Got Local? The Impact of Food Miles on the Demand for Milk

Rigoberto Lopez and Binod Khanal

Selected paper prepared for presentation at the Agricultural and Applied Economics Association Meeting, Kansas City, MO, July 23-26, 2020

Rigoberto Lopez is the DelFavero Professor of Agricultural Economics, and Binod Khanal is a Ph.D. student, at the University of Connecticut. Corresponding author: Rigoberto.Lopez@uconn.edu.

This work was supported by the USDA National Institute of Food and Agriculture, Hatch project accession number 1020636.

Abstract

Increased consumer preferences for specialty and local milk offer marketing opportunities for producers. In contrast to commoditized milk, locally marketed milk generally commands higher prices. This paper examines consumer preferences for local milk, using the concept of food miles and localness as a continuous variable rather than defining local by state boundaries. We apply a random coefficient logit model to fluid milk sales data from six Northeastern U.S. cities, using both distance between bottling and consumption points as well as state boundaries to define localness. Regarding of the measure of localness, empirical results indicate that consumers' valuation of milk products is sensitive to where the milk originates and that the more flexible distance measure of food miles outperforms the conventional definition based on state boundaries.

Keywords: local food, food miles; milk, demand,

JEL code: Q11, Q13.

Introduction

Between 2000 and 2018, per capita U.S. fluid milk consumption declined by 25 percent.[1] At the same time, technological change and consolidation both at the farm and processing levels have expanded supply, resulting in lower prices and often negative profit margins for commoditized milk and accelerating the exit of small dairy farms from the market. Due to economies of scale, in a commoditized milk market small dairies cannot compete unless they receive higher profit margins from higher priced value-added in specialty milk markets that align with consumer shifts in preferences, particularly those for local and organic products. Producers can benefit from shifting to these marketing and production methods to align with shifts in consumer preferences (Wolf et al., 2011; Connolly & Klaiber, 2014). Thus, understanding consumer preferences for local milk is a cornerstone of strategies to enable producers to benefit from these trends.

Although previous studies define local as produced within state boundaries (Khanal, Lopez, and Azzam, 2020), the typical definition of local in the consumer's mind is rather fuzzy, and it varies across consumers by state, region, and within North America (Campbell et al., 2014). In the parallel literature on food miles tied to consumers' concerns with the environment, the concept of distance to the point of production appears to accommodate a variety of definitions of "local." In this study, we introduce a more flexible definition of "localness" than used in previous local food studies, based on distance from the point of production to the point of consumption, using fluid milk as a case study.

[1] During this time, however, the per capita consumption of cheese and yogurt increased. Per capita cheese consumption has gone from mere 14 pounds in 1975 to 37 pounds 2017, while per capita yogurt consumption has increased from two pounds in 1975 to 13 pounds in 2017.

Empirical studies on the demand for local foods so far have shown a higher price premium (Darby et al., 2008; Connolly and Klaiber, 2014). Most of these studies arbitrarily define localness as produced within a state or a county, but this definition of localness might not elucidate the notion of why consumers prefer local products. One of the economic perspectives on localness is that the food travels a just few miles before reaching consumers. This reduces carbon emissions and, as a result, local food is regarded as environmentally friendly. Moreover, the FSMA (Food Safety Modernization Act) assumes that fewer food miles are also associated with lower pathological activity; therefore, local food is assumed to be healthier in terms of risk of foodborne illness. Given the wide range of states' areas across the United States (states in the New England region are relatively smaller compared to states in the Midwest), using state boundaries to define localness does not accommodate consumers' perception of local products. Some demand studies have attempted to incorporate the notion of fewer than 50 miles or a 400-miles radius as common benchmark distances in defining local, but very few studies so far have used the real distance between the point of production and consumption.

