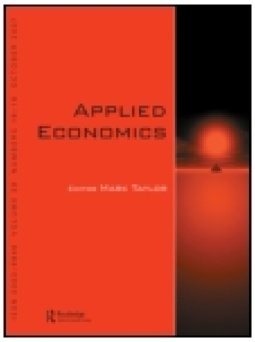

Applied Economics

ISSN: 0003-6846 (Print) 1466-4283 (Online) Journal homepage: http://www.tandfonline.com/loi/raec20

Demand system estimation in the absence of price data: an application of Stone-Lewbel price indices

César E. Castellón, Tullaya Boonsaeng & Carlos E. Carpio

Applied Economics, 2015
Vol. 47, No. 6, 553–568, http://dx.doi.org/10.1080/00036846.2014.975332

Demand system estimation in the absence of price data: an application of Stone-Lewbel price indices

César E. Castellón[a], Tullaya Boonsaeng[b] and Carlos E. Carpio[b,*]

[a]*The John E. Walker Department of Economics, Clemson University, Clemson, SC 29634-0313, USA*
[b]*Department of Agricultural and Applied Economics, Texas Tech University, Texas, TX 79409-2132, USA*

This article evaluates the feasibility of estimating a system of demand equations in the absence of price information using the approach developed by Lewbel (1989). Stone-Lewbel (SL) price indices for commodity groups are constructed using information on the budget shares and the Consumer Price Indices (CPIs) of the goods comprising the commodity groups, which allows for household-level prices to be recovered. This study evaluates how susceptible are elasticities and marginal effects estimates from traditional parametric demand systems to the CPI used in the construction of the SL prices. To do this, three alternative regional CPIs are considered for the construction of the SL prices: monthly, quarterly and a constant (unity) price index. Elasticities and marginal effect estimates are computed for eight food commodity groups using the Exact Affine Stone Index (EASI) model as the parametric demand system and data from the United States Consumer Expenditure Survey. The estimates proved to be robust to the alternative regional CPIs considered in the construction of SL price indices, even to the absence of one. Hence, the results suggest that it is possible to accurately estimate a demand system even in the absence of price information.

Keywords: food demand; Exact Affine Stone Index demand model; Consumer Expenditure Survey; Consumer Price Index

JEL Classification: D12; Q19

I. Introduction

Estimation of demand systems allows economists to compute demand elasticities for composite or individual commodities. These estimates find application in analysing market changes, tax incidence, consumption patterns and international trade, among others. Demand systems parameter estimates are also

*Corresponding author. E-mail: carlos.carpio@ttu.edu

used in policy analysis as most systems of equations allow for indirect utility and cost functions to be recovered. Much of the demand analysis literature uses cross-sectional data from micro-level household surveys because of their higher availability and lower collection cost compared to panel data. However, a common limitation of these data is the lack of price information, an important variable in estimating demand systems.[1] For example, the US Bureau of Labor Statistics (BLS) annual survey of consumer expenditures (Consumer Expenditure Survey (CEX)) does not collect price data for goods and services purchased.

Several approaches are used in the literature to compensate for this lack of price data. Some CEXs collect data on both quantities purchased and expenditures, allowing for unit values to be calculated (expenditure divided by quantities) and used as proxies for prices (e.g., Cox and Wohlgenant, 1986; Deaton, 1988). Another common approach is to incorporate external sources of price variability, such as Consumer Price Indices (CPIs), to account for missing prices (e.g., Kastens and Brester, 1996; Seale *et al.*, 2003). However, studies conducted by Slesnick (2005) and Hoderlein and Mihaleva (2008) have found this approach to be problematic as it does not account for spatial and household variability.

This article empirically evaluates the approach proposed by Lewbel (1989) that allows for the construction of household-level price indices (Stone-Lewbel (SL) prices) for commodity groups using as inputs CPIs and the budget shares of the subgroups of the commodities of interest. Hoderlein and Mihaleva (2008) found that the use of SL price indices results in a more precise and plausible estimated demand model than the use of CPIs only. Nevertheless, a question remains about the selection of the CPIs for the construction of the SL prices as the time period for which a CPI is measured can range from a month to a year, and it can be regionally or demographic specific. Therefore, the question of how dependent the demand estimation results are on the selected CPI for the construction of SL prices becomes relevant in practical settings. This study considers three alternative regional CPIs for the construction of SL prices which, in turn, are used to estimate three demand systems for eight food commodities using household-level data for the United States. Elasticities, marginal effects and parameter estimates are compared across the systems using each of the price series to derive conclusions regarding the effect of the alternative CPIs. The results suggest that it is possible to accurately estimate a demand system even in the absence of price information. This result is driven by the fact that most of the variation in SL prices comes from household heterogeneity and not from the CPIs.

II. Conceptual Framework

Stone-Lewbel price indices

SL price indices exploit the variation in the composition of expenditures within commodity groups to identify household-level commodity price indices. Specifically, if the between-group utility function is weakly separable and the within-group utility functions are assumed to be of the Cobb–Douglas form, then SL prices take the form (Lewbel, 1989)

$$v_{li} = \frac{1}{k_i}\prod_{j=1}^{n_i}\left(\frac{p_{ij}}{w_{lij}}\right)^{w_{lij}} \tag{1}$$

where w_{lij} is the budget share for good (or commodity subgroup) j within commodity group i for a given household $l; k_i$ is a scaling factor for commodity group i constructed using the subgroup budget shares of the reference household[2] ($k_i = \prod_{j=1}^{n_i} \bar{w}_{ij}{}^{-\bar{w}_{ij}}$); and p_{ij} is the price for good j within commodity group i. Equation 1 implies that household-level price indices for each commodity group can be calculated using individual goods' budget shares (w_{lij}) and price indices (p_{ij}). Specifically, we use three alternative p_{ij}'s: monthly, quarterly and unity regional CPIs. These indices are discussed in further detail later on.

The LA/EASI demand system

This study uses the Exact Affine Stone Index (EASI) demand system recently proposed by Lewbel and Pendakur (2009). Some of the advantages of this system relative to other traditional demand systems

[1] Though this problem is characteristic of cross-sectional data, it is not endemic to it; Carliner (1973) experienced the same limitation when working with panel data.
[2] The reference household is the household with average budget shares.

like the AIDS include (1) the EASI model allows for more flexible income expansion paths (Engel curves) since the budget shares can be polynomials of any order in the log of real expenditures and (2) the budget share error terms can be rationalized as unobserved preference heterogeneity. In addition, similar to the AIDS model, the EASI demand system includes a convenient linear approximation (LA) that uses the stone price index[3] to circumvent a nonlinear specification for real expenditures.

