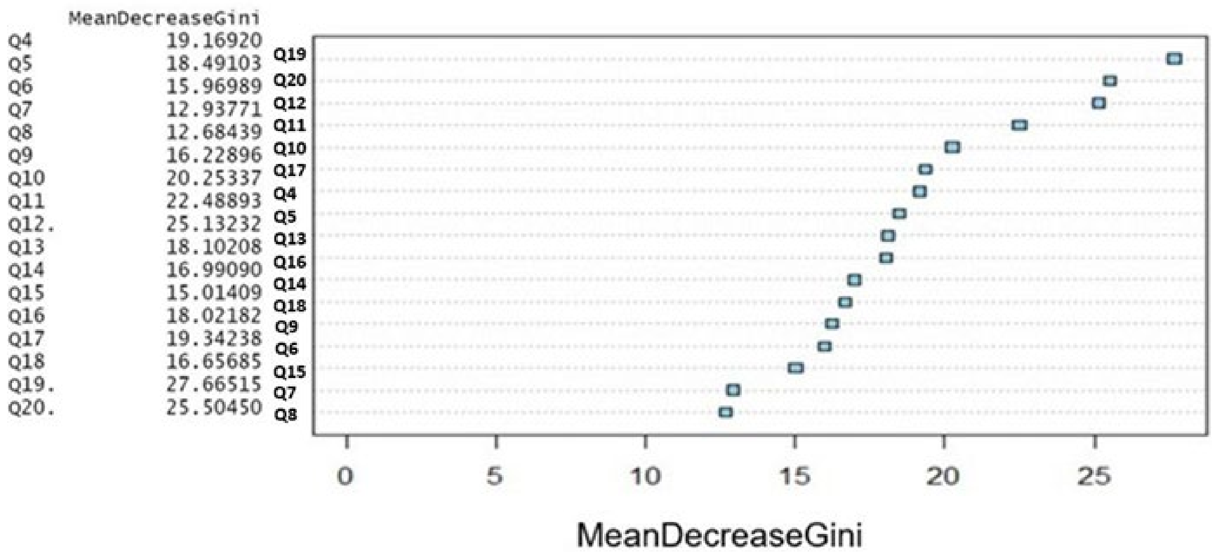


Figure 7. Variable Importance for Target Q22

Formation of the Quality Function Deployment Table

The QFD table is developed using the results of the data analysis and presented in Appendix B. The traditional full implementation of QFD in the manufacturing industry involves four phases: Phase 1 (QFD table) translates customer requirements into technical design requirements; Phase 2 (product design) turns technical requirements into part specifications; Phase 3 (process design) turns part requirements into process requirements; and Phase 4 (process control) turns process requirements into production requirements. Building the QFD table constitutes a critical phase as it captures the voice of the customer and provides a way for efforts toward improving the service design. As such, we primarily focus on the process of QFD table development as detailed in the flowing steps:

Step 1. Identify customer needs and determine their degrees of importance: As shown in Figures 6 and 7, the survey items Q11 (getting and using coupons, promotions, and deals), Q12 (food suggestion and recommendation), Q19 (platform's ability to resolve complaints) and Q20 (handling refund requests) are identified by the Random Forest Method as the most important impactors pertaining to the main targets (i.e., Q21 and Q22). Thus, we identify these four items as the most significant customer needs, as demonstrated in row 2 to row 5 in Appendix B. Their rate of importance is determined by the average of their variable importance (i.e., Mean Decrease Gini) in RFM (i.e., rate of importance of Q19 = (42.6 + 27.7) / 2 = 35.15). We later convert the importance rate to a 5-point scale and insert the numbers into the QFD table under the column titled as "Rate of Importance" (Column 6 in Appendix B).

Step 2. Pinpoint technical requirements and determine inter-relationships: Once the customer needs and their degrees of importance have been identified, technical requirements (i.e., the service design requirements) need to be identified. Service design requirements are the translation of these customer needs to service designs, and each requirement can fulfill one or more customer needs. Our analysis has identified five service design requirements (Column 1-5 in Appendix B) and their inter-relationships, illustrated by color-coded circles.

Step 3. Determine relationships between customer needs and technical requirements: The central grids of QFD, which connect the customer needs listed on the very left column of the QFD Table with the service design requirements listed across the top row, indicate the corresponding relationships (correlation coefficients) between the two (represented as a black single ring, double ring, and triangle in the middle region of Appendix B). There should be at least one service design requirement that has a strong correlation with one of the customer's needs. Otherwise, a particular customer's need may not be properly addressed.

Step 4. Determine plan for customer needs and sales point: In this step, we first set the scores for the current state of the company (i.e., "Company Now" [Column 7 in Appendix B] by using the average values (rounded to the nearest integer) of each survey item. Subsequently, based on the RFM forecasting, we derive the list of the desired values for each survey item to raise to, namely, "Plan" (Column 8 in Appendix B). In this case, we target to increase the averages of customer overall satisfaction (Q21) and recommendation likelihood (Q22) from current value of 4 (Appendix A, "Mean" of Q21 and Q22, rounded to the nearest integer) to 5. According to our model forecasting, if we can raise the means of the customer needs represented by Q11, Q12, Q19, and Q20 from current value 4 ("Company Now") to 5 ("Plan"), then Q21 and Q22, respectively, will have 63.7% and 52%, respectively, chances to reach the "Very Good" overall satisfaction level and the "Extremely Likely" customers' recommendation willingness level. These probabilities are detailed in Table 3. The "Rate of Improvements" (Column 9 in Appendix B) are calculated by dividing the scores under the "Plan" column by the values of "Company Now" column.

Table 3. Predicted Probability for Q21 and Q22 Rankings

	Q21 Rankings						**Q22 Rankings**				
	1	2	3	4	5		1	2	3	4	5
Predicted Probability (Before)	0.0	9.3%	9.6%	61.0%	20.1%	Predicted Probability (Before)	0.3%	3.4%	22.0%	44.0%	30.3%
Predicted Probability (After)	0.0	0.3%	14.6%	33.1%	52.0%	Predicted Probability (After)	1.9%	3.7%	9.0%	21.7%	63.7%

The Sales Point, shown as a single red ring and double red ring under Column 10 in Appendix B in the QFD Table, indicates which customer expectations have more important effects on marketing. Customer needs items with higher marketing importance were assigned 1.5 points, and 1.2 points were assigned to the items with lower importance. The "Absolute Weight" (Column 11 in Appendix B) is the multiplication of Column 6 ("Rate of Importance," column 9 "Rate of

Improvement") and Column 10 ("Sales Point"). The last column, "Demand Weight," (Column 12 in Appendix B) is the percent ratio of "Absolute Weight" for each factor.

Step 5. Develop importance rating and action plan for technical requirements: This step completes the basement of the QFD table where the "total importance rating" is documented. The Total Importance Rating (Row 6 in Appendix B) is the relative weight of each technical requirement in terms of satisfying the customers' demands. The importance ratings determine which technical requirement, in our case the service design requirement, should receive the most attention in the service design and improvement process. Basically, it is calculated by the following expression:

$$w_j = \Sigma_{i=1}^{n} D_i r_{ij} \tag{1}$$

where w_j is the total importance rating of the j^{th} technical requirement; D_i is the "Demand Weight" of i^{th} customer requirement; and r_{ij} is the correlation coefficient (relationship defined in Step 3) between the i^{th} customer requirement and the j^{th} technical requirement. The "Percent" (Row 7 in Appendix B) is then determined by the following equation:

$$P_j = w_j / \sum_{j=1}^{m} w_j \tag{2}$$

where P_j is the importance rating proportion of the j^{th} technical requirement against the total. This percentage indicates which service design requirement has relative higher significance or urgency to be fulfilled. The "Company Situation Now" (Row 8 in Appendix B) assesses the current situations for each service design item. The final row comprises "Action Plans" (Row 9 in Appendix B), which are the perspective actions that should be conducted in the new service design or renovations corresponding to each service design requirement. "Upgrade Food Recommendation System and AI Algorithm" and "Provide More Professional Training to Staffs on Customer Services and Handling Complains" are identified to be the two critical factors (with the highest corresponding "Percent" value) in the service design improvement process. In some QFD implementations, this step could also include the evaluation of market competitors in terms of technical requirements, and the results would usually be recorded in a basement row of the matrix.

Key Managerial Insights

The proposed analysis has examined 19 pivotal factors influencing customer satisfaction in OFD and their propensity to recommend OFD services to others. Among these, four factors stand out as the most influential according to the employed machine learning approach: i) the ease of getting and using coupons, promotions, and deals; ii) the ease of accessing and utilizing coupons, promotions, and deals; ii) the helpfulness of food suggestions and recommendations; iii) the efficacy of resolving complaints; and iv) the handling of refund requests. Consequently, as a key insight, the analysis concludes that enhancing these service elements can elevate customer satisfaction levels and recommendation behaviors. Notably, the analysis suggests that improving the food recommendation system is poised to elicit the most positive responses from customers, as indicated by the highest demand weight in the QFD analysis. Thus, companies should prioritize upgrading their food recommendation systems and embracing relevant AI technologies. Other recommended actions include providing comprehensive professional training to staff in customer service and complaint resolution, continually refining refund policies and procedures, periodically offering coupons, and reducing or eliminating delivery fees where feasible.

Conclusions and Future Work

This study has addressed the intricate challenge of measuring service design and quality within the online food ordering and delivery (OFD) domain. Leveraging a combination of service breakdown and customer experience survey data, we employed three prominent statistical and machine learning algorithms—decision trees, random forests (RFM), and support vector machines (SVM)—to discern the relationships between service components and customer-defined value. Our analysis revealed that the RFM outperformed others, particularly in predicting overall satisfaction and likelihood of recommendation. Utilizing RFM, we constructed a Quality Function Deployment (QFD) table, translating customer needs into actionable service design elements. This strategic integration of customer feedback into service design not only enriches theoretical insights, but also provides a unique framework for enhancing service value.

In this context, this study has examined critical factors influencing customer satisfaction in OFD, identifying four pivotal elements, namely, ease of accessing and utilizing coupons, promotions, and deals; helpfulness of food suggestions and recommendations; efficacy of resolving complaints; and handling refund requests. Our results highlight the significance of enhancing these service elements to elevate customer satisfaction levels and recommendation behaviors. Moreover, our analysis indicates that improving the food recommendation system holds particular promise in eliciting positive responses from customers. Therefore, prioritizing upgrades to food recommendation systems and embracing relevant AI technologies emerge as strategic imperatives for companies in the OFD sector.