Grebitus, Lusk, and Nayga (2013), using an experimental auction game, introduced a food mile concept to explore willingness to pay (WTP) for locally produced wine and locally grown apples, and found a higher WTP for both. A study by Kemp et al. (2010) showed that a quarter of English consumers would stop buying products from New Zealand because of the greater number of food miles associated with those products. Likewise, Caputo, Nayga, and Scarpa (2013) found that consumers' valuation of products that traveled fewer miles was, most of the time, equal to that for emission labeling. No other definitions of localness based on the concept of food mile could be as insightful, however, as using the real distance between production and consumption. As far as we know, ours is the first study to use a flexible definition of localness based on food miles in the study of demand for milk.

This paper makes two contributions to the literature on local food demand. The first is the use of revealed preference data based on actual sale transactions. We fit the data to a random coefficient logit demand model (Berry, Levinsohn, and Pakes, 1995, herein BLP), as done by Lopez and Lopez (2009) for fluid milk. Second, we use a flexible definition of localness based on the distance between the milk bottling plant and the point of consumption. This sidesteps arbitrary definitions of market boundaries to designate local, such as state boundaries, which differ across states, or distances (e.g., 100 miles). Unlike many other studies that use BLP to explore demand in a specific location, this study covers multiple geographical locations.

Economic model and data

We model consumer choices based on the BLP random coefficient logit demand model. An advantage of the BLP is that by virtue of being based on characteristic space, it readily takes into account for product characteristics, such as food miles attributed to a particular product, as well as consumer heterogeneity. Let the indirect utility of consumer i (i=1……….n) from consuming one unit of milk product j (j=1…..J) be given by, [2]

$$U_{ij} = \alpha_i P_j + \beta_i X_j + \Theta_i Local_j + \varepsilon_{ij} \qquad (1)$$

where, P_j is the price of milk product j; X_j is a vector of the milk characteristics other than localness; $Local_j$ is localness associated with that product; α_i, β_i, and $\Theta_i$are parameters unique to consumer I; and ε_{ij} is an i.i.d error term that follows a type I extreme value distribution.

[2] For the sake of simplicity, we omitted the time and city subscripts that we subsequently used in the estimation.

The taste parameters are supposed to be determined by the observed (*D*) and unobserved (*v*) demographic characteristics. Observed characteristics include income, presence of children in the household, and the presence of elderly people in the household. We have created the unobserved variable, which follows a standard normal distribution.

$$\alpha_i = \alpha + \gamma D_i + \delta \nu_i \qquad (2)$$

$$\beta_i = \beta + \rho D_i + \varphi \nu_i \qquad (3)$$

$$\theta_i = \theta + \omega D_i + \tau \nu_i \qquad (4)$$

Now, using (2) and (3) and (4) in equation (1), the indirect utility takes the following form:

$$U_{ij} = \underbrace{\alpha P_j + \beta x_j + \theta Local_j}_{\delta_j} + \underbrace{\lambda D_i P_j + \xi \nu_i P_j + \phi D_i x_j + \psi \nu_i x_j + \upsilon D_i Local_j + \epsilon \nu_i Local_j + +}_{\mu_{ij}} \varepsilon_{ij} \qquad (5)$$

Here, δ_j, and μ_{ij} are, respectively, the mean utility across consumers and a consumer-specific deviation from the mean utility. The model further assumes that each consumer buys only one unit of the good that provides him/her with the highest utility. As ε_{ij} is assumed to be i.i.d with type I extreme value distribution, the probability that consumer *i* purchases product *j,*

Allow for an outside good with utility U_{i0}=0, which offers the option of not buying market brands under consideration. Thus, the probability a consumer *i* buying a unit of brand j is given by:

$$Prob_{ij} = \frac{\exp(\delta_j + \mu_{ij})}{1 + \sum_{k=1}^{J} \exp(\delta_k + \mu_{ij})} \quad . \qquad (6)$$