The LA/EASI system of demand budget share equations is defined as:

$$\boldsymbol{w_l} = \sum_{r=0}^{5} \boldsymbol{b_r} y_l^r + \boldsymbol{C z_l} + \boldsymbol{D z_l} y_l + \boldsymbol{A p_l} + \boldsymbol{B p_l} y_l + \varepsilon_l \tag{2}$$

where index l corresponds to a household; y_l is the total real expenditures ($y_l = \ln x_l - \boldsymbol{p_l}'\boldsymbol{w_l}$); x_l is the total nominal expenditures; $\boldsymbol{p_l}$ is an n vector of commodities' log price indices faced by household l; $\boldsymbol{w_l}$ is an n vector of commodities' demand budget shares; z_l is an m vector of socio-demographic characteristics; and $\boldsymbol{C}, \boldsymbol{D}, \boldsymbol{A}, \boldsymbol{B}$ and $\boldsymbol{b_r}$ are matrices and vectors of parameters to be estimated. Equation 2 is a reduced form of Lewbel and Pendakur's (2009) original demand equation in which an interaction term between socio-demographic characteristics and prices has been omitted to reduce the number of parameters to be estimated.[4] The system of N equations of the form in Equation 2 satisfies adding-up and homogeneity restrictions if

$$1_{\boldsymbol{n}}'\boldsymbol{b}_0 = 1,\ 1_{\boldsymbol{n}}'\boldsymbol{b_r} = 0\ \forall \boldsymbol{r} \neq 0,$$

and

$$1_{\boldsymbol{n}}'\boldsymbol{A} = 1_{\boldsymbol{n}}'\boldsymbol{B} = 0_{\boldsymbol{n}}; 1_{\boldsymbol{m}}'\boldsymbol{C} = 1_{\boldsymbol{m}}'\boldsymbol{D} = 0_{\boldsymbol{m}} \tag{3}$$

where symmetry of the Slutsky matrix is ensured by the symmetry of the nxn matrices **A** and **B**.

In short, the LA/EASI model possesses a set of desirable properties while retaining the familiar features that popularized the AIDS model. However, this model does not yield traditional Marshallian demand functions but rather what Lewbel and Pendakur (2009) describe as implicit Marshallian demand equations. These implicit Marshallian demand equations of the form in Equation 2 are Hicksian demands where the utility term has been approximated using total real expenditures. As a consequence, Marshallian demand elasticities cannot be directly derived from this equation. This study follows Lewbel and Pendakur's (2009) suggestion and estimates Hicksian demand and expenditure elasticities, subsequently recovering the Marshallian demand elasticities using the Slutsky equation (See Lewbel and Pendakur 2009, p. 836 and Appendix VII).

III. Data

Description

This research used data from the BLS, specifically from the CEX in addition to monthly and quarterly CPIs. The CEX data consist of two independent surveys: the Diary Survey and the Interview Survey. For the CEX Diary Survey, the one used in this study, households kept a 2-week diary of all daily food purchases. In addition, this survey also collected information on household characteristics. Households' daily expenditures on specific food products were totalled to obtain biweekly expenditures on aggregate food subgroups and groups (Table 1). Pooled cross-sectional data were constructed by grouping CEX and CPIs data from years 2002 to 2006. This pooled cross-sectional data set initially contained 36 364 households. Households reporting zero or negative values of income and total expenditures were omitted as were those with missing values for socio-demographic variables and outliers (observations in the upper percentile) in commodity group

[3] Lewbel and Pendakur (2009) conducted an empirical comparison between the actual model and its linear approximation without finding any major differences.
[4] To analyse the sensitivity of the results to the exclusion of this interaction, we estimated an LA/EASI model with the interaction terms between prices and socio-demographic variables; however, the results were similar to those using the reduced model in Equation 2.

Table 1. Commodity group composition and summary statistics

Commodity groups	Group composition	Mean budget share (%)	Level of censoring (%)
Cereals and bakery	(1) Cereals (2) Bakery products	15	6
Meats and eggs	(1) Beef (2) Pork (3) Poultry (4) Fish and sea food (5) Eggs (6) Other meats	23	9
Dairy	(1) Milk (2) Cheese (3) Ice cream (4) Other dairy products	12	8
Fruit and vegetables	(1) Fresh fruit (2) Fresh vegetables (3) Processed fruit and vegetables	15	9
Nonalcoholic beverages	(1) Juice and soda (2) Coffee and tea	12	11
Fats and oils	(1) Butter and margarine (2) Salad dressing (3) Fats and oils (4) Other fats	3	35
Sugar and other Sweets	(1) Sugar (2) Candies (3) Other sweets	4	33
Miscellaneous goods	(1) Soups (2) Prepared foods (3) Snacks (4) Seasoning (5) Baby food (6) Other foods	16	11

expenditures.[5] To make sure these observations were not systematically different from the rest of the sample, we compared summary statistics for household characteristics and commodities' budget shares between the initial and final data set; no meaningful differences were found in the composition of the final sample. The resulting final data set contained 30 768 households.

Using established US Department of Agriculture (USDA) nutrition-based guidelines from the Quarterly Food At-Home Price Database, this study considered the following eight commodity groups: (1) cereal and bakery products, (2) meats and eggs, (3) dairy, (4) fruits and vegetables, (5) nonalcoholic beverages, (6) fats and oils, (7) sugar and other sweets and (8) miscellaneous foods. Detailed information on these food groups and their corresponding subgroups is shown in Table 1. This classification is also consistent with the one used by the BLS for the construction of CPIs.

The food subgroups monthly and quarterly CPIs reported by the BLS are only provided at the national level. Monthly and quarterly regional CPIs for the Northeast, Midwest, West and South census regions are only provided for more aggregate good categories (i.e., CPI for all expenditure items). Thus, in order to account for regional food price variation, regional CPIs for the subgroups were constructed by deflating the national monthly and quarterly food subgroups CPIs using the corresponding regional (Northeast, Midwest, West and South) CPIs for all expenditure items.[6] Although this procedure assumes constant relative price differences among all food subgroups between two regions at any time period, the resulting CPIs incorporate all the price information made available by the BLS to reflect both temporal and regional price variation. To produce consistent regional monthly and quarterly regional CPIs series over time, the average CPI from 2002 to 2006 was used as the base period (2002–2006 = 100).

[5] Observations with values of income and total expenditures below or equal to zero correspond to households with severe 'missing data' problems. For example, total zero expenditures correspond to a case where not only all the values of the dependent variables in the system are equal to zero, but also all the prices and the expenditure variable values are 'missing'. The cases of negative and zero income, as described by the BLS, are believed to be mainly due to nonresponse to questions about income. Following previous studies (e.g., Raper *et al.*, 2002), we deleted these observations from the sample.
[6] An alternative to the CPI for all expenditure items is the CPI for food at home which is also available at the regional level. The results were robust to the regional CPI used for deflating the national food subgroups CPIs.

Stone-Lewbel prices

Three series of SL prices were constructed by substituting alternative regional CPIs (monthly, quarterly and unity) for the prices (p_{ij}) in Equation 1. Hence, in addition to the temporal and regional variation reflected in the CPIs, SL prices also account for household-level price variation through the use of the subgroup budget shares w_{lij}. However, in the case of unity CPIs, which aim to simulate a scenario where no price index information is available, SL price indices are only derived from the subgroup budget shares. Although intuitively a more disaggregated CPI would be preferred, there might be situations when this is not possible.[7]

SL commodity price indices for nonconsuming households (i.e., households with subgroup commodity shares w_{lij} equal to zero) were estimated using a regression imputation method (Cox and Wohlgenant, 1986; Alfonzo and Peterson, 2006; Lopez, 2011). It regressed the log SL price indices for uncensored observations obtained from Equation 1 on a set of demographic characteristics; ordinary least square parameters estimates were then used to recover log SL prices for households with subgroup-censored expenditure information.[8]

Summary statistics

Summary statistics and commodity group composition are also shown in Table 1. The degree of purchase censoring at 2-week frequencies ranged from 6% for cereal and bakery products to 35% for fats and oils. Those groups with the highest percentage of purchase censoring are associated with the smallest budget shares.