Our findings underscore the significance of considering service quality throughout the entire OFD process, advocating for a more integrated approach from platform browsing to after-sales service. Embracing a holistic service design approach, which encompasses factors like platform design, order delivery, online assistants, and after-sales service, can foster a customer-centric culture and elevate service quality, thereby increasing customer loyalty. In conclusion, this study contributes a novel methodology framework that integrates service blueprinting, customer surveys, data analysis, and QFD to renovate OFD services, offering valuable insights for businesses striving to enhance customer experiences in the digital age.

Looking ahead, promising avenues for future research include exploring the interrelationships among service components and their collective impact on OFD processes. Additionally, investigating the influence of sociodemographic factors on customer satisfaction and loyalty, as well as extending the scope to include the experiences of employees and backstage service components, offer opportunities for further insights and improvements in service design.

References

Al Amin, M., M.S. Arefin, M.R. Alam, T. Ahammad, and M.R. Hoque. 2021. "Using Mobile Food Delivery Applications during COVID-19 Pandemic: An Extended Model of Planned Behavior." *Journal of Food Products Marketing* 27(2):105–126.

Belanche, D., M. Flavián, and A. Pérez-Rueda. 2020. "Mobile Apps Use and WOM in the Food Delivery Sector: The Role of Planned Behavior, Perceived Security and Customer Lifestyle Compatibility." *Sustainability* 12(10):4275.

Bira, C., J. Jungb, J. Taoc, and N.O. Widmard. 2023. "Online Grocery Shopping Practices and Intentions Shaped by Pandemic-Era Experiences." *Journal of Food Distribution Research* 54(3):1–29.

Bitner, M.J., A.L. Ostrom, and F.N. Morgan. 2008. "Service Blueprinting: A Practical Technique for Service Innovation." *California Management Review* 50(3):66–94.

Bossert, J.L. 2021. *Quality Ffunction Deployment: The Practitioner's Approach*. Boca Raton, FL: CRC Press.

Business Wire. 2020. *Global Online Food Delivery Services Market (2020 to 2030)—COVID-19 Growth and Change*. Available online: https://www.businesswire.com/news/home/20200511005687/en/Global-Online-Food-Delivery-Services-Market-2020-to-2030---COVID-19-Growth-and-Change---ResearchAndMarkets.com.

Buskirk, T.D., A. Kirchner, A. Eck, and C.S. Signorino. 2018. "An Introduction to Machine Learning Methods for Survey Researchers." *Survey Practice* 11(1):1–10.

Centers for Disease Control and Prevention. (2020). *Community, Work, and School*.

Chan, J., and Y.L. Gao. 2021. "Measuring the Up-to-Date Quality of Online Food Delivery: Formative Index Construction." *International Journal of Contemporary Hospitality Management* 33(12):4550–4568.

Chen, J., and Y.-W. Chang. 2023. "How Smart Technology Empowers Consumers in Smart Retail Stores? The Perspective of Technology Readiness and Situational Factors." *Electronic Markets* 33(1):1.

Cheng, C.C., Y.Y. Chang, and C.T. Chen. 2021. "Construction of a Service Quality Scale for the Online Food Delivery Industry." *International Journal of Hospitality Management* 95: 102938.

Cheung, C.M.K., X.L. Shen, Z.W.Y. Lee, and T.K.H. Chan. 2015. "Promoting Sales of Online Games through Customer Engagement." *Electronic Commerce Research and Applications* 14(4):241–250.

Chotigo, J., and Y. Kadono. 2021. "Comparative Analysis of Key Factors Encouraging Food Delivery App Adoption before and during the COVID-19 Pandemic in Thailand." *Sustainability* 13(8):4088.

Chowdhury, R. 2023. "Impact of Perceived Convenience, Service Quality and Security on Consumers' Behavioural Intention towards Online Food Delivery Services: The Role of Attitude as Mediator." *SN Business & Economics* 3(1):29.

Donthu, N., and A. Gustafsson. 2020. "Effects of COVID-19 on Business and Research." *Journal of Business Research* 117:284–289.

El Said, G.R. 2015. "Understanding Knowledge Management System Antecedents of Performance Impact: Extending the Task-Technology Fit Model with Intention to Share Knowledge Construct." *Future Business Journal* 1(1–2):75–87.

Erdil, N.O., and O.M. Arani. 2019. "Quality Function Deployment: More Than a Design Tool." *International Journal of Quality and Service Sciences* 11(2):142–166.

Flavián, C., M. Guinalíu, and R. Gurrea. 2006. "The Role Played by Perceived Usability, Satisfaction and Consumer Trust on Website Loyalty." *Information & Management* 43(1): 1–14.

Gunden, N., C. Morosan, and A. DeFranco. 2020. "Consumers' Intentions to Use Online Food Delivery Systems in the USA." *International Journal of Contemporary Hospitality Management* 32(3):1325–1345.

Haleem, A., M. Javaid, M. Asim Qadri, R. Pratap Singh, and R. Suman. 2022. "Artificial Intelligence (AI) Applications for Marketing: A Literature-based Study." *International Journal of Intelligent Networks* 3:119–132.

Higgins, L.M., M.M. Wolf, R.H.J.N. Bitter, and W.H. Amspacher. 2015. "Winery Distribution Choices and the Online Wine Buyer." *Journal of Food Distribution Research* 46(3):32–49.

Hoang, H., and T. Le Tan. 2023. "Unveiling Digital Transformation: Investigating Technology Adoption in Vietnam's Food Delivery Industry for Enhanced Customer Experience." *Heliyon* 9(9):e19719.

Holmlund, M., Y. Van Vaerenbergh, R. Ciuchita, A. Ravald, P. Sarantopoulos, F.V. Ordenes, and M. Zaki. 2020. "Customer Experience Management in the Age of Big Data Analytics: A Strategic Framework." *Journal of Business Research* 116:356–365.

Hong, C., H. Choi, E.-K. Choi, and H.-W. Joung. 2021. "Factors Affecting Customer Intention to Use Online Food Delivery Services before and during the COVID-19 Pandemic. *Journal of Hospitality and Tourism Management* 48:509–518.

Hossain, M.Z., F. Enam, and S. Farhana. 2017. "Service Blueprint a Tool for Enhancing Service Quality in Restaurant Business." *American Journal of Industrial and Business Management* 07(07):919–926.

Iacobucci, D., and G. Churchill. 2018. *Marketing Research: Methodological Foundations*, 12th ed. Chula Vista, CA: Southwestern College Publishing.

Jeannot, F., E. Jongmans, and M. Dampérat. 2022. "Visual Design and Online Shopping Experiences: When Expertise Allows Consumers to Refocus on Website Attractiveness." *Recherche et Applications En Marketing (English Edition)* 37(1):59–81.

Jenneboer, L., C. Herrando, and E. Constantinides. 2022. "The Impact of Chatbots on Customer Loyalty: A Systematic Literature Review." *Journal of Theoretical and Applied Electronic Commerce Research* 17(1):212–229.

Jin, B., Y.G. Ji, K. Choi, and G. Cho. 2009. "Development of a Usability Evaluation Framework with Quality Function Deployment: From Customer Sensibility to Product Design." *Human Factors and Ergonomics in Manufacturing & Service Industries* 19:177–194.

Jun, K., B. Yoon, S. Lee, and D.-S. Lee. 2021. "Factors Influencing Customer Decisions to Use Online Food Delivery Service during the COVID-19 Pandemic." *Foods (Basel, Switzerland)* 11(1).

Kirchner, A., and C.S. Signorino. 2018. "Using Support Vector Machines for Survey Research." *Survey Practice* 11(1):1–14.

Koay, K.Y., C.W. Cheah, and Y.X. Chang. 2022. "A Model of Online Food Delivery Service Quality, Customer Satisfaction and Customer Loyalty: A Combination of PLS-SEM and NCA Approaches." *British Food Journal* 124(12):4516–4532.

Kostopoulos, I., S. Gounaris, and A. Boukis. 2012. "Service Blueprinting Effectiveness: Drivers of Success." *Managing Service Quality* 22.

Kumar, D.S., K. Purani, and S.A. Viswanathan. 2018. "Influences of 'Appscape' on Mobile App Adoption and M-loyalty." *Journal of Retailing and Consumer Services* 45:132–141.

Kunneman, Y., M. Alves da Motta-Filho, and J. van der Waa. 2022. "Data Science for Service Design: An Introductory Overview of Methods and Opportunities." *The Design Journal* 25(2):186–204.

Kwaku, A.R., and S. Antwi. 2021. "Consumers Attraction to Purchase Online: Website Quality as a Major Influencing Factor." *Open Journal of Business and Management* 09(03):1133–1150.

Lee, D., J. Moon, Y.J. Kim, and M.Y. Yi. 2015. "Antecedents and Consequences of Mobile Phone Usability: Linking Simplicity and Interactivity to Satisfaction, Trust, and Brand Loyalty." *Information & Management* 52(3):295–304.

Lim, W.M., and T. Rasul. 2022. "Customer Engagement and Social Media: Revisiting the Past to Inform the Future." *Journal of Business Research* 148:325–342.

Lin, X., A. Al Mamun, Q. Yang, and M. Masukujjaman. 2023. "Examining the Effect of Logistics Service Quality on Customer Satisfaction and Re-Use Intention." *PloS One* 18(5):e0286382.

Liu, C.-F., and C.-H. Lin. 2020. "Online Food Shopping: A Conceptual Analysis for Research Propositions." *Frontiers in Psychology* 11.

Ma, B., Y.D. Wong, C.C. Teo, and Z. Wang. 2024. "Enhanced Understandings of Online Food Delivery's Service Quality with Online Reviews." *Journal of Retailing and Consumer Services* 76:103588.

Makarius, E.E., D. Mukherjee, J.D. Fox, and A.K. Fox. 2020. "Rising with the Machines: A Sociotechnical Framework for Bringing Artificial Intelligence into the Organization." *Journal of Business Research* 120:262–273.

Markoulidakis, I., I. Rallis, I. Georgoulas, G. Kopsiaftis, A. Doulamis, and N. Doulamis. 2020. "A Machine Learning Based Classification Method for Customer Experience Survey Analysis." *Technologies* 8(4):76.

McHugh, M.L. 2012. "Interrater Reliability: The Kappa Statistic." *Biochemia Medica* 22(3): 276–282.

Meena, P., and G. Kumar. 2022. "Online Food Delivery Companies' Performance and Consumers' Expectations during Covid-19: An Investigation Using Machine Learning Approach." *Journal of Retailing and Consumer Services* 68:103052.

Natarajan, C., S. Gupta, and N. Nanda. 2019. "Food Delivery Services and Customer Preference: A Comparative Analysis." *Journal of Foodservice Business Research* 22:1–12.

