

**Demand and Market Competitiveness of Almond Milk as a Dairy Alternative
Beverage in the United States**

Senarath Dharmasena

Oral Capps, Jr.

Brooke Kosub

Department of Agricultural Economics and
Agribusiness, Food and Consumer Economics Research Center (AFCERC)
Texas A&M University

Email: sdharmasena@tamu.edu, ocapps@tamu.edu, bkhornet2012@tamu.edu

Phone: (979) 862-2894

***Selected Paper Proposal to Southern Agricultural Economics Association Annual Meeting,
Atlanta, Georgia, January 31-February 3, 2015***

Copyright 2015 by Senarath Dharmasena, Oral Capps, Jr and Brooke Kosub. All rights reserved. Readers may make verbatim copied of this document for non-commercial purposes by any means, provided that this copyright notice appears on all such copies.

Demand and Market Competitiveness of Almond Milk as a Dairy Alternative Beverage in the United States

Abstract

There are many different types of nonalcoholic beverages available in the United States today as compared to a decade ago and the functionality and health dimensions of beverages have changed over the years. Recently, calcium and vitamin fortified dairy alternative beverages, such as almond milk and soymilk have entered the market to compete with conventional milk. Knowledge of price sensitivity, substitutes/complements, and demographic profiling with respect to consumption of dairy alternative beverages is important for manufacturers, retailers, advertisers, nutritionists, and other stakeholders from a competitive intelligence and strategic decision-making perspective. Using nationally representative household level data from 65,000 households (Nielsen Homescan), and tobit econometric procedure, factors affecting the demand for almond milk for all households and households grouped by race, ethnicity, region, and income status will be determined. Moreover, own-price, cross-price, and income elasticities for almond milk delineated by selected demographic segments will be estimated. Preliminary analyses reveal that the own-price elasticity of demand for almond milk is -3.50. Soymilk is found to be a substitute for almond milk. This information will be useful for almond milk manufacturers, wholesalers, and retailers for strategic pricing decisions as well as government policy makers to implement policies related to food consumption and nutrition.

Keywords: Almond milk, soymilk, conventional milk, consumer demand, tobit, Nielsen Homescan

JEL Classification: D11, D12

Demand and Market Competitiveness of Almond Milk as a Dairy Alternative Beverage in the United States

Introduction

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. On top of conventional hydration and refreshment functions, beverages now are fortified with numerous vitamins, minerals, proteins, antioxidants, favorable fatty acids, etc (BMC, 2010; 2011, 2012).

Currently, calcium and vitamin fortified dairy alternative beverages are entering the market to compete with dairy milk, providing consumers an alternative, specifically for those who are lactose intolerant. To strengthen the position of this, the new food guidelines developed under the “ChooseMyPlate”, placed dairy alternatives such as soymilk, rice milk and almond milk in the “Dairy Group” (USDA, 2014). This placement raised eyebrows of dairy producers and marketers in the United States, and it is of interest for them to know the competitiveness of dairy alternatives in the dairy marketplace.

Dairy-alternative products represented roughly five percent of dairy launches in 2012, with soy being the primary or secondary ingredient in 78 percent of them (Innova Market Insights, 2013). However, this trend with respect to soy is changing as interest is growing in dairy alternatives made with ingredients including almonds, rice, oats, barley, hazelnuts and walnuts.

According to Chicago based market research firm, Mintel, almond milk has overtaken soymilk over the past two years and has become America’s most popular plant-based milk alternative accounting for 4.1% of total milk sales (KCT.org, 2014). Almond milk now

dominates dairy alternative beverage market with a staggering 60% market share, while soymilk has only about 30% share (Food Navigator, 2014). Growth in almond milk has been attributed to improved health-related claims and consumer perceptions, a flurry of almond milk brands, appealing and convenient packaging, and a plethora of flavors available. Sales of dairy alternative beverages reached nearly \$2 billion in 2013, driven up largely as a result of popularity of almond milk (The Washington Post, 2014).

Given this backdrop, knowledge of price sensitivity, substitutes/complements and demographic profiling with respect to consumption of dairy alternative 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. We are not aware of any past study pertaining to demand for dairy alternative beverages focusing on almond milk in the extant literature. The most comprehensive market research on dairy alternative beverages thus far was conducted by Dharmasena and Capps (2014). However, they focused on soymilk, and did not consider almond milk in their study. Therefore, to our knowledge, our study is the first to examine the market competitiveness and demographic factors determining U.S. demand for almond milk.