Descriptions and summary statistics of the demographic variables used to account for household heterogeneity are detailed in Table 2.[9] In 84% of the households, the reference person[10] is more than 30 years old, while the predominant racial group is Caucasian. In addition, 86% of the households have at least one adult female and 11% of the reference persons self-identify as Hispanics. To assess the representativeness of the CEX data, the statistics presented in Table 2 were compared with summary statistics for the same variables from the United States Census Bureau Current Population Survey for the period from 2003 to 2006. The results from both surveys are similar.

IV. Estimation Procedures

Censored approximated LA/EASI demand model

The high proportion of individuals reporting zero expenditure for some food groups required the use of procedures that account for the censored distribution of these responses. While several methods are available for estimating a system of censored demand equations, this study used the two-step procedure of Shonkwiler and Yen (1999) based on the following system of equations:

$$w_{li}^{*} = f(\boldsymbol{p}, \boldsymbol{z}_l, y_l; \boldsymbol{\theta}_i) + \varepsilon_{li}; \quad d_{li}^{*} = \boldsymbol{s}_l^{'}\boldsymbol{\rho}_i + \mu_{li} \tag{4}$$

$$d_{li} = \begin{cases} 1 \ if \ d_{li}^{*} > 0 \\ 0 \ if \ d_{li}^{*} \leq 0 \end{cases} \quad w_{li} = d_{li} w_{li}^{*} \tag{5}$$

$$(i = 1, 2, \ldots, N; l = 1, 2, \ldots, L)$$

where for the ith commodity group and lth household, w_{li}^{*} is the latent variable for demand budget share; d_{li}^{*} is a latent variable defining the sample selection in Equation 4, w_{li} and d_{li} are the observed dependent variables; $f(\boldsymbol{p}, \boldsymbol{z}_l, y_l; \boldsymbol{\theta}_i)$ represents a demand equation of the form in Equation 2 where

[7] To assess the relevance of SL prices for our data, we also estimated a complete demand system using only monthly CPIs as proxy for prices. Results obtained for this system included positive compensated own-price elasticity for one of the commodity groups.

[8] To test the sensitivity of our results to the presence of censored observations, we ran a full system of equations using only the uncensored observations. We found our estimates to be robust even when using only households with positive expenditures.

[9] Different sets of demographic variables were used at the different stages of the estimation process to avoid multi-collinearity issues. The variables' superscripts in Table 2 denote the model in which the variable was used as control.

[10] The reference person is defined by the BLS as the person who owns or rents the home.

Table 2. Descriptive statistics of household composition and characteristics

Category	Variable	Definition	Mean	SD	Min	Max
Continuous variables						
	Family size*†•	No. of members living in the household	2.56	1.460	1	9
	Proportion of persons below 18†•		0.36	0.481	0	1
	Annual income•	Annual family income before taxes	57 007.23	53 222.170	1	694 723
	Total food expenditures	Biweekly food expenditures	136.18	103.20	0.25	970.99
Dummy variables (yes = 1, no = 0)						
Education level of the reference person	No college*†•	Reference person has less than college education	0.14	0.345	0	1
	Some college*†•	Reference person has some college education	0.56	0.496	0	1
	College	Reference person has at least a college degree	0.30	0.457	0	1
Region of residence	North region*†•	Household is located in the north region of the country	0.18	0.385	0	1
	Midwest region*†•	Household is located in the midwest region of the country	0.26	0.436	0	1
	South region*†•	Household is located in the south region of the country	0.33	0.472	0	1
	West region	Household is located in the west region of the country	0.23	0.421	0	1
Age of the reference person	<25*†•	Reference person is younger than 25	0.06	0.243	0	1
	≥25–30*†•	Reference person is at least 25 but younger than 30	0.07	0.263	0	1
	≥30–40*†•	Reference person is at least 30 but younger than 40	0.20	0.398	0	1
	≥40–50*†•	Reference person is at least 40 but younger than 50	0.22	0.413	0	1
	≥50–60*†•	Reference person is at least 50 but younger than 60	0.24	0.429	0	1
	>60	Reference person is older than 60	0.20	0.402	0	1
Racial group of the reference person	White*†•	Reference person self-identifies as White	0.84	0.368	0	1
	Black*†•	Reference person self-identifies as Black	0.11	0.309	0	1
	Asian*†•	Reference person self-identifies as Asian	0.04	0.192	0	1
	Other	Reference person self-identifies as neither White, Black or Asian	0.02	0.125	0	1
Year in which the survey was collected	2002†•	Household was interviewed in year 2002	0.18	0.385	0	1
	2003†•	Household was interviewed in year 2003	0.19	0.394	0	1
	2004†•	Household was interviewed in year 2004	0.21	0.407	0	1
	2005†•	Household was interviewed in year 2005	0.21	0.409	0	1
	2006	Household was interviewed in year 2006	0.20	0.404	0	1
	Hispanic*†•	Reference person self-identifies as Hispanic	0.11	0.311	0	1
	Female adult unemployment†•	Reference person is female and unemployed	0.13	0.341	0	1
	Presence of a female adult†•	There is at least one female member older than 20 in the household	0.86	0.351	0	1
	Age of female adult†	There is at least one female adult younger than 35 in the household	0.26	0.439	0	1

Notes: *Refers to demographic variables used in the censored LA/EASI model.
†Refers to demographic variables used in the probit model.
•Refers to demographic variables used to regress SL prices.

$\boldsymbol{\theta}_i$ is a vector of parameter estimates, $\boldsymbol{p}$ is a vector of prices, $\boldsymbol{z}_l$ is a vector of socio-demographic characteristics and y represents real expenditures; $\boldsymbol{s}_l$ is a vector of socio-demographic characteristics explaining the sample selection process; and $\boldsymbol{\rho}_i$ is the vector of parameters for the sample selection equation.