Noyes, J., A. Booth, G. Moore, K. Flemming, Ö. Tunçalp, and E. Shakibazadeh. 2019. "Synthesizing Quantitative and Qualitative Evidence to Inform Guidelines on Complex Interventions: Clarifying the Purposes, Designs and Outlining Some Methods." *BMJ Global Health* 4(Suppl 1):e000893.

Pal, D., S. Funilkul, W. Eamsinvattana, and S. Siyal. 2021. "Using Online Food Delivery Applications during the COVID-19 Lockdown Period: What Drives University Students' Satisfaction and Loyalty?" *Journal of Foodservice Business Research* 25:1–45.

Patrício, L., R.P. Fisk, J. Falcão e Cunha, and L. Constantine. 2011. "Multilevel Service Design: From Customer Value Constellation to Service Experience Blueprinting." *Journal of Service Research* 14(2):180–200.

Peters, T., Ö. Işik, O. Tona, and A. Popovič. 2016. "How System Quality Influences Mobile BI Use: The Mediating Role of Engagement." *International Journal of Information Management* 36(5):773–783.

Rathore, S.S., and M. Chaudhary. 2018. "Consumer's Perception on Online Food Ordering." *International Journal of Management & Business Studies* 8(4).

Roesler, P. 2018. "New Study Shows Deals and Promotions Affect Every Part of Shopping Experience." *Inc.Com*. Available online: https://www.inc.com/peter-roesler/new-study-shows-deals-promotions-affect-every-part-of-shopping-experience.html.

Serhan, M., and C. Serhan. 2019. "The Impact of Food Service Attributes on Customer Satisfaction in a Rural University Campus Environment." *International Journal of Food Science* 1–12.

Shahbaznezhad, H., R. Dolan, and M. Rashidirad. 2021. "The Role of Social Media Content Format and Platform in Users' Engagement Behavior." *Journal of Interactive Marketing* 53:47–65.

Shankar, A., C. Jebarajakirthy, P. Nayal, H.I. Maseeh, A. Kumar, and A. Sivapalan. 2022. "Online Food Delivery: A Systematic Synthesis of Literature and a Framework Development." *International Journal of Hospitality Management* 104:103240.

Sharma, R., A. Kumar, and C. Chuah. 2021. "Turning the Blackbox into a Glassbox: An Explainable Machine Learning Approach for Understanding Hospitality Customer." *International Journal of Information Management Data Insights* 1(2):100050.

Shokouhyar, S., S. Shokoohyar, and S. Safari. 2020. "Research on the Influence of After-Sales Service Quality Factors on Customer Satisfaction." *Journal of Retailing and Consumer Services* 56:102139.

Smith, P.A.W., and P. Heriyati. 2023. "The Effect of Online Food Delivery Service Quality on Customer Satisfaction and Customer Loyalty: The Role of Personal Innovativeness." *Business and Entrepreneurial Review* 23(1):33–50.

Statista. 2020. *Change in Online Restaurant Delivery Penetration Share of the Restaurant Market in the United States Due to the Coronavirus Pandemic from 2020 to 2025*. Available online: https://www.statista.com/statistics/1170614/online-food-delivery-share-us-coronavirus/

Tandon, A., P. Kaur, Y. Bhatt, M. Mäntymäki, and A. Dhir. 2021. "Why Do People Purchase from Food Delivery Apps? A Consumer Value Perspective." *Journal of Retailing and Consumer Services* 63:102667.

Tsami, M., G. Adamos, E. Nathanail, E.B. Budiloviča, I.Y. Jackiva, and V. Magginas. 2018. "A Decision Tree Approach for Achieving High Customer Satisfaction at Urban Interchanges." *Transport and Telecommunication Journal* 19(3):194–202.

Vermeir, I., and G. Roose. 2020. "Visual Design Cues Impacting Food Choice: A Review and Future Research Agenda." *Foods* 9(10):1495.

Voorhees, C.M., P.W. Fombelle, Y. Gregoire, S. Bone, A. Gustafsson, R. Sousa, and T. Walkowiak. 2017. "Service Encounters, Experiences and the Customer Journey: Defining the Field and a Call to Expand Our Lens." *Journal of Business Research* 79:269–280.

Wang, J., X. Shen, X. Huang, and Y. Liu. 2021. "Influencing Factors of the Continuous Usage Intention of Consumers of Online Food Delivery Platform Based on an Information System Success Model." *Frontiers in Psychology* 12.

Wang, T.-C., R.-S. Guo, and C. Chen. 2023. "An Integrated Data-Driven Procedure for Product Specification Recommendation Optimization with LDA-LightGBM and QFD." *Sustainability* 15(18):13642.

Xie, M.-P., and W.-Y. Zhao. 2010. "The Analysis of Customers' Satisfaction Degree Based on Decision Tree Model." In *2010 Seventh International Conference on Fuzzy Systems and Knowledge Discovery* 6:2928–2931.

Xu, X. 2021. "What Are Customers Commenting on, and How is Their Satisfaction Affected? Examining Online Reviews in the On-Demand Food Service Context." *Decision Support Systems* 142:f113467.

Zhong, Y., and H.C. Moon. 2020. "What Drives Customer Satisfaction, Loyalty, and Happiness in Fast-Food Restaurants in China? Perceived Price, Service Quality, Food Quality, Physical Environment Quality, and the Moderating Role of Gender." *Foods* 9(4):460.

Zikmund, W.G., B.J. Babin, J.C. Carr, and M. Griffin. 2013. *Business Research Methods,* 9th ed. Boston, MA: Chengage Learning.

Appendix A: Survey Dimensions, Mean and Spearman Correlation Coefficient Values

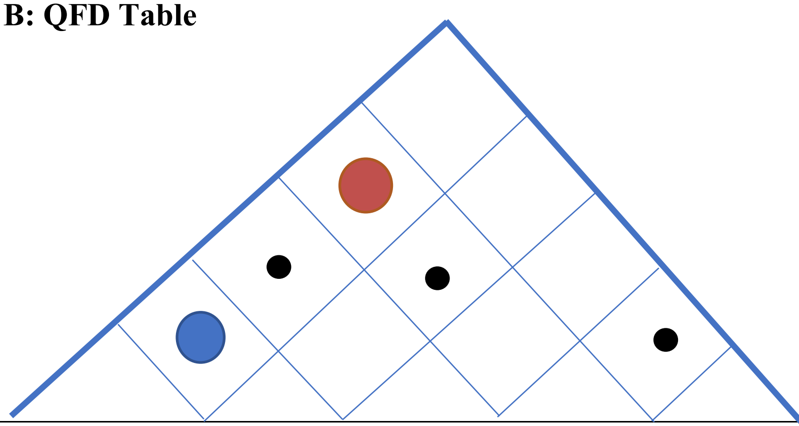
Survey Dimension	Codes and Items (Multiple Choices)
1. Demographic questions	a. Gender b. Age c. Level of education d. What is your current employment status? e. What is your marital status?
2. Consumer behavior questions	Q1 Which platform do you often use when you order food online? Q2 Please indicate the frequency with which you order per week. Q3 How much do you spend each time approximately?

	Codes and Items (all in 1-5 Likert scale, with 1 being the Lowest Satisfaction Level and 5 being the Highest)	Mean	SCC to Q21	SCC to Q22
3. Ordering platform design and operation	Q4 Please rate the overall appearance and structure of the online food ordering platform	4.36	0.589	0.554
	Q5 Please rate the accuracy and trustworthy of the information provided by the platform	4.26	0.496	0.486
	Q6 Please rate the adequacy of the information provided by the platform	4.21	0.525	0.453
	Q7 Please rate the ease of access to the platform	4.49	0.429	0.344
	Q8 Please rate the ease of using the platform	4.45	0.346	0.342
	Q9 Please rate your privacy protection when using the platform	4.13	0.331	0.353
	Q10 Please rate the pricing at the platform	3.88	0.455	0.457
	Q11 Please rate the ease of getting and using coupons, promotions and deals when using the platform	3.89	0.389	0.333
	Q12 Please rate the helpfulness of the food suggestions and recommendations by the platform	4.09	0.517	0.415
	Q13 Please rate the ease of tracking your orders	4.35	0.404	0.409
	Q14 Please rate the timeliness of your orders	4.08	0.497	0.416
4. Order delivery and tracking	Q15 Please rate the appearance of the delivery person	4.15	0.406	0.458
	Q16 Please rate the temperature of the food when you received your order	4.06	0.457	0.407
	Q17 Please rate the presentation of the food you received	4.09	0.51	0.5

Appendix A (cont.)

	Codes and Items (all in 1-5 Likert scale, with 1 being the Lowest Satisfaction Level and 5 being the Highest)	**Mean**	**SCC to Q21**	**SCC to Q22**
5. Online assistants	Q18 Please rate the performance of online assistants when using the platform	4.06	0.472	0.431
6. After-sales service	Q19 Please rate the platform's ability to resolve your complaints	3.77	0.641	0.519
	Q20 Please rate the handling of your refund request	3.82	0.569	0.524
7. Target (dependent variables)	Q21 Please rate you overall satisfaction with using the platform	4.08		0.623
	Q22 How likely are you to recommend the online food ordering platform you often use?	4.21	0.623	

Appendix B: QFD Table

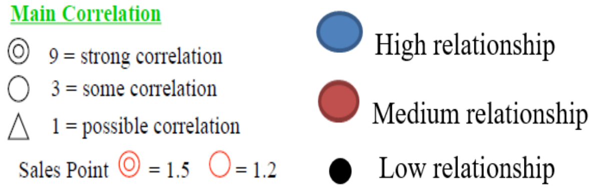


	1	2	3	4	5	6	7	8	9	10	11	12
Customer Needs \ Improvement Options	**Professional Training and Guidelines on Handling Customer Complaints**	**Optimize Refund Polices and Process**	**Improve Food Recommen-dation System and AI Algorithm**	**Offer More Coupons, Promotion Deal**	**Lower Food Delivery Fee**	**Rate of Importance**	**Company Now**	**Plan**	**Rate of Improve -ment**	**Sales Point**	**Absolute Weight**	**Demand Weight**
2. Food suggestion and recommendation (Q12)			◎ 342			5	4	5	1.25	◎	9.38	38
3. Platform's ability to resolve complaints (Q19)	◎ 189	○ 63				4	4	5	1.25		5.00	21
4. Getting and using coupons, promotions, and deals (Q11)				◎ 207	△ 23	3	4	5	1.25	◎	5.63	23

Appendix B (cont.)