Aggregating over all consumers in the market, one obtains the market demand, expressed as a share of product j given by:

$$S_j = \iiint I\{(D_i, v_i, \varepsilon_{ij}): v_{ij} \geq U_{ik}\ \forall k = 0, \ldots .j\} dH(D) dG(V) dF(\varepsilon), \qquad (7)$$

where H(D), G(V), and $F(\varepsilon)$ are cumulative density functions of the respective variables as defined. Since (7) does not have an analytical closed form, it can be approximated numerically through an estimator that uses random draws of *ns* consumers in the market (Berry, 1994; Vincent, (2015). Letting *k* denote the individuals in the draw (*k=1,..ns*), and following BLP and Nevo (2000), approximate (7) by

$$S_j = \frac{1}{ns} \sum_{k=1}^{ns} Prob_{kj} = \frac{1}{ns} \sum_{k=1}^{ns} \frac{exp(\delta j + vk\ \ (v))}{1 + \sum_{r=1}^{J} (\delta r + v\ \ \ (v))} \quad . \qquad (8)$$

Data and estimation strategy

The primary source for this study is IRI (Information Resource, Incorporated) academic data. IRI has store-level weekly sales data for each brand of the product under consideration. Along with the sales data, they have characteristics of the brand, such as whether it is organic and its fat content. Our study uses sales data from January 2004 to December 2011.

For each brand in the sample, the milk bottling plant location was obtained from the U.S. Department of Health and Human Services (USDHHS, 2017). We used Google Maps to estimate the road distance between the point of consumer purchaser (city) and the bottling plant, i.e., thus obtaining our measure of food miles. The number of draws in equation (8) to estimate a closed form of the demand model was obtaining by drawing 500 random observations from market taken from the Consumer Population Survey's March supplement, available from the U.S. Bureau of Census. In addition, the CPS was also use to obtain other demographic data, focusing on three consumer characteristics: age of the household head, annual household income, and number of children in the household. Like Nevo (2001), we also created a vector of unobserved consumer characteristics with standard normal distribution, *N(0, 1).*

Sales data for each market and time period came from the IRI database. Before defining market shares, we define the market size as total milk consumed at home and away from home. This is approximated by multiplying U.S. per capita milk consumption by population in a given city Then market shares are calculated by dividing the quantity of milk of a given brand by the size of the milk market in a given city and time period. The price of the brand is obtained by dividing the total dollar sales value by total volume sold in a given four-week period. Potential price endogeneity was addressed using two different sets of instruments. The first is the interaction between the cost shifters, viz., regional gas price, state-level electricity price, farm price of milk, weekly retail wage rate, and the brand dummies. The second set is the market dummies. So, we have 56 (14*4) variables from the first set of instruments and six variables from the second set, for a total of 62 IVs. Weekly retail wages across different markets come from the Bureau of Labor Statistics' (BLS) Quarterly Census of Employment and Wages (QCEW). The farm price of the fluid milk comes from the National Agricultural Statistics Service (NASS), United States Department of Agriculture (USDA). Regional gas and state-level electricity prices are from the Environmental Protection Agency (EPA). We also checked the retail price of milk in other cities, as suggested by Nevo (2000), along with interactions between cost shifters and brand dummies (57 IVs) as a set of instrumental variables and got comparable results. But as the statistical tests[3] suggested that the instruments created by the interaction between brand and cost shifters and market dummies perform better than the milk price from

[3] The set of the instruments is strong and statistically shown using under-identification test by Anderson LM statistics, the weak instrument test by Cragg-Donald Wald F statistics, and Stock-Yogo weak identification test under 2SLS framework. As we tested the validity of the instruments, the price is regressed against the instruments in the first stage, while market share of brands are regressed against the predicted price and other exogenous regressors in the second stage regression. For the overidentification test, we fit the model with a GMM estimator, and the Hansen J statistics proves the exogeneity of the additional instruments.

other markets and interaction between brand dummies and cost shifters, we have presented the results estimated by using the former set of IVs.

We have used 14 different brands of milk in this study across six cities in the northeastern United States, viz., Boston, Buffalo-Rochester, Hartford, New York, Philadelphia, and Providence. The completion of the BLP model is achieved by defining an outside good, which in our study is defined as other fluid milk brands and milk sold in stores not included in our study. Following a common norm, we have defined the market as a city-time. As this data is aggregated to four-week timespans, there are 105 time periods, yielding 105 markets for one city and 630 markets in total. However, since not all milk brands are sold in all the cities covered in this study, the panel is not balanced. Moreover, some local brands in a city have a very low market share in other cities, which we did not include in our study, resulting in an unbalanced panel. We dropped brands that have a market share of less than 0.1 percent. In the end, we used 2,843 observations in the estimation of the main model.