The procedure involves the following three steps: (1) maximum likelihood (ML) probit estimates are obtained for $\boldsymbol{\rho}_i$; (2) the vector of parameter estimates $\hat{\boldsymbol{\rho}}_i$ is then used to calculate $\hat{\Phi}_{li}$ and $\hat{\phi}_{li}$, which represent estimates for the cdf and pdf of μ_{li}; and (3) estimates for the parameters in θ_i are obtained using equations of the form

$$\boldsymbol{w_l} = \hat{\boldsymbol{\Phi}}_{\boldsymbol{l}}\left(\sum_{r=0}^{5} \boldsymbol{b}_r y_l^r + \boldsymbol{C}\boldsymbol{z_l} + \boldsymbol{D}\boldsymbol{z}_l y_l + \boldsymbol{A}\boldsymbol{p_l} + \boldsymbol{B}\boldsymbol{p}_l y_l\right) + \hat{\phi}_{\boldsymbol{l}}\boldsymbol{\delta} + \boldsymbol{\varepsilon_l} \tag{6}$$

where $\hat{\boldsymbol{\Phi}}_{\boldsymbol{l}}$ and $\hat{\phi}_l$ are nxn identity matrices where the ones have been replaced by the cdf and pdf values of μ_{li}, and $\boldsymbol{\delta}$ is an n vector of parameters to be estimated. We will refer to Equation 6 as the censored LA/EASI demand system for household l. Economic theory does not provide any guidance regarding the selection of socio-demographic variables for the sample section probit model (s_l vector) and those included in the demand equation (z_l vector). However, additional demographic variables are included in the s_l vector to avoid potential multicollinearity problems in the estimation of Equation 6 (see footnotes in Table 2). Moreover, our main results are robust to changes in the composition of the sets of socio-demographic variables.

Elasticities and demographic effects are derived from Equation 6 (Yen *et al.*, 2002; Yen and Lin, 2006). It can be shown that compensated price elasticities (e_{ij}^*) in the censored LA/EASI demand system are given by

$$\xi = \varpi^{-1}\boldsymbol{\Phi}(\boldsymbol{A} + \boldsymbol{B}u) + \boldsymbol{\Omega}\varpi - \boldsymbol{I} \tag{7}$$

where ξ is an nxn matrix of compensated demand elasticities, ϖ is an identity matrix where the ones have been replaced by the commodities' budget shares, $\boldsymbol{\Omega}$ is an nxn matrix of ones and $\boldsymbol{I}$ is an identity matrix. Similarly, expenditure elasticities (η_i) from the implicit Marshallian demand equations are given by

$$\boldsymbol{\eta} = \varpi^{-1}(\boldsymbol{I} + \boldsymbol{\Phi\Theta p}')^{-1}\boldsymbol{\Phi\Theta} + 1_{\boldsymbol{n}} \tag{8}$$

where $\boldsymbol{\eta}$ is an n vector of expenditure elasticities and $\boldsymbol{\Theta}$ is the derivative of Equation 6 with respect to the real expenditures y_l, such that $\boldsymbol{\Theta} = \sum_{r=1}^{5} \boldsymbol{r}\boldsymbol{b_r} y^{r-1} + \boldsymbol{Dz} + \boldsymbol{Bp}$. The marginal effects of socio-demographic characteristics can also be derived from Equation 6; however, the formula is dependent upon the presence of the socio-demographic characteristics either in the budget share equation or in the probit model only, or in both equations.

In this study, the SAS MODEL procedure imposing the symmetry and adding-up restrictions was used to estimate the Seemingly Unrelated Regression estimators of the parameters in Equation 6 using all N equations. The use of all N equations is possible since the system in Equation 6 does not have a singular variance–covariance residual matrix (Yen *et al.*, 2002; Drichoutis *et al.*, 2008).

Given the likelihood of the correlation between error terms in each equation and the total real expenditures (y) (LaFrance, 1991; Lewbel and Pendakur, 2009, p. 834), the approach suggested by Blundell and Robin (2000) where each equation in the system in Equation 6 is augmented with the error term v from a reduced form of real expenditures was used here. Thus, each element in the vector of error terms ε_l in Equation 6 is rewritten as the orthogonal decomposition $\varepsilon_{li} = \omega_i v_l + u_{li}$, where $E(u_{li}|lnx_l, z_{li}, \ldots, z_{lm}, \log p_{li}, \ldots, \log p_{ln}) = 0$. The reduced form of y follows Blundell and Robin's (2000) specification and is defined as a function of a linear trend, log prices, demographic variables, interaction terms between socio-demographic characteristics and linear and higher-order terms of log income. The hypothesis that the ω_i parameters are different from zero was used to test the endogeneity of y (Blundell and Robin, 2000; Boonsaeng *et al.*, 2008).

To account for the use of two-step estimation procedures and the heteroscedasticity of the disturbances in the system of equations of the form in Equation 6, this study computed SEs for parameter, elasticities and marginal effect estimates using the nonparametric bootstrapping procedure outlined by Wooldridge (2002, p. 379) with 900 replications.

Comparison of elasticities and marginal effects

Compensated and expenditure elasticities were estimated for the average household using Equations 7 and 8. Uncompensated elasticities were recovered using the Slutsky equation. Marginal effects were also estimated for the average household. Two procedures were used here to assess differences across the demand systems' estimates. First, the percentage error between the elasticities and marginal effects obtained when using monthly CPI-based SL prices and those obtained when using quarterly and unity CPI-based SL prices were computed and compared. Second, to formally analyse the statistical difference between parameter estimates, bootstrapping procedures were employed because the samples used to estimate the SEs for these values were not drawn from independent populations but from the same population, and, therefore, statistical methods of comparison of means such as the student's *t*-test were inappropriate.

The comparison using bootstrapping procedures involved the following three steps: (1) the parameter estimates from the bootstrapping samples were used to obtain the elasticities and marginal effect estimates for each sample; (2) the difference in parameters, elasticities and marginal effects between the systems was calculated for each bootstrap sample using quarterly and unity CPI-based SL prices and the estimates of the system with monthly CPI-based SL prices (these estimates were used as benchmark); and (3) using the distributions of differences, 95% confidence intervals for all parameter, elasticity and marginal effect estimates were constructed.

V. Results

Endogeneity tests

The null hypothesis that real expenditure is exogenous is rejected (5% level) in five of the eight demand equations for the systems using monthly and quarterly CPI-based SL prices, and in six of the eight demand equations for the system using unity CPI-based SL prices. Endogeneity of expenditures is problematic since it can result in biased and inconsistent estimates; however, is worth noting that the parameter, elasticity and marginal effect estimates of the models were robust to the correction for endogeneity.

Comparison of models

Because monthly CPI-based SL prices are the most disaggregated proxy for prices, the results from the demand system using these indices are used as benchmarks throughout this section. Elasticities obtained using the monthly and unity CPI-based SL prices are shown in Tables 3 and 4, respectively. Elasticity estimates for the specification using quarterly CPI-based SL prices are not reported due to constraints on space; however, these values were always in between those from the specifications using monthly and unity CPI-based SL prices.

The percentage error for the expenditure elasticities ranged in absolute terms from 0.002% to 0.05% for the quarterly CPI-based SL prices and from 0.02% to 0.86% for the unity CPI-based SL prices. For own-price elasticities, the percentage error derived from absolute differences ranged from 0.004% to 0.20% for the quarterly CPI-based SL prices and from 0.09% to 2.09% for the unity CPI-based SL prices.

