ECONOMIC AND DEMOGRAPHIC FACTORS AFFECTING THE CONSUMER DEMAND FOR SUPERFRUIT BEVERAGES IN THE UNITED STATES

Rachel Victoria
Department of Agricultural Economics
Texas A&M University
College Station, TX 77843-2124
rachelmaria@tamu.edu

Senarath Dharmasena
Department of Agricultural Economics
Texas A&M University
College Station, TX 77843-2124
sdharmasena@tamu.edu

Selected Paper prepared for presentation at the Southern Agricultural Economics Association's 2016 Annual Meeting, San Antonio, Texas, February 6-9, 2016

ABSTRACT

Economic and Demographic Factors Affecting the Consumer Demand for Superfruit Beverages in the United States

There are many different types of nonalcoholic beverages available in the United States today compared to decade ago. Functionality and health dimensions of beverages have changed over the years. Currently, exotic superfruit beverages are serving as a healthier alternative to more traditional fruit juices, such as orange juice and apple juice, which contain high levels of sugar. Superfruit beverages strong emergence in the marketplace has created a major competition with traditional beverages and is providing consumers an alternative. Knowledge of price sensitivity, substitutes/complements and demographic profiling with respect to consumption of superfruit beverages is important for manufacturers, retailers, advertisers, nutritionists and other stakeholders from a competitive intelligence perspective as well as from a strategic decision-making perspective. Using nationally representative household level data from nearly 65,000 households, factors affecting the consumer demand for superfruit beverages was estimated. Moreover, we estimated own- and cross-price and income elasticities for superfruit beverages delineated by selected demographic segments, hence determine the competitiveness of superfruit beverages vis-à-vis conventional apple and orange juice. This information will reflect the market competitiveness and profiles of demographics consuming superfruit beverages in the United States.

Keywords: Superfruit beverages, Nielsen data, consumer demand
JEL Classification: D11, D12

INTRODUCTION

There are a variety of beverages offered in the United States beverage market in both alcoholic and non-alcoholic form. In response to a growing health trend brought on by both millennial and boomer generations, the beverage market has responded by providing an increasing amount of beverage options that have been functionally altered in terms of health dimensions (Landi). Some of these health dimensions include beverages rich in antioxidants, vitamins, minerals, and beneficial omega-3 fatty acids ("Superfoods to the…").

Superfruit is defined as any fruit supposed to confer remarkable health benefits (Gross).Currently, exotic superfruit beverages are serving as a healthier alternative to more traditional fruit juices, such as orange juice and apple juice, which contain high levels of high fructose corn syrup and calories ("Fruit and Vegetable…"). Superfruit beverages strong emergence in the marketplace has created a major competition with traditional beverages and is providing consumers an alternative. To strengthen the position of this, a massive decline in the consumption of liquid and frozen orange juice occurred between 2007 and 2012 ("Fruit and Vegetable…"). At the same time, Superfruit juices, such as POM Wonderful pomegranate juice, filled up an increasing amount of retail shelf space at a rapid rate throughout the United States ("Fruit and Vegetable…"). As a result of this new competitive beverage category, producers of more traditional beverages including alcoholic beverages have begun to incorporate small amounts of superfruits into their products creating a variety of new flavors for consumers (Landi). Some producers have even gone as far as falsely advertising their beverages in the same light as their competitors in terms of health benefits and ingredients. For example, Purely

Juice was found guilty of deceiving customers with adulterated pomegranate juice and falsely advertising that their product was made from 100 percent pomegranate juice even though the formulation primarily consisted of cane sugar and corn sweeteners ("POM Wonderful Announces…"). This increase in demand for superfruit beverages could probably be due to change in consumer perception as well as the presence of a wide array of fruit juice alternatives and new exotic flavors now available in the market.