Objectives:

The general objective of this study is to develop models that uncover the demand for almond milk by a diverse set of consumers. The specific objectives of this study are to: (1) determine the factors affecting the decision to purchase almond milk for all households as well as for households grouped by race, ethnicity, region and income/poverty status; (2) to estimate own- and cross-price and income elasticities for almond milk delineated by selected

demographic segments; and (3) to determine the competitiveness of almond milk vis-à-vis dairy milk and soymilk.

Data and Methodology:

Household purchases of soymilk, almond milk, white milk and flavored milk (expenditure and quantity) and socio-economic-demographic characteristics will be generated for each household in the Nielsen Homescan Panel for calendar year 2011. These data represent the most recent data available to us. Expenditure and quantity data for each household will be aggregated to form quarterly observations for each type of milk and dairy alternative beverages, subsequently generating a panel dataset. In other words we will be using pooled time-series (four quarters) and cross-sectional data (61,000 households). Quantity data will be standardized in terms of liquid ounces and expenditure data will be expressed in terms of dollars. Then taking the ratio of expenditure to volume, we will generate unit values (prices in dollars per ounce) for each beverage.

Trying to use a standard Heckman-type (Heckman, 1979) sample selection correction for censored panel data could induce problems with respect to calculation of the inverse mills ratio (fixed effects or random effects model inverse mills ratio) as well as for the second-stage conditional demand model (Vella, 1992; Vella, 1998; Verbeek and Nijman, 1992; Wooldridge, 2002). Therefore, using the aforementioned panel in the presence of censored observations, we will estimate demand models for dairy alternatives using panel tobit specification taking care of censoring issue associated with this data (Wooldridge 2002). Panel tobit procedure available in the Stata statistical package will be used to model aforementioned tobit specifications. Quantity of soymilk, almond milk, white milk and flavored milk are considered as dependent variables in each demand equation. Price of milk and dairy alternative beverages, and host of demographic

variables (age of household head, education, household income, race, ethnicity, region, presence of children, poverty status) are considered as explanatory variables in each tobit specification.

Finally, we will generate forecasts of quantities of milk and dairy alternative beverages consumed for all households, and households grouped by selected demographic segments.

Preliminary analysis was performed using the tobit procedure (Tobin, 1958) for almond milk purchases for the year 2011. We generated both conditional and unconditional demand estimates pertaining to almond milk. Heckman (1979) model only will be able to speak to conditional demand estimates, although in the first stage probit analysis will provide information on consumer's probability to purchase or not to purchase the product. Also, we use the decomposition of the "beta" coefficient estimates of tobit model suggested by McDonald and Moffitt (1980) to shed light on changes in probability of being above the limit (limit being zero in this paper) and changes in the value of the dependent variable if it is already above the limit. This is the McDonald and Moffitt decomposition associated with tobit parameter estimates.

For all those transactions associated with zero quantities and hence zero expenditures, we do not observe any unit value or price. However, since we are expecting to use price of each beverage category as explanatory variables in the tobit model, we have to impute price for those observations where no price is observed. Price imputation is done using an auxiliary regression, where observed prices for each beverage are regressed on household income, household size and region where the household is located. These variables are used extensively in the price imputation literature as good instruments in imputing prices. Once the price for almond milk is imputed, we use them and aforementioned explanatory variables to estimate the tobit model pertaining to almond milk consumption. Table 1 shows different categories of explanatory variables used in this study along with base categories for dummy variables.

The Tobit Model

The stochastic model underlying the tobit model can be expressed as follows:

$$(1) \quad y_i = \begin{cases} X_i\beta + u_i, & X_i\beta + u_i > 0 \\ 0, & X_i\beta + u_i \leq 0 \end{cases}$$

where $i = 1, 2, 3, \dots, N$, the number of observations. y_i is the censored dependent variable; X_i is the vector of explanatory variables; β is the vector of unknown parameters to be estimated; $E[u_i|X] = 0$ and $u_i \sim N(0, \sigma^2)$. The unconditional expected value for y_i is expressed in equation (2) and the corresponding conditional expected value for y_i is shown in equation (3), where the normalized index value z is shown as $z = \frac{X\beta}{\sigma}$. Also, $F(z)$ is the cumulative distribution function (CDF) associated with z and $f(z)$ is the corresponding probability density function (pdf).

$$(2) \quad E(y) = X\beta F(z) + \sigma f(z)$$

$$(3) \quad E(y^*) = X\beta + \sigma \frac{f(z)}{F(z)}$$

The unconditional marginal effect is represented by,

$$(4) \quad \frac{\partial E(y)}{\partial X} = \beta F(z)$$

The conditional marginal effete is shown by,

$$(5) \quad \frac{\partial E(y^*)}{\partial X} = \beta \left(1 - z \frac{f(z)}{F(z)} - \frac{f(z)^2}{F(z)^2}\right)$$