The mean percentage errors derived from absolute differences for cross-price elasticities were 1.36% and 10.36% for the quarterly and the unity CPI-based SL prices, respectively. Similarly, the mean percentage errors of the marginal effects were 5.91% and 12.57% for the quarterly and unity CPI-based SL prices.[11] The higher mean percentage errors for cross-price elasticities and marginal effects than for own-price and expenditure elasticities can be explained by the higher number of parameter estimates not statistically different from zero (5% level) for the cross-price elasticities and marginal effects. The fact that the elasticity estimates using quarterly CPI-based SL prices are closer to the estimates obtained using monthly CPI-based SL prices than those obtained using unity CPI-based SL prices comes as no surprise, as quarterly CPIs are more disaggregated than the unity CPIs which do not vary across time or region. Nonetheless, the elasticity and marginal effects estimates obtained using the

[11] We also estimated percentage errors for parameter estimates. Mean percentage errors for quarterly and unity CPI-based SL prices were 1.08% and 415%, respectively. The high mean percentage error for unity CPI-based SL prices can be explained by the presence of parameter estimates not statistically different from zero. A table of parameter estimates is available from the authors upon request.

Table 3. Estimated uncompensated and expenditure elasticities when employing monthly CPI-based SL price index

Quantity demanded	Prices: Cereal and bakery	Meats and eggs	Dairy	Fruit and vegetables	Nonalcoholic beverages	Fats and oils	Sugar and other sweets	Miscellaneous goods	Expenditure
Cereal and bakery	−0.7208**	−0.1818**	−0.0563**	−0.1049**	−0.0408**	−0.0066	0.0254**	−0.0382**	1.1241**
	(0.0131)	(0.0119)	(0.0093)	(0.0097)	(0.0100)	(0.0059)	(0.0065)	(0.0100)	(0.0222)
Meats and eggs	−0.0879**	−0.5287**	−0.0949**	−0.0548**	−0.042**	−0.0019	−0.0170**	−0.1198**	0.9471**
	(0.0056)	(0.0105)	(0.0055)	(0.0062)	(0.0057)	(0.0033)	(0.0038)	(0.0074)	(0.0152)
Dairy	−0.0178*	−0.1485**	−0.5789**	−0.0354**	0.0144	−0.0156**	0.0199**	−0.0055	0.7675**
	(0.0099)	(0.0124)	(0.0157)	(0.0103)	(0.0103)	(0.0060)	(0.0069)	(0.0104)	(0.0234)
Fruit and vegetables	−0.1041**	−0.1301**	−0.0712**	−0.5971**	−0.1021**	−0.0148**	−0.0052	−0.1276**	1.1524**
	(0.0085)	(0.0110)	(0.0081)	(0.0134)	(0.0089)	(0.0051)	(0.0057)	(0.0095)	(0.0180)
Nonalcoholic beverages	−0.0182	−0.0718**	−0.0050	−0.0886**	−0.7502**	−0.0069	0.0192**	−0.0077	0.9293**
	(0.0115)	(0.0130)	(0.0109)	(0.0115)	(0.0187)	(0.0070)	(0.0089)	(0.0119)	(0.0235)
Fats and oils	−0.0397**	−0.0662**	−0.0846**	−0.0667**	−0.0525**	−0.7928**	0.0303	−0.1342**	1.2066**
	(0.0199)	(0.0213)	(0.0173)	(0.0186)	(0.0202)	(0.0397)	(0.0263)	(0.0176)	(0.0350)
Sugar and other sweets	0.0919**	−0.0679**	0.0312**	0.0120	0.0477**	0.0335*	−1.1087**	0.0259	0.9343**
	(0.0177)	(0.0210)	(0.0168)	(0.0170)	(0.0205)	(0.0217)	(0.0350)	(0.0177)	(0.0367)
Miscellaneous goods	−0.0172**	−0.1810**	−0.0311**	−0.0962**	−0.0159*	−0.0268**	0.0046	−0.6456**	1.0092**
	(0.0070)	(0.0111)	(0.0068)	(0.0081)	(0.0079)	(0.0040)	(0.0047)	(0.0124)	(0.0187)

Notes: SEs in parentheses.
*Denotes significance at the 10% level.
**Denotes significance at the 5% level.

Table 4. Estimated uncompensated and expenditure elasticities when employing unity CPI-based SL price index

	Prices								
Quantity demanded	Cereal and bakery	Meats and eggs	Dairy	Fruit and vegetables	Nonalcoholic beverages	Fats and oils	Sugar and other sweets	Miscellaneous goods	Expenditure
Cereal and bakery	−0.7220**	−0.1768**	−0.0519**	−0.1062**	−0.0452**	−0.0039	0.0246**	−0.0385**	1.1200**
	(0.0118)	(0.0048)	(0.0092)	(0.0106)	(0.0090)	(0.0051)	(0.0054)	(0.0103)	(0.0246)
Meats and eggs	−0.0857**	−0.5271**	−0.0958**	−0.0547	−0.0435	−0.0027**	−0.0180	−0.1216**	0.9490**
	(0.0047)	(0.0053)	(0.0041)	(0.0046)	(0.0042)	(0.0017)	(0.0019)	(0.0057)	(0.0231)
Dairy	−0.0127**	−0.1497**	−0.5814**	−0.0361**	0.0136*	−0.0173**	0.0198	−0.0032**	0.7669**
	(0.0096)	(0.0056)	(0.0151)	(0.0108)	(0.0089)	(0.0063)	(0.0060)	(0.0108)	(0.0254)
Fruit and vegetables	−0.1064**	−0.1301**	−0.0721**	−0.5918**	−0.0982**	−0.0151**	−0.0067**	−0.1349**	1.1554**
	(0.0084)	(0.0061)	(0.0089)	(0.0138)	(0.0084)	(0.0049)	(0.0051)	(0.0114)	(0.0271)
Nonalcoholic beverages	−0.0247**	−0.0755**	−0.0066**	−0.0845**	−0.7509**	−0.0049**	0.0211	−0.0099**	0.9360**
	(0.0106)	(0.0074)	(0.0106)	(0.0122)	(0.0140)	(0.0061)	(0.0060)	(0.0134)	(0.0343)
Fats and oils	−0.0314**	−0.0700	−0.0890**	−0.0675**	−0.046**	−0.8094**	0.0364**	−0.1295**	1.2064**
	(0.0179)	(0.0073)	(0.0187)	(0.0183)	(0.018)	(0.0280)	(0.0129)	(0.0181)	(0.0374)
Sugar and other sweets	0.0883**	−0.0733	0.0302	0.0070	0.0520	0.0383**	−1.1071**	0.0238**	0.9408**
	(0.0152)	(0.0072)	(0.0152)	(0.0160)	(0.0134)	(0.0108)	(0.0139)	(0.0167)	(0.0362)
Miscellaneous goods	−0.0168**	−0.1812**	−0.0285**	−0.1011**	−0.0156**	−0.0253**	0.0046**	−0.6367**	1.0006**
	(0.0074)	(0.0053)	(0.0068)	(0.0087)	(0.0093)	(0.0039)	(0.0043)	(0.0125)	(0.0198)

Notes: SEs in parentheses.
*Denotes significance at the 10% level.
**Denotes significance at the 5% level.

three alternative specifications are approximately the same.