Results and discussion

The summary statistics of for the variables used in the model are shown in Table 1. The average household income is around $58,000, the average age of the household head is around 51 years, and the average number of children in the household is 0.76. All the variables from the samples are comparable to the average demographics of the region. Since there is not enough information on the bottling plant for private labels, we did not use private labels in this study. Most of the brands are either regional or local. The average market share of the brands is slightly above 3 percent, while the average price of the milk is USD 6.76 per gallon. Thirty percent of the brands are branded as organic. The average food miles for these milk brands are 152, which is fewer than the 400-mile benchmark that has been widely used in defining local in several studies (Low et al. 2015).

Once the empirical variables were operational, we used Stata to produce the parameter estimates in Table 2. Two alternative model specifications are presented based on how localness is defined. The first model uses the food mile concept of localness, while the second model uses the traditional state boundary. All consumer and product characteristics are statistically significant at less than the 5 percent level of significance in both models, with the exception of the price coefficient in the state-local model. As expected, price has a negative mean utility in both models. However, only the distance-based model shows a statistically significant estimate for milk price.

Independent of demographics, consumers prefer milk with fewer food miles, a finding directly opposite of studies that use the state-local definition in which local milk is valued less than out-of-state milk. Thus, the latter is contrary to expectations and inconsistent with the proposed distance-based model. However, to fully evaluate the effect of localness, the role of demographics must be considered, as captured by the interaction terms.

Note that in Table 2, all the coefficients for the interaction terms between localness and consumer characteristics are statistically significant. The interaction coefficients are consistent across both models, with the exception of the one for age. For instance, in both models, higher income consumers value localness more than do consumers without or with a lower number of children. The distance model shows that younger consumers value local milk more, while the state-local model shows that they value it less, although the associated coefficient is significant only at the 10 percent level. The age result for the distance model is in line with many findings that young consumers are more inclined toward local products as compared to older consumers.

The following are the equations that show the marginal valuation of localness (distance and state local) with respect to demographic variables:

Marg. Utility w.r.t. distance=-36.44+0.44*Age-0.38*Children-10.8*Income+0.154*Unobserved

Marg. Utility w.r.t. State local=-79.55+0.17*Age-0.15*Children+39.44*Income+1.36*Unobserved

Conclusion

Using the IRI dataset and a random coefficient logit demand model, we estimate two models with alternative definitions of localness. Empirical results show that a more flexible, novel model of food miles (distance between the bottling and consumption points) outperforms the more conventional model based on state boundaries. It also obviates the need to lock the analysis into a specific definition based on fixed distances (e.g., 100-mile radius), as done in previous work. More specifically, the empirical results show that consumers value milk with fewer food miles, particularly those who are younger, with higher incomes, and fewer children. Since low prices for commoditized fluid milk have accelerated the disappearance of small dairy farmers, our study suggests that focusing on 'localness' could be a potential marketing strategy that small dairy farms could utilize to harness the shifting consumer preferences toward specialty and local milk.

References

Berry, S. T. (1994). Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, 25(2), 242-262.

Berry, S., Levinsohn, J., & Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica: Journal of the Econometric Society*, 63(4), 841-890.

Campbell, B. L., Khachatryan, H., Behe, B. K., Dennis, J., & Hall, C. (2014). US and Canadian consumer perception of local and organic terminology. *International Food and Agribusiness Management Review*, 17(2), 21-40.

Caputo, V., Nayga Jr, R. M., & Scarpa, R. (2013). Food miles or carbon emissions? Exploring labeling preference for food transport footprint with a stated choice study. *Australian Journal of Agricultural and Resource Economics,* 57(4), 465-482.