The change in consumer demand pertaining to the beverage market as a whole has caused institutions, such as schools, to alter their retailed beverage option s. For example, in early 2010, the Alliance School Beverage Guidelines Final Progress Report was issued mandating the elimination of high in sugar and full- calorie drink options in schools ("240_School Beverage…"). This mandate is a response to parents concerns for their children's' health and bad dietary practices while at school ("240_School Beverage…"). The mandate has successfully altered the beverage landscape in schools across the country, which is notable by the 90 percent decrease in beverage calories shipped to schools ("240_School Beverage…"). Soft drinks, energy drinks, and high in added sugar fruit juices are now either offered in smaller quantities and proportions and or substituted with 100 percent fruit juices and water ("240_School Beverage…"). These actions in turn create a massive market share loss in schooling institutions.

In response significant market share losses, competitors of superfruit beverages are altering their business practices to stay afloat. These competitors are doing so by investing in superfruit beverage companies and buying a hefty amount of their stocks ("State of the…"). Drink Maple,

Concord, Mass., is one such company trying to make a splash in the plant water category ("State of the…"). Also, the Coca-Cola Co. took a minority stake in Suja Life L.L.C., San Diego, a manufacturer of organic, cold-pressed juice products ("State of the…"). The transaction is expected to increase distribution of Suja's products and improve operational efficiencies ("State of the…"). In addition to these investment practices, competitors of superfruit beverages are starting to offer "clean" versions of their popular beverage products, meaning that they are free of artificial colors, flavors, sweeteners and preservatives ("State of the…"). The clean beverage movement's goal is to reduce the negative health impacts that occur from consuming unclean products in an effort to be stronger competition against superfruit beverages.

Some of the top superfruit beverages that are gaining rapid growth in the market are acai berry, cranberry, coconut, elderberry, and goji berry (Reuteman). Additionally, Pomegranate continues to maintain the superfruit market share, "… account[ing] for more than 40 percent of tracked beverage launches featuring superfruit flavors from June 2008 to May 2013, ahead of açai and lychee with 12.5 percent and 12 percent, respectively, according to Innova data" ("Healthy Flavors Boom…"). Aside from overall market share data, there is a lack of data that suppliers consider necessary for making effective business decisions. In other words, superfruit's new emergence in the beverage market has led to a marketplace that is operating blindly do to a lack of economic and quantitative data for competitors to access.

By the same token, it is crucial for fruit beverage producers to understand the economic impacts of emerging superfruit beverages in the marketplace. Growth in fruit juice alternatives

has been attributed to improved health-related claims and consumer perceptions, a flurry of brands, appealing and convenient packaging, and a plethora of flavors available. This increasing demand for fruit alternative beverages and declining demand for traditional high in sugar beverages in the United States could negatively affect non-superfruit beverage producers in terms of low prices for more traditional fruits juices as well as reduced income. Therefore, it is of interest for non-superfruit beverage producers in the United States to know the competitiveness and elasticities of fruit drink alternatives in the beverage marketplace and their implications on fruit drink prices and supplier income.

OBJECTIVES

The general objective of the study is to determine socio-economic-demographic factors affecting demand for superfruit beverages in the United States. The specific objectives are to (1) determine own-price, cross-price and income elasticities of demand for superfruit beverages; and (2) determine demographic factors affecting demand for superfruit beverages in the United States.

DATA AND METHODS

The data for this analysis will be obtained from 201 1 Nielsen Homescan Consumer Panel. Household expenditure, quantity, and other data on the amount superfruit beverage consumption will be collected by Nielsen Homescan in 2011 and contains approximately 65,000 household level data points. Statistical Analysis Software (SAS 9.4) will be used to analyze the data. A simple demand model will be estimated for those households who purchased superfruit beverages. Various functional forms, such as log-log, log-linear, linear-log will be

tested to find the best fitting model. Also, other than quantity and price for these beverage products, host of demographic characteristics such as household income, household size (the number of people per house), age, region, race, ethnicity, education status, age and presence of children, and location in the United States will be used as additional covariates.