Even though the elasticities obtained using the alternative specifications are similar, the tests of differences using bootstrapping procedures revealed statistically significant differences at the 5% confidence level across the models. Specifically, seven of eight own-price and expenditure elasticities from the model using quarterly CPI-based SL prices were statistically different from those obtained from the model using monthly CPI-based SL prices. All the own-price elasticities and four of eight expenditure elasticities obtained from the demand model using the unity CPI-based SL prices were statistically different from those obtained from the model using monthly CPI-based SL prices. For the cross-price elasticities, 22 of the 56 were statistically different between the systems using quarterly and monthly CPI-based SL prices. Similarly, 20 of the 56 cross-price elasticities were statistically different among the models employing monthly and unity CPI-based SL prices.

Estimates of the marginal effects from the system using monthly CPI-based SL prices are provided in Table 5.[12] Results from the bootstrapping procedure indicate that at the 5% level, 102 of 120 marginal effects were not statistically different between the systems using quarterly versus monthly CPI-based SL prices. Similarly, 94 of 120 marginal effect estimates were not statistically different between the models using monthly versus unity CPI-based SL prices.

Another consideration is whether the use of different CPIs affected the precision of the parameter, elasticity and marginal effect estimates. This does not seem to be the case as estimated SEs for the elasticities and marginal effects were similar. While the number of significant parameters is smaller for the system using unity CPI-based SL prices, differences in the number of significant elasticities and marginal effects are small across the three systems (5% level).

Sources of variation in the SL price indices

The similarity in the three series of SL price indices as well as the similarity in the elasticities and marginal effects estimated using these series might be better understood when looking individually at the components of the SL price index.

Recall from Equation 1 that the SL price index for a specific good i, for a given household l, is constructed as the product of three components: (1) a scaling factor k_i, that is invariant across households, (2) a subgroup price component $\prod_{j=1}^{n_i} (p_{ij})^{w_{lij}}$ that varies across households because of its exponent and (3) a subgroup budget share component $\prod_{j=1}^{n_i} (1/w_{lij})^{w_{lij}}$ that is specific to each household. Table 6 presents the coefficient of variation (CV) for the monthly SL price indices for all commodity groups as well as for the price and budget share components of those price indices. Correlations between each component and its respective price index are also provided.

Notice that the variation in the price components is approximately 8% of the total variation of the price indices, whereas the variation in the budget share components is almost identical to that of the price indices. Also, there is a weak correlation between price components and the price indices and almost a perfect positive relation between the budget share components and the SL price indices. Thus, most of the variability of the price indices is being generated by the information in the subgroup budget shares with little or insignificant participation of the subgroup price indices, thereby reducing the relevance of the level of disaggregation of the price indices used in the analysis.

The previous result can be rationalized as follows, current CPIs are not disaggregated enough to generate meaningful differences in the variability of SL price indices, especially when compared with the variability in households' budget shares. For instance, consider the case of monthly regional CPIs (the benchmark for model comparison); a monthly regional CPI for meats implies that all households within the same region face the same price for meats at a given month, which is rather restrictive when compared with a price index that varies with household's food expenditures, such that only households with the same socio-

[12] Given space limitations we only report: (1) marginal effects of socio-demographic variables included in the demand equations and (2) marginal effects from models estimated using monthly regional CPI-based SL prices. Marginal effects of variables included only in the sample selection probit models, and marginal effects from the models estimated using quarterly or unity CPI-based SL prices are available from the authors upon request.

Table 5. Estimated socio-demographic marginal effects when employing monthly CPI-based SL price index

	Education		Region			Age of household head in years					Race				
Quantities demanded	No college	Some college	Northeast	Midwest	South	<25	≥25–30	≥30–40	≥40–50	≥50–60	White	Black	Asian	Family size	Hispanic
Cereal and bakery	−1.480*	−0.920	4.140**	1.100	1.550*	−10.380**	−8.690**	−8.410**	−7.360**	−5.100**	1.840	−2.150	−2.970	2.780**	−2.970**
	(0.799)	(0.563)	(0.852)	(0.711)	(0.719)	(2.333)	(1.646)	(1.319)	(1.029)	(0.794)	(1.755)	(1.884)	(2.231)	(0.605)	(0.807)
Meats and eggs	2.510**	1.940**	3.110**	0.990*	3.480**	−11.320**	−5.690**	−5.040**	−1.580**	0.300	−1.100	6.300**	1.150	3.220**	7.000**
	(0.724)	(0.485)	(0.681)	(0.618)	(0.583)	(1.831)	(1.216)	(0.969)	(0.754)	(0.670)	(1.677)	(1.837)	(2.110)	(0.427)	(0.922)
Dairy	−2.260**	−1.620**	0.940**	−0.370	−0.850**	−4.970**	−2.740**	−2.140**	−2.040**	−1.560**	2.580**	−4.680**	−5.690**	2.180**	−1.580**
	(0.539)	(0.362)	(0.472)	(0.409)	(0.379)	(1.281)	(0.844)	(0.664)	(0.565)	(0.466)	(1.081)	(1.194)	(1.448)	(0.343)	(0.523)
Fruit and vegetables	−5.320**	−4.600**	−0.460	−1.710**	−1.990**	−13.380**	−10.330**	−9.440**	−8.050**	−4.560**	1.140	0.710	13.170**	1.000**	6.090**
	(0.587)	(0.410)	(0.491)	(0.436)	(0.425)	(1.255)	(0.913)	(0.754)	(0.610)	(0.501)	(1.273)	(1.326)	(1.669)	(0.260)	(0.580)
Nonalcoholic beverages	1.370**	1.140**	0.990**	0.450	0.760**	6.110**	4.400**	5.320**	5.080**	4.170**	−1.460	−2.430**	−3.600**	0.680**	0.170
	(0.457)	(0.332)	(0.458)	(0.378)	(0.380)	(0.843)	(0.637)	(0.515)	(0.499)	(0.471)	(1.040)	(1.124)	(1.436)	(0.282)	(0.500)
Fats and oils	0.950**	0.500**	−0.450**	−0.040	−0.160	0.920**	0.730**	0.610**	0.330	0.220	−0.250	0.650	−1.050**	−0.530**	−0.270*
	(0.168)	(0.121)	(0.170)	(0.157)	(0.159)	(0.369)	(0.287)	(0.209)	(0.162)	(0.149)	(0.361)	(0.384)	(0.479)	(0.087)	(0.196)
Sugar and other sweets	0.570	0.580**	−1.440**	0.000	−0.480	−1.150**	−0.060	0.850*	1.310**	0.820**	0.310	0.230	−1.870*	0.040	−1.600**
	(0.337)	(0.237)	(0.324)	(0.293)	(0.295)	(0.557)	(0.457)	(0.369)	(0.341)	(0.337)	(0.754)	(0.818)	(0.968)	(0.158)	(0.385)
Miscellaneous goods	−3.490**	−1.890**	−4.190**	−0.490	−1.230**	5.120**	4.590**	3.060**	2.960**	1.630**	0.300	−4.500**	−4.040**	0.550**	−4.020**
	(0.514)	(0.357)	(0.516)	(0.393)	(0.376)	(1.019)	(0.709)	(0.523)	(0.449)	(0.406)	(0.968)	(1.085)	(1.209)	(0.241)	(0.460)

Notes: SEs in parentheses.
*Denotes significance at the 10% level.
**Denotes significance at the 5% level.