	1	2	3	4	5	6	7	8	9	10	11	12
Customer Needs / Improvement Options	**Professional Training and Guidelines on Handling Customer Complain**	**Optimize Refund Polices and Process**	**Improve Food Recommen-dation System and AI Algorithm**	**Offer More Coupons, Promotion Deal**	**Lower Food Delivery Fee**	**Rate of Importance**	**Company Now**	**Plan**	**Rate of Improve-ment**	**Sales Point**	**Absolute Weight**	**Demand Weight**
5. Handling of refund request (Q20)	○ 54	◎ 162		△ 18		3	4	5	1.25	○	4.50	18
6. Total importance rating	243	225	342	225	23	1,058				Total	24.51	100
7. Percent	22.96	21.27	32.33	21.27	2.17	100						
8. Company situation now	Lack of professional training on handling customer complaints	Imperfect refund polices and process	Mediocre food recommen-dation system and AI algorithm	Offer coupons to customers quarterly	Relatively high delivery fees							
9. Action plans	Provide more professional training to staffs on customer services and handling complaints	Continuo usly improve refund polices and process	Upgrade food recommen-dation system and AI algorithm	Offer coupons to customers monthly	Cutting delivery fees							

Appendix B (cont.)



Col.11 = Col.6 x Col.9 x Col.10
Col. 9 = Col.8/Col.7 (Col stands for Column)

Food Distribution Research Society

Journal of Food Distribution Research
Volume 55, Issue 2, pp. 96–120

Effects of Organic and Origin Labels on Consumer Willingness to Pay for Kale: A Case Study in Southeastern United States

Joan U. Ureta[a] , Felipe de Figueiredo Silva[b], Michael M. Vassalos[c], and George Strickland[d]

[a]*Post-Doctoral Fellow, Forestry and Environmental Conservation,*
142 Jersey Ln., Clemson University,
Clemson, SC 29634, USA

[b]*Assistant Professor, Agriculture Sciences,*
821 McMillan Rd., Clemson University,
Clemson, SC 29634, USA

[c]*Associate Professor, Agriculture Sciences,*
821 McMillan Rd., Clemson University,
Clemson, SC 29634, USA

[d]*Graduate Student, Policy Studies,*
321 Calhoun Dr., Clemson University,
Clemson, SC 29634, USA

Abstract

As consumers become increasingly conscious of the health and environmental impacts of their dietary decisions, the demand for "superfoods" has surged. Using data from an online survey in the seven states in the southeastern United States and a choice experiment approach, this study investigated the effects of organic and product origin attributes on respondents' willingness to pay (WTP) for kale. A mixed logit in WTP space was utilized for the analysis. Results showed that respondents are willing to pay approximately a 35% premium for organic kale and a 27% premium for kale produced from the southeastern United States. Policy recommendations are also discussed.

Keywords: kale, choice experiment, WTP space, mixed logit, superfoods

Corresponding author: Email: joanu@g.clemson.edu

Introduction

During the last two decades, as consumers have become more interested in the health impact of their diets, several trends regarding food preferences have emerged. Among the most recent ones is the increased interest and marketing share of "superfoods" (Liu et al., 2021; Magrach and Sanz 2020; Cobos and Díaz, 2023). Generally, superfoods are considered and marketed as food products that are nutritionally dense and beneficial to a variety of health goals. Nevertheless, despite their popularity, there is currently no widely accepted legal definition of "superfoods" (Driessche, Plat, and Mensink, 2018; Liu et al., 2021; Franco Lucas et al., 2022).

Although a relatively rich literature regarding the factors affecting the consumption of functional foods exists (e.g., Pappalardo and Lusk, 2016; Plasek and Temesi, 2019; Szakály et al., 2019), to the best of our knowledge, the research regarding consumers' willingness to pay for superfoods such as kale in the United States is rather limited. This study is an effort to add to this research area by examining factors influencing consumer preferences for kale, a superfood product with a substantial increase in demand, as noted by several news outlets.[1] Moreover, Cobos and Díaz (2023) found that kale stands out as the most frequently mentioned superfood on websites due to its health-promoting properties.

The surge in demand for kale can be attributed to consumers' favorable perception of its numerous health advantages. As a low-calorie food with high levels of phytochemicals, vitamins, and minerals (Šamec, Urlić, and Salopek-Sondi, 2019), kale improves gut and metabolic health (Raychaudhuri et al., 2021; Thavarajah et al., 2016) and could be beneficial for preventing obesity (Reda et al., 2021). Traditionally, kale has been a natural remedy for treating stomach ulcers, diabetes mellitus, rheumatism, bone weakness, ophthalmologic problems, hepatic diseases, anemia, and obesity (Šamec, Urlić, and Salopek-Sondi, 2019). Recent studies found that kale supplementation could reduce risks of coronary artery disease (Kim et al., 2008), intestinal inflammation (Lima de Albuquerque et al., 2010), stomach ulcer (Lemos et al., 2011), cognitive decline and age-related oxidative damage (Kushimoto et al., 2018), and other diseases (Satheesh and Workneh Fanta, 2020). Alfawaz et al. (2022) found that more than 60% of their participants self-reported improvements in their health after adding kale to their diet.

In the United States, kale's domestic availability tripled in the last two decades and grew by 47% between 2020 and 2022 (USDA-ERS, 2023). California, South Carolina, New Jersey, Texas, and Georgia are the biggest kale producers, respectively (USDA-NASS, 2023). In terms of production practices, more than half of the kale sold in the United States is labeled as organic (Reda et al., 2021).

As a superfood that has gained popularity in western markets, several studies have examined kale and its attributes. Research conducted in the United States has centered predominantly on the sensory characteristics of kale (e.g., Swegarden et al., 2019). However, little is known about how labeling strategies (e.g., organic, place of origin) impact U.S. consumer preferences for kale. The extensive meta-analysis by Kilduff and Tregeagle (2022) also identified a limited number of

[1] *The New York Times* (Eddy, 2019) and Winsight (Sidrane 2015).

studies estimating consumers' willingnes s to pay (WTP) for organic and origin labels of leafy greens. To the best of our knowledge, studies of a similar nature were mostly conducted in Kenya (e.g., Ngigi et al., 2011; Lagerkvist et al., 2013). This study aims to bridge the knowledge gap by assessing the consumers' WTP for kale with value-added attributes. Following the literature on the estimation of WTP for fresh produce (Yue and Tong, 2009; Onozaka and McFadden, 2011), we estimate the WTP for kale with organic and origin attributes for consumers in seven states in the southeastern region[2] of the United States using a choice experiment approach.

Several studies have found that consumers are willing to pay a premium for organic produce (Bond, Thilmany, and Keeling Bond, 2008; Yue and Tong, 2009; Costanigro et al., 2014). Another strand of the literature found that consumers are also willing to pay for locally or regionally grown produce (Onozaka and McFadden, 2011; Gumirakiza and Choate, 2018). However, divergences exist. Kilduff and Tregeagle (2022) found that organic labels had no significant impact on WTP for fresh produce, whereas the local attribute increases WTP. Moreover, while the intersection of organic and local attributes was fairly well-studied in literature, consumers' WTP for fresh produce with organic and origin labels (i.e., country, region, or state) was much less explored.

As one of the superfoods, kale was selected because of its surging popularity in the United States (Thavarajah et al., 2016; Cobos and Díaz, 2023). We included organic and product origin attributes because they often indicate food preferences (Pappalardo and Lusk, 2016). The data for the study were obtained from an online survey of 199 consumers, and the mean WTP was estimated using mixed logit in WTP space. As a robustness check, the mean WTP was also obtained using a conditional logit and mixed logit in preference space. The secondary objective is to discuss in-depth the characteristics of kale consumers, such as purchase behavior and beliefs about organic and regional products.

We find that WTP estimates do not change substantially among the methods used. For the estimates using WTP space, results suggest that on the average, consumers are willing to pay a 35% premium for organic kale ($0.470 per bunch) and a 27% premium for regionally sourced kale ($0.359 per bunch). These estimates show that consumers are willing to pay a premium price for organic and regionally grown produce. Moreover, these estimates are similar to what has been found in the literature for other food products (Printezis, Grebitus, and Hirsch 2019; Li and Kallas, 2021).

Method

Data

The data for the study were obtained from an online survey distributed through Qualtrics to the southeastern region of the United States[3] in October 2022. The questionnaire was divided into six sections—screening questions, consumer grocery shopping habits, preferences for local options,

[2]States included are North Carolina, South Carolina, Georgia, Florida, Alabama, Mississippi, and Tennessee.
[3]States included are North Carolina, South Carolina, Georgia, Florida, Alabama, Mississippi, and Tennessee. As per interviews with kale producers, these states share homogeneous production practices for kale.

awareness and perception of kale, discrete choice experiment (DCE), and demographics. To reduce potential bias and ensure relatability of the choice experiment scenario, participation was restricted to individuals who met a set of criteria—being 18 years or older, serving as the primary grocery shopper for their household, purchasing vegetables monthly, and consuming kale at least once every quarter. Following Onozaka and McFadden (2011), our survey was restricted to kale buyers and consumers due to its online nature, precluding physical product inspection. Through this approach, we can ensure that respondents derive utility from consuming kale, with their additional WTP directly tied to organic and region of origin[4] labels. Prior to actual data collection, the survey was pretested with agricultural professionals and students, and 20 pilot responses were also gathered by Qualtrics to test for survey functionality. The final sample size included 199 participants.

An online platform was utilized considering that 83.69% of the households in the study site have broadband internet subscriptions (U.S. Census Bureau, 2023). Moreover, utilization of online surveys enabled us to randomize the sequence of choice cards presented to each respondent, thus avoiding an order effect (Carson et al., 1994). Lastly, we utilized the page break feature of Qualtrics to deter the respondents from reading ahead and comparing choice sets that should be evaluated independently (Champ, Boyle, and Brown, 2017).

In the DCE section, a description of USDA organic products was presented to the respondents (see Appendix 1). Given the potential for hypothetical bias in stated preferences studies, a cheap talk script was included to reduce this risk. This script reminded respondents of the tendency for consumers to overestimate their WTP when presented with hypothetical product descriptions (Champ, Boyle, and Brown, 2017). Each respondent was then asked to select among kale products with varying prices (i.e., $0.99, $1.33, or $1.67) [5] and different combinations of organic and region-of-origin attributes or indicate that they would not purchase the product (see Figure 1). For this study, "regional"[6] refers to kale that was produced in the southeastern United States. To help survey participants visualize this area, instead of a regional label, a map was included in the choice tasks. The use of real-world kale pricing and two uncorrelated attributes (organic, regional)[7] provided a choice experiment with reliably maximized marginal utility responses (Gao and Schroeder, 2009).