Furthermore, the McDonald and Moffitt (1980) decomposition relating both change in conditional expectations and unconditional expectations can be shown below. In other words, the total change in unconditional expected value of the dependent variable, y can be represented by the sum of the change in the expected value of y being above the limit, weighted by the

probability of being above the limit and change in probability of being above the limit weighted by the expected value of y being above the limit.

$$(6) \quad \frac{\partial E(y)}{\partial x} = F(z) \left(\frac{\partial E(y^*)}{\partial x} \right) + E(y^*) \left(\frac{\partial F(z)}{\partial x} \right)$$

Empirical Estimation

We tried several functional forms such as liner, quadratic and linear-log to find that Linear-Log model (we used logged price variables in the model) outperforms other functional forms as far as the model fit, significance of variables and loss matrices such as AIC and Schwarz criteria are concerned. The tobit model for almond milk can be represented as follows,

$$(7) \quad \begin{aligned} (Q_Almond_Milk) = & \beta_1 + \beta_2 \log PRICE_AL_i + \beta_2 \log PRICE_SOY_i + \\ & \beta_3 \log PRICE_WMILK_i + \beta_5 PRICE_FMILK_i + \beta_6 AGEHH2529_i + \\ & \beta_7 AGEHH3034_i + \beta_8 AGEHH3544_i + \beta_9 AGEHH4554_i + \beta_{10} AGEHH5564_i + \\ & \beta_{11} AGEHHGT64_i + \beta_{12} EMPHHPT_i + \beta_{13} EMPHHFT_i + \beta_{14} EDUHHHS_i + \\ & \beta_{15} EDUHHU_i + \beta_{16} EDUHHPC_i + \beta_{17} MIDWEST_i + \beta_{18} SOUTH_i + \\ & \beta_{19} WEST_i + \beta_{20} BLACK_i + \beta_{21} ASIAN_i + \\ & \beta_{22} OTHER_i + \beta_{23} HISP_YES_i + \beta_{24} AGEPCLT6_ONLY_i + \\ & \beta_{25} AGEPC6_12ONLY_i + \beta_{26} AGEPC13_17ONLY_i + \\ & \beta_{27} AGEPCLT6_6_12ONLY_i + \beta_{28} AGEPCLT6_13_17ONLY_i + \\ & \beta_{29} AGEPC6_12AND13_17ONLY_i + \beta_{30} AGEPCLT6_6_12AND13_17_i + \\ & \beta_{31} MHONLY_i + \beta_{32} FHONLY_i + \beta_{33} INCOME_i \end{aligned}$$

As such, we will calculate both conditional and unconditional 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. The equations for unconditional and conditional marginal effects for the Linear-Log model and corresponding unconditional and conditional own- and cross-price elasticity estimates are explained below.

The unconditional marginal effect for the Linear-Log model is as follows,

$$(8) \quad \frac{\partial E(y)}{\partial X} = \frac{\beta}{P^U} F(z)$$

where P^U is the average price of all observations (unconditional price) considered. The conditional marginal effect for the Linear-Log model is as follows,

$$(9) \quad \frac{\partial E(y^*)}{\partial X} = \frac{\beta}{P^C} \left(1 - z \frac{f(z)}{F(z)} - \frac{f(z)^2}{F(z)^2}\right)$$

Where, P^C is the average price of non-censored sample (conditional price).

The unconditional own- and cross-price demand elasticities are represented by equations (9) and (10) respectively.

$$(10) \quad \varepsilon_{ii}^U = \frac{P_i^U}{Q_i^U} \frac{\beta}{P^U} F(z)$$

$$(11) \quad \varepsilon_{ij}^U = \frac{P_j^U}{Q_i^U} \frac{\beta}{P^U} F(z)$$

The conditional own- and cross-price demand elasticities are represented by equations (11) and (12) respectively,

$$(12) \quad \varepsilon_{ii}^C = \frac{P_i^C}{Q_i^C} \frac{\beta}{P^C} \left(1 - z \frac{f(z)}{F(z)} - \frac{f(z)^2}{F(z)^2}\right)$$

$$(13) \quad \varepsilon_{ij}^C = \frac{P_j^C}{Q_i^C} \frac{\beta}{P^C} \left(1 - z \frac{f(z)}{F(z)} - \frac{f(z)^2}{F(z)^2}\right)$$

Expected Results and Discussion:

Once the conditional demand functions are estimated for all households, and households grouped by race, ethnicity, region and income/poverty status, we are in position to calculate own- and cross-price and income elasticities for these segments. This information will reflect

the market competitiveness and profiles of demographics consuming almond milk in the United States. Preliminary analysis of almond milk data reveal that the own-price elasticity of demand for almond milk in the United States is -3.50. The cross price elasticity of almond milk with soymilk is 0.22, making soymilk a substitute for almond milk in consumption. In the end, these results are useful for almond milk manufacturers, wholesalers and retailers for strategic pricing decisions as well as government policy makers to implement policies related to food and nutrition. Also, this information will be useful for dairy processors to understand the competitive position of dairy milk in the dairy alternative beverage marketplace.