Table 6. Variation and correlation of the monthly SL price index components

Variable	Coefficient of variation	Correlations		
		SL price – cereals	$\Pi j(p_{ij})^{w_{lij}}$	$\Pi j(1/w_{lij})^{w_{lij}}$
SL price – cereals	0.2479	1.000		
$\Pi j(p_{ij})^{w_{lij}}$	0.0190	0.098	1.000	
$\Pi j(1/w_{lij})^{w_{lij}}$	0.2468	0.997	0.017	1.000
		SL price – meats	$\Pi j(p_{ij})^{w_{lij}}$	$\Pi j(1/w_{lij})^{w_{lij}}$
SL price – meats	0.4231	1.000		
$\Pi j(p_{ij})^{w_{lij}}$	0.0363	0.068	1.000	
$\Pi j(1/w_{lij})^{w_{lij}}$	0.4219	0.997	0.001	1.000
		SL price – dairy	$\Pi j(p_{ij})^{w_{lij}}$	$\Pi j(1/w_{lij})^{w_{lij}}$
SL price – dairy	0.4037	1.000		
$\Pi j(p_{ij})^{w_{lij}}$	0.0342	0.074	1.000	
$\Pi j(1/w_{lij})^{w_{lij}}$	0.4026	0.997	0.000	1.000
		SL price – fruit	$\Pi j(p_{ij})^{w_{lij}}$	$\Pi j(1/w_{lij})^{w_{lij}}$
SL price – fruit	0.2943	1.000		
$\Pi j(p_{ij})^{w_{lij}}$	0.0266	0.097	1.000	
$\Pi j(1/w_{lij})^{w_{lij}}$	0.2928	0.996	0.014	1.000
		SL price – beverages	$\Pi j(p_{ij})^{w_{lij}}$	$\Pi j(1/w_{lij})^{w_{lij}}$
SL price – beverages	0.2992	1.000		
$\Pi j(p_{ij})^{w_{lij}}$	0.0243	0.055	1.000	
$\Pi j(1/w_{lij})^{w_{lij}}$	0.2990	0.997	−0.024	1.000
		SL price – fats	$\Pi j(p_{ij})^{w_{lij}}$	$\Pi j(1/w_{lij})^{w_{lij}}$
SL price – fats	0.4741	1.000		
$\Pi j(p_{ij})^{w_{lij}}$	0.0354	0.084	1.000	
$\Pi j(1/w_{lij})^{w_{lij}}$	0.4723	0.997	0.018	1.000
		SL price – sugar	$\Pi j(p_{ij})^{w_{lij}}$	$\Pi j(1/w_{lij})^{w_{lij}}$
SL price – sugar	0.4051	1.000		
$\Pi j(p_{ij})^{w_{lij}}$	0.0220	0.062	1.000	
$\Pi j(1/w_{lij})^{w_{lij}}$	0.4043	0.999	0.012	1.000
		SL price – miscelaneous	$\Pi j(p_{ij})^{w_{lij}}$	$\Pi j(1/w_{lij})^{w_{lij}}$
SL price – miscelaneous	0.4400	1.000		
$\Pi j(p_{ij})^{w_{lij}}$	0.0279	0.031	1.000	
$\Pi j(1/w_{lij})^{w_{lij}}$	0.4395	0.998	−0.032	1.000

demographic characteristics (i.e., family composition, region of residence, occupation of head of household, etc.) face the same price for a given good.

Elasticities and marginal effects

This section focuses on the elasticities and marginal effects obtained from the system using monthly CPI-based SL prices since these results were used as benchmarks. Nonetheless, as shown above, the elasticity values and marginal effects across the three alternative specifications were similar.

Consistent with general economic theory, all own-price uncompensated elasticities are negative and statistically significant at the 5% level. Expenditure elasticities indicate no commodity group is inferior, an expected result given the broad level of aggregation. The absolute values for estimated cross-price elasticities are less than one, and cross-price effects indicate complementary relations across goods. Again, this result can be seen as a consequence of the high level of aggregation.

Marginal effect estimates of the dummy variables reported in Table 5 are measured, *ceteris paribus*, relative to those of the base (omitted) category: a household located in the West region whose reference person has at least a college degree, is older than 60, non-Hispanic and self-identifies as neither White/ Black nor Asian. For instance, the coefficient in Table 5, column 1 for fruits and vegetables, indicates that a household with a reference person with no college education spends \$5.32 less on fruits and vegetables biweekly than a household with a reference person with at least a college degree.

The results from the marginal effects are consistent with general expectations. Households with a less-educated reference person tend to spend less on fruits and vegetables and more on sweets than those with a higher-educated one. Larger households spend more on all commodity groups with exception of the fats and oils group than smaller ones. White households spend the most on the dairy and sweets commodity groups, Asian households, the most on the fruit and vegetables commodity groups, while black households spend the most on the meats commodity group. When age is used to identify the reference person, households with a younger reference person spend the most on the miscellaneous group, a result associated with a higher consumption level of ready-to-eat food and snacks; on the other hand, households with an older reference person seem to spend more in most of the categories, possibly due to larger household size or/and a higher income.

The estimated own-price elasticities reported here for the cereals, meats, dairy, fruits and vegetables, and fats and oils groups are not as elastic as those found in past research (see Raper *et al.*, 2002). Differences can also be seen in the estimates for the expenditure elasticities. In particular, the expenditure elasticity here for the Meats group is more inelastic than that reported by Raper *et al.* (2002), perhaps a consequence of the differences in the commodity groups chosen to be included in the system as well as the within-group aggregation. In addition, the data used by Raper *et al.* (2002) is from 10 years prior to this study. The magnitude of demand responsiveness of United States consumers may have changed over time

The estimates reported here were also compared with those found by Leffler (2012), who used US Homescan data from the AC Nielsen database to estimate a demand system with the same eight commodity groups considered in this study. The own-price elasticities reported here for the cereals, nonalcoholic beverages, fats, sweets and miscellaneous groups are similar to those obtained by Leffler (2012). Larger differences were observed between the own-price elasticities for the meats, dairy, and fruits and vegetables groups, with the elasticities reported being more inelastic than the ones presented by Leffler (2012). A second major difference involves the estimates for expenditure elasticities, as Leffler (2012) classified the meats and fruit and vegetables groups as luxury goods, whereas this study classified the cereals, fruit and vegetables, fats and miscellaneous groups as luxuries. These inconsistencies could be due to differences in the data used in the two studies. For instance, the AC Nielsen Homescan data provide information on market prices for all individual commodities, circumventing the price identification issue of this study. In addition, Homescan data are an annual record, while the CEX data used in this study are limited to a biweekly period.

Andreyeva *et al.* (2010) conducted a survey of reported uncompensated price elasticities for food commodities in the United States from 1838 to 2007, considering a total of 160 food demand studies and providing mean values and ranges for the uncompensated own-price elasticities of 16 commodity groups. Their study did not account for differences in methodology, year or data sources as their intention was to provide a benchmark of the reported price elasticities for major groups of food consumption in the literature. The estimated own-price elasticities for fruits and vegetables, dairy, nonalcoholic beverages and cereals groups were similar to the mean values reported for these groups by Andreyeva *et al.* (2010), and the estimates of this study for the own-price elasticities for fats and oils, sugar and other sweets, and meats were also within the ranges reported for these groups (Andreyeva *et al.*, 2010).