Model Development, Procedures and Variables

Choice to purchase or not to purchase superfruit beverages could be affected by price, income and various demographic factors. This type of choice is a dichotomous discrete (buy or not-to buy or “one” if buy and “zero” if do not buy) and a probit model is used generally to model such a choice decision. The dependent variable is a zero one type dummy variable which is created to reflect the non-purchase or purchase respectively of superfruit beverages. It is regressed on price and a host of demographic factors. Probit analysis will provide statistically significant findings of the decision to purchase superfruit beverages.

Demographic and economic factors hypothesized to be affecting the decision to buy superfruit bcheverages are listed on Table 1. Also, we provide different categories used in each factor along with base category for dummy variables.

The probit model for superfruit beverages can be written as follows:

$$
\begin{aligned}
&\Pr(Y=1 \mid x_i'\beta) = \beta_1 + \beta_2 PRICE_i + \beta_3 AGEHH2529_i + \beta_4 AGEHH3034_i + \\
&\beta_5 AGEHH3544_i + \beta_6 AGEHH4554_i + \beta_7 AGEHH5564_i + \beta_8 AGEHHGT64_i + \\
&\beta_9 EMPHHPT_i + \beta_{10} EMPHHFT_i + \beta_{11} EDUHHHS_i + \beta_{12} EDUHHU_i + \\
&\beta_{13} EDUHHPC_i + \beta_{14} MIDWEST_i + \beta_{15} SOUTH_i + \\
&\beta_{16} WEST_i + \beta_{17} BLACK_i + \beta_{18} ASIAN_i + \\
&\beta_{19} OTHER_i + \beta_{20} HISP\_YES_i + \beta_{21} AGEPCLT6\_ONLY_i + \\
&\beta_{22} AGEPC6\_12ONLY_i + \beta_{23} AGEPC13\_17ONLY_i + \\
&\beta_{24} AGEPCLT6\_6\_12ONLY_i + \beta_{25} AGEPCLT6\_13\_17ONLY_i + \\
&\beta_{26} AGEPC6\_12AND13\_17ONLY_i + \beta_{27} AGEPCLT6\_6\_12AND13\_17_i + \\
&\beta_{28} MHONLY_i + \beta_{29} FHONLY_i + \beta_{30} INCOME_i
\end{aligned}
\tag{1}
$$

where $i = 1, \ldots\ldots, n$ is the number of households. *Y* corresponds to the decision to buy superfruit beverages. Variables are defined in Table 1.

A common characteristic in micro level data (data gathered at consumer level such as at the individual or household level) is a situation where some consumers do not purchase some items during the sampling period and presence of them in the sample creates a zero consumption level for that data period. The data used in this study are gathered at household level and due to that it suffers from zero consumption data. As such we face a censored sample of data. Application of ordinary least squares (OLS) to estimate a regression with a limited dependent variable (such as in a censored sample like ours) usually give rise to biased estimates, even asymptotically (Kennedy, 2003). Removing all observations pertaining to zero purchases and estimating regression functions only for non-zero purchases too creates a bias in the estimates. This phenomenon also is known as *sample selection bias*. Heckman (1979) stated that not adjusting for sample selection may result in biased estimates of the demand parameters. Furthermore, Heckman (1979), discussed the sample selection bias as a specification error, and developed a simple consistent estimation method that eliminates the

specification error for the case of censored samples. It is known as Heckman-type correction procedure.

The first stage of the Heckman-two-step sample selection procedure, involves in decision to purchase superfruit beverages. It is modeled through a probit model. A binary dependent variable is observed (purchase or not purchase), where purchase is represented by one (1) and not purchase is given by a zero (0). The latent selection equation can be written as follows;

$$Z_h = w_h'\gamma + \varepsilon_h \qquad (2)$$

where Z_k represents a latent selection variable (buy or not to buy type dichotomous variable),

$$Z_h = \begin{cases} 1 \ if \ \ Z_h > 0 \\ 0 \ if \ \ Z_h < or = 0 \end{cases} \qquad (3),$$