[4]A regional origin label was chosen over a local label, given that kale produced in the Southeast is generally marketed within the region and nationwide.
[5]Based on the USDA commodity reports, the lowest kale price in the Southeast during the development of the survey was $0.99, and the highest was $1.67 (February 2022), resulting in a mean of $1.33 per bunch (USDA Agricultural Marketing Service, 2022).
[6]Given that kale produced in the Southeast is typically marketed in the region and across the country, a place-of-origin label was chosen over a local label.
[7]The correlation analysis revealed that organic and local attributes have a weak correlation of 0.25.

Figure 1. A sample question from the discrete choice section of the survey, in which respondents indicate their purchase decision based on label information and price.

The choice sets were generated using Ngene.[8] In generating the choice sets, a condition was set to guarantee that the price of alternatives labeled as organic is consistently equal to or higher than the price of non-organic alternatives.[9] Given this condition, Ngene generated 12 choice sets with two alternatives resulting in a D-efficiency score of 89.16%. A no-purchase option was added as the third alternative to avoid conditional situation and to estimate the "true" demand (Louviere, Hensher and Swait 2000). To avoid survey fatigue, the samples were split into two groups, and each group was presented with only six choice tasks.

Empirical Strategy

McFadden's (1974) random utility model was utilized to evaluate the consumers' responses to organic and regional value-added attributes of kale. As shown in equation 1, the indirect utility (U) experienced by each individual ($I = 1, 2, \ldots, 199$) when choosing a product with $j = 1,2,3$ alternative in choice set ($n = 1,2,\ldots,6$)[10] is determined by a linear function of attributes (X_{ijn}):

[8]Ngene 1.3.0 was utilized for generating the choice sets.
[9]Although organic product prices may exhibit seasonal fluctuations, they consistently command a premium when compared to their conventional counterparts (USDA Economic Research Service, 2023).
[10]The model includes 3,582 observations (199 respondents x 6 choice tasks x 3 alternatives).

$$U_{ijn} = \boldsymbol{\beta}' X_{ijn} + e_{ijn} \tag{1}$$

where $\boldsymbol{\beta}$ represents the vector with unknown parameters of marginal utilities associated with the attributes of product $\boldsymbol{X}$ with alternate *j* in choice set *n*. The last term, e_{ijn}, denotes the random error of the computed utilities. The rational, utility-maximizing consumer has a choice probability of selecting alternative *j* in the *n*th choice set.

This study uses a fixed effects conditional logit[11] approach as a baseline model, similar to the approach used by Soley, Hu, and Vassalos (2019), Güney and Giraldo (2019), and Hu, Woods, and Bastin (2009). Conditional logit (CL) is especially well-suited for use in discrete choice experiments (DCE), accepting that the independent and identical distribution (IID) of error terms and independence of irrelevant alternatives (IIA) assumptions hold. CL assesses a binary dependent variable, such as a purchase or no-purchase choice option, modeled through a logistic regression (McFadden, 1974). The IIA assumption restricts the participant's substitution within the model, suggesting that the choice probabilities of one product relative to another must hold, regardless of the introduction of new alternatives. Under these conditions, the probability of alternative *j* being selected by individual *i*, in choice set *n* can be modeled using equation 2:

$$P_{ijt} = \frac{\exp(X_{ijn}\,\beta')}{\sum_{j=1}^{J} \exp\,(X_{ijn}\;\beta')} \tag{2}$$

where the coefficients $\boldsymbol{\beta}$ are weights that represent indefinite marginal utilities derived from different attributes of kale (i.e., organic, regional). Given the inherent limitations of the conditional logit, alternative regression models are employed to analyze DCE data (Train, 2009). Mixed logit (MXL) relaxes IIA assumption and accounts for preference heterogeneity by modeling the choice probability as follows:

$$P_{ijt} = \frac{\exp(X_{ijn}\,\beta')}{\sum_{j=1}^{J} \exp\left(X_{ijn}\;\beta'\right)} f(\beta|\theta)d\beta \tag{3}$$

where the density function of $\boldsymbol{\beta}$ is represented as $f(\beta|\theta)$ in which θ pertains to a parameter vector that characterizes the distribution of preferences in the population.

Two types of mixed logit models were utilized in this study—preference space (PS) and WTP space. In order to differentiate between these two models, equation 1 was expanded to emphasize that the utility derived by respondent *i* from selecting alternative *j* in a choice set *n* is a function of both the monetary $(Price_{ijn})$ and non-monetary (X_{ijn}) attributes of kale, resulting in the following expression:

$$U_{ijn} = \boldsymbol{\alpha} Price_{ijn} + \boldsymbol{\beta}' X_{ijn} + e_{ijn} \tag{4}$$

[11]Stata 15.1 was used to estimate the models using Clogit and mixlogit followed by WTP by Hole (2007), and mixlogitwtp by (Hole, 2016) for the mixed logit on willingness to pay space.

In the PS model, the utility coefficients are presumed to conform to a normal distribution, enabling the estimation of mean and standard deviation for each coefficient. Hence, the marginal WTPs for organic and origin attributes are then calculated using equation 5 (Louviere, Hensher, and Swait, 2000):

$$WTP_x = -\frac{\beta_x}{\alpha} \tag{5}$$

where α and $\boldsymbol{\beta}$ represent monetary and non-monetary coefficients, respectively.

Unlike the PS model, WTP is directly estimated in the WTP space model (Lim and Hu, 2023).[12] As such, the utility is further specified as:

$$U_{ijn} = \lambda(\boldsymbol{WTP}'X_{ijn} - Price_{ijn}) + e_{ijn} \tag{6}$$

where WTP is a vector of non-monetary parameters (e.g., organic, regional) with dollar units, and λ is a scale parameter. In equation 3, the θ in the density function of PS model contain α and $\boldsymbol{\beta}$, whereas WTP and λ for the WTP space model (Bazzani, Palma, and Nayga, 2018; Helveston, 2022) noted that the WTP coefficients generated by a PS model are prone to inaccurate interpretation due to the fixed specification of price and scale parameters. While a PS model specifies the price as a fixed parameter, suggesting that the standard deviation of unobserved utility remains constant across observations, the price/scale coefficient in a WTP space model can be considered random (Bazzani, Palma, and Nayga, 2018). Previous studies found that WTP space models outperform PS models in generating more stable and reasonable WTP estimates (Train and Weeks, 2005; Balcombe, Chalak, and Fraser, 2009; Thiene and Scarpa, 2009; Bazzani, Palma, and Nayga, 2018).

Results and Discussion

The demographic characteristics of our sample and a comparison with the 2021 American Community Survey are reported in Table 1. The distribution of the sample closely resembles that of the household population in the study site. Most of the respondents were from Florida, North Carolina, and Georgia. The difference in the average age of the sample and the population is attributed to the survey design, which excluded residents under the age of 18 from participating. The sample also included a higher proportion of females, likely due to the filtering of primary household grocery shoppers. This distribution is consistent with previous studies. For example, Fonner and Sylvia (2015) found that females represented 60% of shoppers, whereas Soley, Hu, and Vassalos (2019) found that females comprised 69% of their sample. Moreover, previous studies found that women are more likely to respond to web surveys than men (Keusch, 2015; Becker, 2022).

[12]Also see Train and Weeks (2005) and Scarpa, Thiene, and Train (2008) for discussion and initial applications of this method.

The demographic characteristics of the respondents are also comparable with the profile of "superfoodies," as determined by Franco Lucas et al. (2022) in their consumer segmentation study. For instance, compared to other clusters of consumers, superfoodies are mostly female, employed, and have a relatively higher household income.

Table 1. Sociodemographic Characteristics of the Respondents (n = 199)

Characteristics	Sample	Population*
Distribution of household population (%)		
Alabama	7.00	7.96
Florida	33.70	34.41
Georgia	15.10	17.06
Mississippi	4.50	4.66
North Carolina	17.60	16.67
South Carolina	13.10	8.20
Tennessee	9.00	11.02
Age (year)	47.64	39.64
Male (%)	28.14	48.79
Employed fulltime or part-time (%)	55.78	55.89**
Low income (<$25k/yr) (%)	22.11	20.90
Middle income ($25k -$50k/ yr) (%)	31.66	22.03
Homeowners (%)	67.34	68.47
Has a four-year degree or higher (%)	33.67	28.62
County resident for 5+ years (%)	78.39	Unknown
White (%)	66.83	69.03

Notes: Determined using state-level data from 2021 American Community Survey 1-year estimates.
**The percentage of individuals aged 16 years and above who are employed.

Awareness and Perception of Organic and Regional Produce

The majority of the respondents spend between $25 to $100 per month on purchasing fresh vegetables (see Table 2). Kale emerges as a favored choice as it is typically included in the weekly diet of 58% of the respondents. Among the varieties of kale available in the market, green kale is the most sought after variety as reported by 91% of the respondents. Moreover, more than half of the respondents prefer to buy organic kale grown in the Southeast, and a significant percentage indicated that they would be willing to pay a premium for these products (see Table 2).

Table 2. Respondents' Awareness and Perception of Organic Products ($N = 199$)

Consumer Behavior	Sample (%)
Spending habits on fresh vegetables	
Spends less than $25 per month on fresh vegetables	14.57
Spends between $25 and $100 per month on fresh vegetables	76.88
Spends more than $100 per month on fresh vegetables	8.54
Preferences of kale	
Eats kale at least once a week	58.29
Typically eats red kale	32.66
Typically eats green kale	91.46
Typically eats kale lacinato/Tuscan	15.58
Prefers organic kale	64.82
Prefers kale grown in southeastern United States	64.82
Willing to pay a premium for organic and local kale	61.31
Awareness and engagement with organic products	
Heard the term "organic food products"	98.49
Able to find organic produce at their regular stores	87.44
Shops at different location because of their organic food selection	58.79
Buys organic products at least once a month	80.90
USDA organic label seeking behavior	
Seek USDA organic label all the time	28.64
Seek USDA organic label most of the time	27.14
Seek USDA organic label sometimes	36.68
Never sought USDA label before	7.54
Perception of organic products	
Organic products are healthier or more nutritious.	73.87
Organic products taste better.	58.29
Organic products are more fresh.	59.30
Organic products are better for the environment.	69.85
Organic products contain no artificial ingredients and additives.	72.36
Organic products have less chemical or pesticide residue.	74.87
Organic products promote animal welfare.	63.32
Organic products are better for the health of farmers/farm workers.	67.84
Organic products support local farmers.	67.34

Almost all participants indicated that they were familiar with organic products, and organic produce is typically available at their regular grocery venue. When organic products are unavailable in their usual stores, 58.79% opt to shop elsewhere, primarily due to their wider selection of organic food. This amount is much higher than the 20% that was reported in the study of Govindasamy, DeCongelio, and Bhuyan (2006) in the northeastern United States.