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

Variable	Explanation
PRICE	Price of Almond milk
<i>AGEHHLT25</i>	<i>Age of Household Head less than 25 years (Base category)</i>
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
<i>EMPHHNP</i>	<i>Household Head not employed for full pay (Base category)</i>
EMPHHPT	Household Head Part-time Employed
EMPHHFT	household Head Full-time Employed
<i>EDUHHLTHS</i>	<i>Education of Household Head: Less than high school (Base category)</i>
EDUHHHS	Education of Household Head: High school only
EDUHHU	Education of Household Head: Undergraduate only
EDUHHPC	Education of Household Head: Some post-college
<i>EAST</i>	<i>Region: East (Base category)</i>
MIDWEST	Region: Central (Midwest)
SOUTH	Region South
WEST	Region West
<i>WHITE</i>	<i>Race White (Base category)</i>
BLACK	Race Black
ASIAN	Race Oriental
RACE_OTHER	Race Other (non-Black, non-White, non-Oriental)
<i>HISP_NO</i>	<i>Non-Hispanic Ethnicity (Base category)</i>
HISP_YES	Hispanic Ethnicity

Table 1 Continued....

Variable	Explanation
<i>NPCLT_18</i>	<i>No Child less than 18 years (Base category)</i>
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_17 ONLY	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
<i>FHMH</i>	<i>Household Head both Male and Female (Base category)</i>
MHONLY	Household Head Male only
FHONLY	Household Head Female only

References:

- BMC, 2010, 2011, 2012, Beverage Marketing Corporation Multiple Beverage Marketplace Reports
- Dharmasena, S., and O.Capps, Jr. (2014) “Unraveling Demand for Dairy Alternative Beverages: The Case of Soymilk.” *Agricultural and Resource Economics Review*, 43(1): 140-157.
- Food Navigator, (2014), “Almond Milk Accounts for Two-Thirds of Sales in Plant-Based Milk Category.” Internet access: <http://www.foodnavigator-usa.com/Manufacturers/Almond-milk-accounts-for-two-thirds-of-sales-in-plant-based-milk-category-says-WhiteWave-Foods> (accessed on September 3, 2014).
- Heckman, J. J., (1979). “Sample Selection Bias as a Specification Error”. *Econometrica* 47: 153-161
- Innova Market Insights, FoodBusinessNews.net, accessed January 7, 2013.
- KCT.org, (2014) “Demand for Almond Milk is Growing, but is it Bad for the Planet?” Internet access: <http://www.kcet.org/living/food/the-nosh/demand-for-almond-milk-is-soaring-but-is-it-bad-for-the-planet.html> (accessed on September 3, 2014)
- Soyfoods Association of North America, (2013) “Soy Information: Sales and Trends.” Internet Access: <http://www.soyfoods.org/soy-information/sales-and-trends> (accessed on February 12, 2013)
- STATA, 2013, STATA Statistical Software, College Station, USA
- The Washington Post, (2014), “Got Milk? From a Cow or a Plant?” Internet access: <http://www.washingtonpost.com/news/to-your-health/wp/2014/06/10/got-milk-from-a-cow-or-a-plant/> (accessed on September 3, 2014).
- USDA, (2014). “What foods are included in the Dairy Group?”. ChooseMyPlate, (internet access on September 3, 2014: <http://www.choosemyplate.gov/food-groups/dairy.html>)
- Vella, F., (1992). “Simple Tests for Sample Selection Bias in Censored and Discrete Choice Models”, *Journal of Applied Econometrics* 7:413-421
- Vella, F., (1998). “Estimating Models with Sample Selection Bias: A Survey”. *Journal of Human Resources* 33: 127-169
- Verbeek, M., and T. Nijman, (1992). “Testing for Selectivity Bias in Panel Data Models”, *International Economic Review* 33: 681-703
- Wooldridge, J.M., (2002). “Econometric Analysis of Cross Section and Panel Data”. The MIT Press, Cambridge, Massachusetts