w_h is a vector of explanatory variables in the latent decision making variable, γ_h is a vector of parameters to be estimated in the decision making equation, ε_h is the error term, and $h = 1,2,....., N$ is the number of observations (in our work the number of households in the sample) in the sample. Modeling above equation 2 through probit model gives us following relationships;

$$\Pr[Z_h = 1] = \phi(w_h, \gamma) \qquad (4) \text{ and}$$

$$\Pr[Z_h = 0] = 1 - \phi(w_h, \gamma) \qquad (5)$$

where ϕ is the normal cumulative probability distribution function (cdf). The first stage estimation provides estimates of γ and the inverse of the Mills Ratio (IMR hereinafter). We also

generate the associated probability density function (pfd). Inverse of Mills Ratio is calculated taking the ratio of pdf to cdf. Mathematically, it is as follows;

$$\text{for } Z_k = 1,\ IMR_h = \frac{\varphi(w_h\hat{\gamma})}{\phi(w_h\hat{\gamma})} \qquad (6),$$

where φ represents the probability density function. Inverse mills ratio is a monotone decreasing function of the probability that an observation is selected into the sample, $\phi(w_k\hat{\gamma}_k)$ (Heckman, 1979). In particular,

$$\lim_{\phi(Z_{hi})\to 1} IMR_h = 0 \qquad (7)$$

$$\lim_{\phi(Z_{hi})\to 0} IMR_h = \infty \qquad (8)$$

$$\frac{\partial IMR_h}{\partial \phi(Z_h)} < 0 \qquad (9)$$

The calculated IMR, will be used as an additional explanatory variable in the second stage volume equation, which takes care of the sample selection bias in the data. Second stage equation is given as follows;

$$E[Y_h \mid Z_h = 1] = X_h'\beta + \alpha\frac{\varphi(w_h\hat{\gamma})}{\phi(w_h\hat{\gamma})} \qquad (10)$$

$$E[Y_h \mid Z_h = 1] = X_h'\beta + \alpha IM\hat{R}_h \qquad (11)$$

where X_k is a vector of explanatory variables considered in the second stage. Importantly, only observations associated with non-zero observations on Y_k are considered here. The IMR calculated using information retrieved from first stage probit model is used as an explanatory variable in the second stage (see equations 10 and 11 above). Presence of a sample selection bias in data will be communicated through statistical significance of the coefficient associated

with IMR, i.e. α_k . If α_k is statistically not different from zero, we conclude that there is no sample selection bias in the data and result in the following regression model;

$$E[Y_h \mid Z_h = 1] = X'_h \beta_i \qquad (12)$$

It is important to know that the explanatory variables in first stage and second stage equations may or may not be the same. In our work, the price variables in both equations do not. However, rest of the demographic variables is exactly the same in the first stage and second stage.

Choice of explanatory variables in the first and second stage has an implication on the derivation and interpretation of marginal effects associated with variables in the second stage. This is because in the second stage, we have the IMR term augmenting the regular regression function with other explanatory variables. Therefore, in calculating marginal effects, the influence of IMR and its associated regression coefficient on other regression coefficients have to be taken into consideration.

Suppose X_{kj} denote the *j*th regressor that is common to both first stage regressors, w_k and, second stage regressors, X_j . Differentiating equation 11 with respect to *j*th regressor, the marginal effect is given by the following relationship (following explanation is borrowed from Saha, Capps and Byrne (1997));

$$\frac{\partial E[Y_h \mid Z_h = 1]}{\partial X_{kj}} = \beta_{ij} + \alpha_i \frac{\partial (IM\hat{R}_{hi})}{\partial X_{hj}} \qquad (13)$$