In line with the prevailing perception of "superfoodies" (Franco Lucas, Costa, and Brunner, 2021; Franco Lucas et al., 2022), the majority of survey participants believed that organic products are healthier, have less chemical or pesticide residue, contain no artificial ingredients, and are more environmentally friendly than non-organic products. These findings align with the systematic review conducted by Katt and Meixner (2020), highlighting that consumers' environmental and health concerns drive their consumption of organic products. Also, more than half of the respondents believed that organic production is beneficial for local farmers and for the health of agricultural workers. This finding is consistent with the findings of Bond, Thilmany, and Keeling Bond (2008), which demonstrated that consumers' support for local farmers positively impacts their purchases of fresh produce.

Given the respondents' positive perceptions of organic products and their availability at their regular stores, 80.9% reported that they buy organic products at least once a month, with 34.67% buying them even on a weekly basis. When asked if they look for a USDA organic label when purchasing these products, 29% reported that they always seek this label, and very few reported that they never looked for this label before (7.54%). In total, 92.46% of the respondents indicated that they seek out the USDA label at least occasionally. This percentage is higher than the findings of McFadden and Huffman (2017) in the midwestern United States, where only 66% of respondents reported noticing the USDA organic seal prior to their study.

When it comes to product origin, 62.81% indicated that they prefer to buy fruits and vegetables grown from the Southeast over those from other regions. This finding is similar to the response by Hasselbach and Roosen (2015) regarding the local attribute, who found that 65% of consumers were conscious of the origin of the products they bought. It is noteworthy that only 61.31% of respondents were willing to pay a premium for both organic and regional (grown in the southeastern United States) attributes of kale.

Estimation Results

Table 3 presents the estimation results of the four models—conditional logit (Model 1), mixed logit in preference space (Model 2), and mixed logit in WTP space (Models 3 and 4). Although the CL model was presented, it should be noted that the Hausman test revealed that the IIA property is violated. This suggests that MXL estimation is a more suitable approach for the analysis. Based on the AIC and BIC values, WTP space model exhibited superior model fit when contrasted with conditional logit and mixed logit in preference space models. It should be noted that Models 1 and 2 present utility coefficients, whereas the utility in Models 3 and 4 is expressed in dollar units.

Table 3. Estimation Results

Attributes	**Model 1 (Coef. / S.E.)**	**Model 2 (Mean/S.E.)**	**Model 3 (Mean/S.E.)**	**Model 4 (Mean/S.E.)**
Price	-1.753***	-2.517***	1.128***	1.323***
	(0.186)	(0.244)	(0.152)	(0.158)
Organic	0.799***	1.161***	0.470***	0.375***
	(0.101)	(0.183)	(0.064)	(0.062)
Regional	0.644***	0.907***	0.359***	0.237***
	(0.070)	(0.119)	(0.043)	(0.046)
Organic*Regional				0.204*
				(0.081)
No purchase	-3.208***	-5.833***	-2.747***	-2.210***
	(0.246)	(0.526)	(0.311)	(0.185)
Standard deviation				
Price			0.975***	-0.819***
			(0.257)	(0.211)
Organic		1.693***	0.612***	0.530***
		(0.211)	(0.066)	(0.076)
Regional		1.041***	0.375***	0.272***
		(0.136)	(0.039)	(0.045)
Organic*Regional				-0.586***
				(0.095)
No purchase		2.795***	1.401***	0.914***
		(0.426)	(0.227)	(0.133)
AIC	2,104.841	1,912.118	1,902.239	1,887.591
BIC	2,129.576	1,955.404	1,951.708	1,949.428
Log likelihood	-1,048.4207	-949.05918	-943.11939	-933.79575
N	199	199	199	199
Observations	3,582	3,582	3,582	3,582

Notes: Model 1: Conditional logit; Model 2: Mixed logit in Preference space; Model 3 & 4: Mixed logit in WTP space.
Level of significance: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Across the models, price, organic, and product origin variables are statistically significant, confirming their influence on consumers' decisions regarding kale selection. A "No Purchase" variable was also added to the analysis, wherein 1 represents the third alternative (I would not buy), whereas 0 corresponds to alternatives A and B with different attribute levels. It is an alternative specific constant that holds across all choice sets, providing participants with the option of not purchasing any of the presented kale options. The significant negative coefficient associated with "No Purchase" indicates that the lack of a purchase by consumers significantly reduces their utility.

Table 4 presents the mean WTP for the organic and regional attributes of kale. Based on the point estimates computed through the three models (Models 1–3), respondents' mean WTP for organic kale ranges from $0.456 to $0.470. On average, this amount represents a 35% premium price over the average price of kale at $1.33 per bunch. The estimated WTP for regionally grown kale ranges from $0.359 to $0.368, which represents a 27% to 28% premium. The calculated premium for the origin attribute closely aligns with findings from Kilduff and Tregeagle (2022), who estimated a 28.39% premium for locally sourced sustainable food. However, it was lower than the results of Printezis, Grebitus, and Hirsch (2019), who identified a premium ranging from 41.4% to 52.5% for food products labeled as local, which could originate from nearby local, state, or regional sources. It is important to highlight that these studies employed meta-regression analysis and encompassed a diverse range of products, whereas our estimates were specifically derived for kale. Furthermore, our estimates were specifically focused on regionally sourced kale, whereas previous studies often link the local label with products grown or produced either within a defined distance, state, or region (Printezis, Grebitus, and Hirsch, 2019).

Table 4. Willingness to Pay Point Estimates for Organic and Origin Attributes of Kale

Attributes	**Mean WTP (95% confidence intervals) ($)**			
	Model 1	**Model 2**	**Model 3**	**Model 4**
Organic	0.456 (0.353–0.558)	0.461 (0.328–0.595)	0.470 (0.344–0.595)	0.375 (0.253–0.497)
Regional	0.368 (0.282–0.453)	0.360 (0.269–0.452)	0.359 (0.275–0.443)	0.237 (0.147–0.327)
Organic*Regional				0.204 (0.04–0.362)

Notes: Model 1: Conditional logit; Model 2: Mixed logit in preference space; Model 3 and 4: Mixed logit in WTP space.

In Model 4, we interacted the organic with the regional variable to estimate the respondents' mean WTP on kale that possesses both attributes. Following the approach of Meas et al. (2014) in computing for combined WTP,[13] our results showed that respondents are willing to pay $0.816 per bunch of organic and regionally grown kale, which represents a 61.35% premium. This WTP estimate exceeds the one calculated by Meas et al. (2014), which determined a $0.4 premium for an organic and regionally produced jar of blackberry jam. Additionally, Meas et al. (2014) found that the total premium for combined regional (produced in Ohio Valley) and organic attributes was lower than the sum WTP of individual attributes. On the contrary, we found that combining these two attributes generates a higher premium for kale, which suggests complementary effects between these two kinds of label.

At the conclusion of the discrete choice experiment, participants were asked to report the highest priority attribute of the choices they made. They reported that price (38.7%) and organic attribute (38.2%) were their main priorities, followed by the regional attribute with 23.1%. These findings show that organic labels draw a higher premium than regional labels. This result is consistent with the study of Kilduff and Tregeagle (2022), which reported that organic labels command higher WTP for sustainable food products than a local label. Previous literature has shown that these

[13]The aggregate WTP is derived from the sum of individual WTP values and the WTP associated with interaction variables (Meas et al., 2014).

attributes are not consistently prioritized, and, for any given study, the items and samples may cause one to have higher priority than another (Bond, Thilmany, and Keeling Bond; 2008; Gao and Schroeder, 2009; Hu, Woods, and Bastin, 2009; Yue and Tong, 2009; Li and Kallas, 2021).

Conclusions

Over the past two decades, alongside the rising demand for local and organic products, there has been a notable trend where consumers are increasingly eager to explore and purchase larger quantities of food items commonly referred to as "superfoods." Although there is not an official definition, superfoods are often considered as food products with high concentrations of nutrients or bioactive chemicals beneficial to human health (Liu et al., 2021; Franco Lucas et al., 2022). Consumers' growing desire for a healthier diet is a key driving force behind the increased popularity of superfoods. Although there is a rich literature regarding the chemical characteristics of superfoods (Franco Lucas, Costa, and Brunner, 2021), and substantial research efforts exist regarding consumers' WTP for functional foods, to the best of our knowledge, there is limited literature related to WTP for superfoods in the United States. This study is an effort to add to this literature.

Specifically, we estimated the WTP of consumers across seven states in the southeastern United States for value-added kale products that are produced using organic practices or come from farms within the region. Kale was selected because of a noticeable surge in demand within the United States and its growing popularity as a superfood (Cobos and Díaz, 2023). Using a discrete choice experiment, the results indicate that consumers have a significant positive response to both organic and origin attributes. Despite price being one of the major concerns of shoppers, both organic and origin attributes are able to draw premiums individually. Organic and origin attributes in kale draw premiums of 35% and 27%, respectively. Moreover, when these labels were shown together, they generated a combined premium of 61.35%. This result implies that as the demand for kale continues to grow, producers may consider shifting from conventional to organic farming practices and exploring regional distribution options, especially if they can do so at costs lower than the price premiums outlined in this analysis.

The estimates obtained in this study can help beginning farmers consider venturing into organic kale farming and distributing the produce regionally. Additionally, they provide essential guidance for Extension professionals who are pivotal in assisting producers with their decision-making processes. Premium-priced products, such as organic and regionally grown kale, can give producers more realized revenue (USDA-ERS, 2023). While our study contributes to the growing body of literature affirming the positive effects of the organic and origin labels on food prices, it is important to acknowledge the presence of conflicting results in the literature. Moreover, producers may not realize the price premiums stated by consumers in research studies, especially those derived from the stated preference approach (Davidson, Khanal, and Messer, 2023).

It should be noted that the findings of this study are limited to understanding the consumer preferences for organic and regionally grown kale in the southeastern United States. Consequently, these findings may not necessarily align with broader national perspectives on kale. Future studies should aim to capture an equitable representation of respondents from various southeastern states, given that our paper does not capture state-level differences. Since kale produced in the Southeast is marketed nationwide, future studies could also include a more diverse audience from various regions across the country. Moreover, as the top 5 kale producers span various U.S. regions (i.e., West, Southeast, Northeast, Southwest), exploring whether product origin impacts superfood preferences is valuable to inform targeted marketing strategies. It will also be interesting to explore additional superfood commodities and to conduct comparative analyses based on the findings of this study. Lastly, although we tried to control for hypothetical bias using cheap talk scripts, future studies could supplement the survey using field experiments with real products or online surveys coupled with improved visualization of alternatives. Yue and Tong (2009) and Lizin et al. (2022) found that using real products instead of pictures reduces any hypothetical bias in CE studies.