VI. Summary and Conclusions

Lewbel (1989) developed an approach for the construction of household level commodity price indices (SL prices) using only the budget shares and CPIs of the goods comprising the commodity groups. This study considers three alternative CPIs for the construction of the SL prices used in the estimation of a demand system: monthly, quarterly and unity, with the unity CPI being used to simulate a scenario where no price index information is available. The analysis of the performance of the three SL prices involved the comparison of the estimated elasticities, marginal effects and parameters obtained from demand models using household level data from the United States.

The results suggest that the low variability on current measures of CPIs from the BLS restricts their influence on the performance of SL price indices. Elasticities and marginal effect estimates from the demand models proved to be robust for the alternative CPIs considered in this study. Though statistical differences were found across estimates from the models using different CPI-based SL price indices, the empirical differences found across the models used here are quite small. Specifically, these differences are substantially smaller in comparison with those from other studies; that is, differences in elasticity estimates due to changes in the construction of SL prices are smaller than those found when employing different data sets (Leffler, 2012) or methodologies (Raper *et al.*, 2002).

Based on the results, this study concludes that incorporation of CPI data in the calculation of SL prices plays a limited role, thereby making it possible to accurately estimate a demand system in the absence of price information. However, more research is needed to evaluate the performance of unity CPI-based SL prices with other data sets.

The study has several limitations. Currently, the BLS does not provide regional CPIs for groups or subgroups of commodities. The regional CPIs used in this study were approximated using the national commodities CPIs and the aggregate regional CPIs. Even though this approximation represents a more disaggregate measure than the national price indices used by Hoderlein and Mihaleva (2008), future studies could use more disaggregated CPIs provided by the national statistical entities in several countries. For instance, the Ecuadorian National Institute of Statistics and the Mexican National Institute of Statistics both provide food commodity CPIs for the largest metropolitan areas.

In addition, the use of household-level surveys with information on expenditures and consumed quantities for individual commodities allows for the estimation of quality-corrected unit values (Cox and Wohlgenant, 1986; Deaton, 1988). A comparison of SL price indices relative to the use of quality-corrected unit values would provide another measure of the performance of SL indices as approximations for unobserved prices. A further comparison could be conducted using a privately owned database such as data from the AC Nielsen Homescan, which provides market price information for all the commodities within the survey.

Besides these limitations, this study provides a good foundation for further analysis of SL price indices as a way to incorporate price variation in demand studies where prices are unobserved or are observed with error. Specifically, it shows that the estimation of SL prices is not constrained by the availability of CPIs and that under current CPI estimates, the difference in results is negligible from a practical setting.

Funding

This project was supported by Agriculture and Food Research Initiative Competitive Contract no. 2011-67023-30058 from the USDA National Institute of Food and Agriculture.

References

Alfonzo, L. and Peterson, H. H. (2006) Estimating food demand in Paraguay from household survey data, *Agricultural Economics*, **34**, 243–57. doi:10.1111/j.1574-0864.2006.00122.x

Andreyeva, T., Long, M. W. and Brownell, K. D. (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**, 216–22. doi:10.2105/AJPH.2008.151415

Blundell, R. and Robin, J.-M. (2000) Latent separability: grouping goods without weak separability, *Econometrica*, **68**, 53–84. doi:10.1111/1468-0262.00093

Boonsaeng, T., Fletcher, S. M. and Carpio, C. E. (2008) European Union import demand for in-shell peanuts, *Journal of Agricultural and Applied Economics*, **40**, 941–51.

Carliner, G. (1973) Income elasticity of housing demand, *The Review of Economics and Statistics*, **55**, 528–32. doi:10.2307/1925683

Cox, T. L. and Wohlgenant, M. K. (1986) Prices and quality effects in cross-sectional demand analysis, *American Journal of Agricultural Economics*, **68**, 908–19. doi:10.2307/1242137

Deaton, A. (1988) Quality, quantity, and spatial variation of price, *The American Economic Review*, **78**, 418–30.

Drichoutis, A. C., Klonaris, S., Lazaridis, P. *et al*. (2008) Household food consumption in Turkey: a comment, *European Review of Agricultural Economics*, **35**, 93–98. doi:10.1093/erae/jbn010

Hoderlein, S. and Mihaleva, S. (2008) Increasing the price variation in a repeated cross section, *Journal of Econometrics*, **147**, 316–25. doi:10.1016/j.jeconom.2008.09.022

Kastens, T. L. and Brester, G. W. (1996) Model selection and forecasting ability of theory-constrained food demand systems, *American Journal of Agricultural Economics*, **78**, 301–12. doi:10.2307/1243704

LaFrance, J. T. (1991) When is expenditure “exogenous” in separable demand models?, *Western Journal of Agricultural Economics*, **16**, 49–62.

Leffler, K. (2012) Temporal aggregation and treatment of zero dependent variables in the estimation of food demand using cross-sectional data, Unpublished Master Thesis, Clemson University, Clemson, SC.

Lewbel, A. (1989) Identification and estimation of equivalence scales under weak separability, *The Review of Economic Studies*, **56**, 311–16. doi:10.2307/2297464

Lewbel, A. and Pendakur, K. (2009) Tricks with hicks: the EASI demand system, *The American Economic Review*, **99**, 827–63. doi:10.1257/aer.99.3.827

Lopez, J. A. (2011) A comparison of price imputation methods under large samples and different levels of censoring. Paper presented at the 2011 AAEA Annual Meeting, July 24–26, Pittsburgh, PA.

Raper, K. C., Wanzala, M. N. and Nayga Jr, R. M. (2002) Food expenditures and household demographic composition in the US: a demand systems approach, *Applied Economics*, **34**, 981–92. doi:10.1080/00036840110061959

Seale Jr, J. L., Marchant, M. A. and Basso, A. (2003) Imports versus domestic production: a demand system analysis of the US red wine market, *Review of Agricultural Economics*, **25**, 187–202. doi:10.1111/1467-9353.00053

Shonkwiler, J. S. and Yen, S. T. (1999) Two-step estimation of a censored system of equations, *American Journal of Agricultural Economics*, **81**, 972–82. doi:10.2307/1244339

Slesnick, D. T. (2005) Prices and demand: new evidence from micro-data, *Economics Letters*, **89**, 269–74. doi:10.1016/j.econlet.2005.05.034

Wooldridge, J. M. (2002) *Econometric Analysis of Cross-Section and Panel Data*, The MIT Press, Cambridge, MA.

Yen, S. T., Kan, K. and Su, S.-J. (2002) Household demand for fats and oils: two-step estimation of a censored demand system, *Applied Economics*, **34**, 1799–806. doi:10.1080/00036840210125008

Yen, S. T. and Lin, B.-H. (2006) A sample selection approach to censored demand systems, *American Journal of Agricultural Economics*, **88**, 742–9. doi:10.1111/j.1467-8276.2006.00892.x