It is evident from 13 that marginal effect of the *j*th regressor on Y_{ki} consists of two parts: a change in X_j which affects the probability of consuming the commodity (this effect is

represented by $\frac{\partial(IM\hat{R}_{hi})}{\partial X_{hj}}$ in 13); a change in X_j which affects the level of consumption (or expenditure of consumption) which is conditional upon the household choosing to consume the *i*th commodity (this is represented by β_{ij} in 13). The former of the above two expression is important, because the sign and magnitude of the marginal effect depends not only on the β_{ij}, but also that of the $\frac{\partial(IM\hat{R}_{hi})}{\partial X_{hj}}$. According to Saha, Capps and Byrne (1997), after some simplification we get arrive at the following relationship for the Heckman second stage marginal effects,

$$M\hat{E}_{kj} = \frac{\partial E[y_k \mid Z=1]}{\partial X_{kj}} = \beta_j - \alpha\gamma_{ij}\{W\gamma IMR_k + (IMR_k)^2\} \qquad (14)$$

In general the marginal effect $M\hat{E}_{kj} \neq \hat{\beta}_j$; however the only case where $M\hat{E}_{kj} = \hat{\beta}_j$ is where $\hat{\alpha} = 0$ which is a situation where the errors in the first-stage and second-stage estimation equations have zero covariance. It must be noted that the $M\hat{E}_{kj}$ estimation depends on a local set of co-ordinates. Therefore, we estimate the $M\hat{E}_{kj}$ at the sample means. Following equation 14 shows this result. For simplicity, let us denote IMR in the letter λ.

$$M\hat{E}_{kj} \mid_{samplemean} = \hat{\beta}_j - \hat{\alpha}_i\hat{\gamma}_j\{(\overline{W}\hat{\gamma})\overline{\hat{\lambda}} + \overline{\hat{\lambda}}^2\} \qquad (15)$$

where $\overline{W}$ denotes the vector of regressor sample means in the probit equation (the first stage equation of the Heckman two-step model and

$$\overline{\hat{\lambda}} = \frac{\varphi(\overline{W}\hat{\gamma})}{\phi(\overline{W}\hat{\gamma})} \qquad (16)$$

is the inverse Mills ratio evaluated at those means.

The Heckman two-step demand model for superfruit beverages can be written as follows:

$$
\begin{aligned}
q_i = {} & \beta_1 + \beta_2 P_i + \beta_3 AGEHH2529_i + \beta_4 AGEHH3034_i + \\
& \beta_5 AGEHH3544_i + \beta_6 AGEHH4554_i + \beta_7 AGEHH5564_i + \beta_8 AGEHHGT64_i + \\
& \beta_9 EMPHHPT_i + \beta_{10} EMPHHFT_i + \beta_{11} EDUHHHS_i + \beta_{12} EDUHHU_i + \\
& \beta_{13} EDUHHPC_i + \beta_{14} REG\_CENTRAL_i + \beta_{15} REG\_SOUTH_i + \\
& \beta_{16} REG\_WEST_i + \beta_{17} RACE\_BLACK_i + \beta_{18} RACE\_ORIENTAL_i + \\
& \beta_{19} RACE\_OTHER_i + \beta_{20} HISP\_YES_i + \beta_{21} AGEPCLT6\_ONLY_i + \\
& \beta_{22} AGEPC6\_12ONLY_i + \beta_{23} AGEPC13\_17ONLY_i + \\
& \beta_{24} AGEPCLT6\_6\_12ONLY_i + \beta_{25} AGEPCLT6\_13\_17ONLY_i + \\
& \beta_{26} AGEPC6\_12AND13\_17ONLY_i + \beta_{27} AGEPCLT6\_6\_12AND13\_17_i + \\
& \beta_{28} MHONLY_i + \beta_{29} FHONLY_i + \beta_{30} INCOME_i + \alpha_i IMR + \varepsilon_i
\end{aligned}
\tag{17}
$$

where $i = 1,\ldots\ldots, n$ is the number of observations (households in our work) in the model. q_i corresponds to the quantity of purchase of superfruit beverages and P_i variable represent the price of superfruit beverages. We have defined the variables in the above equation 17 in Table 1. In the equation 17, *IMR* stands for the inverse Mills ratio and α_i corresponds to the coefficient associated with *IMR*. Presence of sample selection bias is determined looking at the significance of α_i. If we have sample selection bias, we have to do an adjustment to the coefficient estimates in the second stage estimation in trying to get at correct marginal effects. Procedure to adjust for marginal effects was elaborated in the preceding section.