References

Alfawaz, H.A., K. Wani, H. Alrakayan, A.M. Alnaami, and N.M. Al-Daghri. 2022. "Awareness, Knowledge and Attitude towards 'Superfood' Kale and Its Health Benefits among Arab Adults." *Nutrients* 14(2):245.

Balcombe, K., A. Chalak, and I. Fraser. 2009. "Model Selection for the Mixed Logit with Bayesian Estimation." *Journal of Environmental Economics and Management* 57(2):226–237.

Bazzani, C., M.A. Palma, and R.M.J. Nayga. 2018. "On the Use of Flexible Mixing Distributions in WTP Space: An Induced Value Choice Experiment." *Australian Journal of Agricultural and Resource Economics* 62(2):185–198.

Bond, C.A., D. Thilmany, and J. Keeling Bond. 2008. "Understanding Consumer Interest in Product and Process-Based Attributes for Fresh Produce." *Agribusiness* 24(2):231–252.

Becker, R. 2022. "Gender and Survey Participation: An Event History Analysis of the Gender Effects of Survey Participation in a Probability-based Multi-wave Panel Study with a Sequential Mixed-mode Design." *Methods, Data, Analyses* 16(1):30.

Carson, R.T., J.J. Louviere, D.A. Anderson, P. Arabie, D.S. Bunch, D.A. Hensher, R.M. Johnson, W.F. Kuhfeld, D. Steinberg, J. Swait, H. Timmermans, and J.B. Wiley. 1994. "Experimental Analysis of Choice." *Marketing Letters* 5(4):351–367.

Champ, P.A., K.J. Boyle, and T.C. Brown. 2017. *The Economics of Non-Market Goods and Resources: A Primer on Nonmarket Valuation*. New York, NY: Springer.

Cobos, Á., and O. Díaz. 2023. "'Superfoods': Reliability of the Information for Consumers Available on the Web." *Foods* 12(3):546.

Costanigro, M., S. Kroll, D. Thilmany, and M. Bunning. 2014. "Is It Love for Local/Organic or Hate for Conventional? Asymmetric Effects of Information and Taste on Label Preferences in an Experimental Auction." *Food Quality and Preference* 31:94–105.

Davidson, K.A., B. Khanal, and K.D. Messer. 2023. "Are Consumers No Longer Willing To Pay More for Local Foods? A Field Experiment." *Agricultural and Resource Economics Review*:1–21.

Driessche, J.J. van den, J. Plat, and R.P. Mensink. 2018. "Effects of Superfoods on Risk Factors of Metabolic Syndrome: A Systematic Review of Human Intervention Trials." *Food & Function* 9(4):1944–1966.

Eddy, M. 2019. "Where Kale Is King (at Least, When It's Stewed in Schmaltz and Bacon)." *The New York Times*. [New York, NY]. Available online: https://www.nytimes.com/2019/02/26/world/europe/germany-kale-festival.html.

Fonner, R., and G. Sylvia. 2015. "Willingness to Pay for Multiple Seafood Labels in a Niche Market." *Marine Resource Economics* 30(1):51–70.

Franco Lucas, B., J.A.V. Costa, and T.A. Brunner. 2021. "Superfoods: Drivers for Consumption." *Journal of Food Products Marketing*. Available online: https://www.tandfonline.com/doi/full/10.1080/10454446.2020.1869133 [Accessed September 19, 2023].

Franco Lucas, B., F. Götze, J.A. Vieira Costa, and T.A. Brunner. 2022. "Consumer Perception Toward 'Superfoods': A Segmentation Study." *Journal of International Food & Agribusiness Marketing*:1–19.

Gao, Z., and T.C. Schroeder. 2009. "Effects of Label Information on Consumer Willingness-to-Pay for Food Attributes." *American Journal of Agricultural Economics* 91(3):795–809.

Govindasamy, R., M. DeCongelio, and S. Bhuyan. 2006. "An Evaluation of Consumer Willingness to Pay for Organic Produce in the Northeastern U.S." *Journal of Food Products Marketing* 11(4):3–20.

Güney, O.I., and L. Giraldo. 2019. "Consumers' Attitudes and Willingness to Pay for Organic Eggs: A Discrete Choice Experiment Study in Turkey." *British Food Journal* 122(2):678–692.

Gumirakiza, J.D., and T. Choate. 2018. "The Willingness to Pay for Local, Domestic, and Imported Bundled Fresh Produce by Online Shoppers." *Journal of Agricultural Science* 10(12):15.

Hasselbach, J.L., and J. Roosen. 2015. "Consumer Heterogeneity in the Willingness to Pay for Local and Organic Food." *Journal of Food Products Marketing* 21(6):608–625.

Helveston, J.P. 2022. “logitr: Fast Estimation of Multinomial and Mixed Logit Models with Preference Space and Willingness to Pay Space Utility Parameterizations.” *ArXiv*. Available online: http://arxiv.org/abs/2210.10875.

Hole, A.R. 2007. “Fitting Mixed Logit Models by Using Maximum Simulated Likelihood.” *Stata Journal* 7(3):388–401.

Hole, A.R. 2016. “Mixlogitwtp: Stata Module to Estimate Mixed Logit Models in Wtp Space.” *Statistical Software Components*. Available online: https://ideas.repec.org//c/boc/bocode/s458037.html.

Hu, W., T. Woods, and S. Bastin. 2009. “Consumer Acceptance and Willingness to Pay for Blueberry Products with Nonconventional Attributes.” *Journal of Agricultural and Applied Economics* 41(1):47–60.

Katt, F., and O. Meixner. 2020. “A Systematic Review of Drivers Influencing Consumer Willingness to Pay for Organic Food.” *Trends in Food Science & Technology* 100:374–388.

Keusch, F. 2015. “Why Do People Participate in Web Surveys? Applying Survey Participation Theory to Internet Survey Data Collection.” *Management Review Quarterly* 65(3):183–216.

Kilduff, A., and D. Tregeagle. 2022. “Willingness-to-Pay for Produce: A Meta-Regression Analysis Comparing the Stated Preferences of Producers and Consumers.” *Horticulturae* 8(4):290.

Kim, S.Y., S. Yoon, S.M. Kwon, K.S. Park, and Y.C. Lee-kim. 2008. “Kale Juice Improves Coronary Artery Disease Risk Factors in Hypercholesterolemic Men.” *Biomedical and Environmental Sciences* 21(2):91–97.

Kushimoto, S., Y. Uchibori, S. Yanai, H. Makabe, S. Nakamura, and S. Katayama. 2018. “Kale Supplementation Up-Regulates Hsp70 and Suppresses Cognitive Decline in a Mouse Model of Accelerated Senescence.” *Journal of Functional Foods* 44:292–298.

Lagerkvist, C.J., S. Hess, J. Okello, and N. Karanja. 2013. “Consumer Willingness to Pay for Safer Vegetables in Urban Markets of a Developing Country: The Case of Kale in Nairobi, Kenya.” *The Journal of Development Studies* 49(3):365–382.

Lemos, M., J.R. Santin, L.C.K. Júnior, R. Niero, and S.F. de Andrade. 2011. “Gastroprotective Activity of Hydroalcoholic Extract Obtained from the Leaves of Brassica Oleracea Var. Acephala Dc in Different Animal Models.” *Journal of Ethnopharmacology* 138(2):503–507.

Li, S., and Z. Kallas. 2021. “Meta-analysis of Consumers’ Willingness To Pay for Sustainable Food Products.” *Appetite* 163:105239.

Lim, K.H., and W. Hu. 2023. “Contextual Reference Price in Choice Experiments.” *American Journal of Agricultural Economics* 105(4):1288–1306.

Lima de Albuquerque, C., M. Comalada, D. Camuesco, M.E. Rodríguez-Cabezas, A. Luiz-Ferreira, A. Nieto, A.R. Monteiro de Souza Brito, A. Zarzuelo, and J. Gálvez. 2010. “Effect of Kale and Papaya Supplementation in Colitis Induced by Trinitrobenzenesulfonic Acid in the Rat.” *e-SPEN, the European e-Journal of Clinical Nutrition and Metabolism* 5(3):e111–e116.

Liu, H., Y. Meng-Lewis, F. Ibrahim, and X. Zhu. 2021. “Superfoods, Super Healthy: Myth or Reality? Examining Consumers’ Repurchase and Wom Intention Regarding Superfoods: A Theory of Consumption Values Perspective.” *Journal of Business Research* 137:69–88.

Lizin, S., S. Rousseau, R. Kessels, M. Meulders, G. Pepermans, S. Speelman, M. Vandebroek, G. Van Den Broeck, E.J. Van Loo, and W. Verbeke. 2022. “The State of the Art of Discrete Choice Experiments in Food Research.” *Food Quality and Preference* 102:104678.

Louviere, J.J., D.A. Hensher, and J.D. Swait. 2000. *Stated Choice Methods: Analysis and Applications*. Cambridge, United Kingdom: Cambridge University Press.

Magrach, A., and M.J. Sanz. 2020. “Environmental and Social Consequences of the Increase in the Demand for ‘Superfoods’ World-Wide.” *People and Nature* 2(2):267–278.

McFadden, D. 1974. “Conditional Logit Analysis of Qualitative Choice Behavior.” In P. Zarembka, ed. *Frontiers in Econometrics*. New York, NY: Academic Press, pp. 105–142.

McFadden, J.R., and W.E. Huffman. 2017. “Willingness-to-Pay for Natural, Organic, and Conventional Foods: The Effects of Information and Meaningful Labels.” *Food Policy* 68:214–232.

Meas, T., W. Hu, M. Batte, T.A. Woods, and S. Ernst. 2014. “Substitutes or Complements? Consumer Preference for Local and Organic Food Attributes.” *American Journal of Agricultural Economics* 97(4):1044–1071

Ngigi, M.W., J.J. Okello, C.L. Lagerkvist, N.K. Karanja, and J. Mburu. 2011. “Urban Consumers’ Willingness to Pay for Quality of Leafy Vegetables along the Value Chain: The Case of Nairobi Kale Consumers, Kenya.” *International Journal of Business and Social Science* 2(7):208–216.

Onozaka, Y., and D.T. McFadden. 2011. “Does Local Labeling Complement or Compete with Other Sustainable Labels? A Conjoint Analysis of Direct and Joint Values for Fresh Produce Claim.” *American Journal of Agricultural Economics* 93(3):693–706.

Pappalardo, G., and J.L. Lusk. 2016. “The Role of Beliefs in Purchasing Process of Functional Foods.” *Food Quality and Preference* 53:151–158.