Connolly, C., & Klaiber, H. A. (2014). Does organic command a premium when the food is already local? *American Journal of Agricultural Economics*, *96*(4), 1102-1116.

Darby, K., Batte, M. T., Ernst, S., & Roe, B. (2008). Decomposing local: A conjoint analysis of locally produced foods. *American Journal of Agricultural Economics*, 90(2), 476-486.

Grebitus, C., Lusk, J. L., & Nayga Jr, R. M. (2013). Effect of distance of transportation on willingness to pay for food. *Ecological Economics,* 88, 67-75.

Kemp, K., Insch, A., Holdsworth, D. K., & Knight, J. G. (2010). Food miles: Do UK consumers actually care? *Food Policy,* 35(6), 504-513.

Khanal, B., Lopez, R.A., and Azzam, A. (2020). Testing local bias in food consumption: The case of fluid milk. *Agribusiness: An International Journal*. 36(2), 339-344.

Lopez, E., & Lopez, R. A. (2008). Demand for differentiated milk products: implications for price competition. *Agribusiness: An International Journal*, 25(4), 453-465.

Low, S. A., Adalja, A., Beaulieu, E., Key, N., Martinez, S., Melton, A., ... & Jablonski, B. B. (2015). *Trends in US Local and Regional Food Systems: A Report to Congress*. USDA, Economic Research Services, Administrative Publication No. AP-068.

Nevo, A. (2000). A practitioner's guide to estimation of random-coefficients logit models of demand. *Journal of Economics & Management Strategy*, 9(4), 513-548.

Nevo, A. (2001). Measuring market power in the ready-to-eat cereal industry. *Econometrica*, 69(2), 307-342.

USDHHS. (2017). *Sanitation Compliance and Enforcement Ratings of Interstate Milk Shippers*. U.S. Department of Health and Human Services, Public Health Service, Food and Drug Administration. Retrieved 2/10/2020 from https://www.fda.gov/media/104934/download.

Vincent, D. W. (2015). The Berry–Levinsohn–Pakes estimator of the random-coefficients logit demand model. *The Stata Journal*, *15*(3), 854-880.

Wolf, C. A., Tonsor, G. T., & Olynk, N. J. (2011). Understanding US consumer demand for milk production attributes. *Journal of Agricultural and Resource Economics*, 36(2), 326-342.

Table 1: Descriptive statistics

Variable	Mean	Min	Max	Std. Dev
Product characteristics				
Price (per gallon)	6.76	3.19	10.27	1.38
Distance ('00 miles)	1.52	0.10	8.31	1.63
State local (1/0)	0.27	0	1	0.45
Organic (1/0)	0.307	0	1	0.46
Fat %	1.44	0	3.5	0.28
Share	0.033	0.001	0.43	0.06
Consumer characteristics				
Income ('000 dollars)	58.69	-11.4	1169	78.00
Age (years)	51.59	15.0	85.0	16.06
Children	0.76	0	9	1.08

Table 2: Demand parameter estimates for localness of milk

	Mean	SE	Mean	SE
	Distance-based		State-Local	
Product characteristics				
Price	-0.21**	0.10	-0.24	0.80
Local	-36.44***	3.26	-79.55***	24.30
Organic	-1.35***	0.09	-1.91***	0.10
Fat	1.15***	0.10	1.72***	0.11
Consumer characteristics				
Children (number)	-24.62***	5.09	-13.14***	3.71
Income (in $1,000)	0.20***	0.013	0.15***	0.03
Age (years)	-34.63***	3.26	-26.27***	0.42
Unobserved	0.01	0.03	-0.003	0.04
Interactions				
Local*Children	-0.38***	0.12	-0.15*	0.07
Local*Income	-10.8***	2.17	39.44***	4.52
Local*Age	0.44	1.67	0.17	0.42
Local*unobserved	0.15	0.62	1.36	0.03
Constant	17.8***	1.38	13.11***	1.52
Year Fixed Effects		Yes		
First stage F statistics		211.27		

*, **, and *** indicate the level of significance at less than 10%, 5%, and 1%, respectively.