As such, we will calculate marginal effects associated with each explanatory variable. The level of significance we will be using in this study is 0.05. We further conduct an *F*-test for demographic variable categories to find statistically significant demographics.

RESULTS AND DISCUSION

Once the demand is estimated, we are in position to calculate own- and cross-price and income elasticities for these segments for superfruit beverages, conventional fruit beverages, and others. This information will reflect the market competitiveness and profiles of demographics consuming superfruit beverages in the United States. In the end, they useful for superfruit beverage manufacturers, wholesalers and retailers for strategic pricing decisions as well as government policy makers to implement policies related to food and nutrition.

Table 1 Description of the Right-Hand Side Variables Used in the Econometric Analysis

Variable	Explanation
PRICE	Price of superfruit beverages
AGEHHLT25	*Age of Household Head less than 25 years (Base category)*
AGEHH2529	Age of Household Head between 25-29 years
AGEHH3034	Age of household Head between 30-34 years
AGEHH3544	Age of household Head between 35-44 years
AGEHH4554	Age of household Head between 45-54 years
AGEHH5564	Age of household Head between 55-64 years
AGEHHGT64	Age of household Head greater than 64 years
EMPHHNFP	*Household Head not employed for full pay (Base category)*
EMPHHPT	Household Head Part-time Employed
EMPHHFT	household Head Full-time Employed
EDUHHLTHS	*Education of Household Head: Less than high school (Base category)*
EDUHHHS	Education of Household Head: High school only
EDUHHU	Education of Household Head: Undergraduate only
EDUHHPC	Education of Household Head: Some post-college
EAST	*Region: East (Base category)*
MIDWEST	Region: Central (Midwest)
SOUTH	Region South
WEST	Region West
WHITE	*Race White (Base category)*
BLACK	Race Black
ASIAN	Race Oriental
RACE_OTHER	Race Other (non-Black, non-White, non-Oriental)
HISP_NO	*Non-Hispanic Ethnicity (Base category)*
HISP_YES	Hispanic Ethnicity

Table 1 Continued….

Variable	Explanation
NPCLT_18	*No Child less than 18 years (Base category)*
AGEPCLT6_ONLY	Age and Presence of Children less than 6-years
AGEPC6_12ONLY	Age and Presence of Children between 6-12 years
AGEPC13_17ONLY	Age and Presence of Children between 13-17 years
AGEPCLT6_6_12ONLY	Age and Presence of Children less than 6 and 6-12 years
AGEPCLT6_13_17ONLY	Age and Presence of Children less than 6 and 13-17 years
AGEPC6_12AND13_17ONLY	Age and Presence of Children between 6-12 and 13-17 years
AGEPCLT6_6_12AND13_17	Age and Presence of Children less than 6, 6-12 and 13-17 years
FHMH	*Household Head both Male and Female (Base category)*
MHONLY	Household Head Male only
FHONLY	Household Head Female only

REFERENCES

"Fruit and Vegetable Juices: U.S. Market Trends - Market Research Report."

Gross, "Superfruits: Defining an Emergent Category."

"Healthy Flavors Boom in US."

Landi, "Beverage Brands Innovate with Super Fruits."

"240_School Beverage Guidelines Final Progress Report Executive Summary.pdf."

"POM Wonderful Announces Purely Juice Found Guilty of Deceiving Customers With Adulterated... -- Re> LOS ANGELES, July 22 /PRNewswire/ --."

"State of the Industry."

"Superfruits — Super Sales and Super Claims."

"Superfoods To The Rescue

Kennedy, P. (2003). Limited Dependent Variables. A Guide to Econometrics, MIT Press.

Heckman, J. J., (1979). "Sample Selection Bias as a Specification Error". *Econometrica* 47: 153-161