Plasek, B., and Á. Temesi. 2019. "The Credibility of the Effects of Functional Food Products and Consumers' Willingness to Purchase/Willingness to Pay–Review." *Appetite* 143:104398.

Printezis, I., C. Grebitus, and S. Hirsch. 2019. "The Price Is Right!? A Meta-Regression Analysis on Willingness to Pay for Local Food" *PLOS ONE* 14(5):e0215847.

Raychaudhuri, S., S. Fan, O. Kraus, M. Shahinozzaman, and D.N. Obanda. 2021. "Kale Supplementation during High Fat Feeding Improves Metabolic Health in a Mouse Model of Obesity and Insulin Resistance." *PLOS ONE* 16(8):e0256348.

Reda, T., P. Thavarajah, R. Polomski, W. Bridges, E. Shipe, and D. Thavarajah. 2021. "Reaching the Highest Shelf: A Review of Organic Production, Nutritional Quality, and Shelf Life of Kale (brassica Oleracea Var. Acephala)." *Plants People Planet* 3(4):308–318.

Šamec, D., B. Urlić, and B. Salopek-Sondi. 2019. "Kale (Brassica oleracea var. acephala) As a Superfood: Review of the Scientific Evidence Behind the Statement." *Critical Reviews in Food Science and Nutrition* 59(15):2411–2422.

Satheesh, N., and S. Workneh Fanta. 2020. "Kale: Review on Nutritional Composition, Bio-Active Compounds, Anti-Nutritional Factors, Health Beneficial Properties and Value-Added Products" *Cogent Food & Agriculture* 6(1):1811048.

Scarpa, R., M. Thiene, and K. Train. 2008. "Utility in Willingness to Pay Space: A Tool to Address Confounding Random Scale Effects in Destination Choice to the Alps." *American Journal of Agricultural Economics* 90(4):994–1010.

Sidrane, A. 2015. "National Kale Day Becomes a Movement." *Winsight Grocery Business*.

Soley, G., W. Hu, and M. Vassalos. 2019. "Willingness to Pay for Shrimp with Homegrown by Heroes, Community-Supported Fishery, Best Aquaculture Practices, or Local Attributes." *Journal of Agricultural and Applied Economics* 51(4):606–621.

Swegarden, H., A. Stelick, R. Dando, and P.D. Griffiths. 2019. "Bridging Sensory Evaluation and Consumer Research for Strategic Leafy Brassica (Brassica oleracea) Improvement." *Journal of Food Science* 84(12):3746–3762.

Szakály, Z., S. Kovács, K. Pető, P. Huszka, and M. Kiss. 2019. "A Modified Model of the Willingness to Pay for Functional Foods." *Appetite* 138:94–101.

Thavarajah, D., T. Lawrence, S. Powers, B. Jones, N. Johnson, J. Kay, A. Bandaranayake, E. Shipe, and P. Thavarajah. 2021. "Genetic Variation in the Prebiotic Carbohydrate and Mineral Composition of Kale (Brassica oleracea L. var. acephala) Adapted to an Organic Cropping System." *Journal of Food Composition and Analysis* 96:103718.

Thavarajah, D., P. Thavarajah, A. Abare, S. Basnagala, C. Lacher, P. Smith, and G.F. Combs. 2016. "Mineral Micronutrient and Prebiotic Carbohydrate Profiles of Usa-Grown Kale (Brassica oleracea L. var. acephala)." *Journal of Food Composition and Analysis* 52:9–15.

Thiene, M., and R. Scarpa. 2009. "Deriving and Testing Efficient Estimates of WTP Distributions in Destination Choice Models." *Environmental and Resource Economics* 44(3):379–395.

Train, K., and M. Weeks. 2005. "Discrete Choice Models in Preference Space and Willingness-to-Pay Space." In R. Scarpa and A. Alberini, eds. *Applications of Simulation Methods in Environmental and Resource Economics*. The Economics of Non-Market Goods and Resources. Berlin/Heidelberg, Germany: Springer-Verlag, pp. 1–16.

Train, K.E. 2009. *Discrete Choice Methods with Simulation*. Cambridge, United Kingdom: Cambridge University Press.

U.S. Census Bureau. 2023. *QuickFacts*. Washington, DC: U.S. Department of Commerce, Bureau of the Census. Available online: https://www.census.gov/quickfacts/fact/table/FL,GA,SC,NC,VA,WV/INC110221

U.S. Department of Agriculture. 2022. "Growth in the Number of U.S. Farmers Markets Slows in Recent Years." *Data Products*. Washington, DC: U.S. Department of Agriculture, Economic Research Service.

U.S. Department of Agriculture. 2023. "Organic Agriculture." *Natural Resources and Environment*. Washington, DC: U.S. Department of Agriculture, Economic Research Service.

U.S. Department of Agriculture. 2023. *Quick Stats*. Washington, DC: U.S. Department of Agriculture, National Agricultural Statistics Service.

Yue, C., and C. Tong. 2009. "Organic or Local? Investigating Consumer Preference for Fresh Produce Using a Choice Experiment with Real Economic Incentives." *HortScience* 44(2):366–371.

Appendix

Appendix 1. A redacted version of the questionnaire.

1. Which state do you currently reside in?

2. What is your five-digit zip code?

3. What is your year of birth?

4. Are you the primary grocery shopper of your household (at least 50% of the time)?

☐ Yes ☐ No

5. Do you buy vegetables every month? ☐ Yes ☐ No

6. On average, how much do you usually spend on fresh vegetables per month?

☐ Less than $25 ☐ $76- $100

☐ $25- $50 ☐ More than $100

☐ $51- $75

7. How often do you eat kale?

☐ Daily ☐ Once a month

☐ 4-6 times a week ☐ Once a quarter

☐ 2-3 times a week ☐ Rarely or Never

☐ Once a week

8. Have you heard the term 'organic food products'?

☐ Yes ☐ No

9. Is organic produce available at the stores where you typically buy your groceries?

☐ Yes ☐ No ☐ Sometimes ☐ I do not know

10. Please indicate how often you purchase organic food products.

☐ Weekly ☐ Once every quarter

☐ Twice a month ☐ Rarely

☐ Once a month ☐ Never

11. Do you shop at different location if you are specifically seeking out organic food products?

☐ Yes ☐ No

12. Do you look for this label when purchasing products?

☐ Never ☐ Sometimes ☐ Most of the time ☐ Always

13. What are your reasons for purchasing organic products? Please select your level of agreement to each statement.

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
Organic products are healthier or more nutritious.					
Organic products taste better.					
Organic products are more fresh.					
Organic products are better for the environment.					

Organic products contain no artificial ingredients and additives.	○	○	○	○	○
Organic products have less chemical or pesticide residue.	○	○	○	○	○
Organic products promote animal welfare.	○	○	○	○	○
Organic products are better for the health of farmers/ farm workers.	○	○	○	○	○
Organic products support local farmers.	○	○	○	○	○

14. Do you prefer to buy fruits and vegetables that were grown from Southeastern US?

☐ Yes ☐ Sometimes ☐ No ☐ Product origin does not matter to me

15. What are the varieties of kale that you typically eat? (please check all that apply)

☐ Red kale ☐ Green kale ☐ Kale Lacinato/ Tuscan

16. Do you prefer eating organic kale?

☐ Yes ☐ No ☐ Method of production does not matter to me

17. Do you prefer to eat kale grown from Southeastern US?

☐ Yes ☐ No ☐ Source of kale does not matter to me

In this section, we would like to know your willingness to pay for locally-sourced and organic kale. Please read the following information on how USDA defines organic products and local food.

Organic Products

According to USDA, organic products have to meet the following requirements to be certified as "USDA organic":

Must be produced using agricultural production practices that foster resource recycling, promote ecological balance, maintain and improve soil and water quality, minimize the use of synthetic materials, and conserve biodiversity.

Products must be:

Overseen by a USDA NOP (National Organic Program)-authorized certifying agent, following all USDA organic regulations

Produced without excluded or prohibited methods, (e.g., genetic engineering, ionizing radiation, or sewage sludge)

Produced using allowed substances

The experience from previous similar surveys is that people often state a higher willingness to pay than what one is actually willing to pay for the good. For instance, a recent study asked people whether they would purchase a new food product similar to the one you are about to be asked about. This purchase was hypothetical (as it will be for you) in that no one actually had to pay money when they indicated a willingness to purchase. In the study, 80% of people said they would buy the new product, but when a grocery store actually stocked the product, only 43% of people actually bought the new product when they had to pay for it. This difference (43% vs. 80%) is what we refer to as hypothetical bias. Accordingly, it is important that you make each of your upcoming selections like you would if you were actually facing these exact choices in a store, i.e., noting that buying a product means that you would have less money available for other purchases.

18. Would you be willing to pay a premium for **organic kale grown in the Southeastern US**?

☐ Yes ☐ No

19. Assume that you are willing to purchase green kale. Given the set of choices below, please select your most preferred option:

Attribute	Option 1	Option 2
Method of production	USDA ORGANIC	USDA ORGANIC
Product origin		Not from the Southeast
Price per bunch	$0.99	$1.33

☐ I prefer Option 1 ☐ I prefer Option 2 ☐ I would buy neither

(Note: Six choice sets were shown to each respondent)

20. When you were deciding which option to choose, what was the most important attribute that you considered?

☐ Organic ☐ Product origin ☐ Price

21. What is your gender?

☐ Male ☐ Female ☐ non-binary ☐ Prefer not to say

22. What ethnicity do you most identify with?

☐ White

☐ Native Hawaiian or Pacific Islander

☐ Black or African American

☐ Hispanic or Latino or Spanish Origin of any race

☐ American Indian or Alaska Native

☐ Other

☐ Asian

23. What is the highest level of education you have completed?

☐ Some high school or less

☐ High school diploma or GED

☐ Some college, but no degree

☐ Associates or technical degree

☐ Bachelor's degree

☐ Graduate or professional degree (MA, MS, MBA, PhD, JD, MD, DDS etc.)

☐ Prefer not to say

24. Do you have any agricultural background?
☐ Yes, we are managing a farm.

☐ Yes, I earned a degree related to agriculture.

☐ Yes, I worked in an agricultural-related company.

☐ Yes (please specify) ____

☐ No

25. How many people are currently living in your household?

26. Which of the following categories best describes your employment status?

☐ Employed full time ☐ Student

☐ Employed part time ☐ Disabled

☐ Self-employed ☐ Unemployed

☐ Retired

27. What was your total household income before taxes during the past 12 months?

☐ Less than $25,000 ☐ $100,000-$149,999

☐ $25,000-$49,999 ☐ $150,000 or more

☐ $50,000-$74,999 ☐ Prefer not to say

☐ $75,000-$99,999