Estimation of Price Elasticities from Cross-Sectional Data*

by

Chanjin Chung[a]
Diansheng Dong[a]
Todd M. Schmit[a]
Harry M, Kaiser[a]
Brian Gould[b]

[a]Department of Applied Economics and Management
Cornell University

and

[b]Department of Agricultural and Applied Economics
University of Wisconsin

First written: May 15, 2001
First revised: July 31, 2001

---

*A selected paper presented at the American Agricultural Economics Association Annual Meeting, Chicago, Illinois, August 5-8, 2001.

Estimation of Price Elasticities from Cross-Sectional Data

Abstract

This study develops an empirical framework that can be used to estimate quality-adjusted price elasticities from cross-sectional data, which are theoretically consistent and comparable to elasticities from time-series data. The new approach shows the importance of properly adjusting for quality variation in demand analysis.

JEL subject codes: Demand and Price Analysis

Estimation of Price Elasticities from Cross-Sectional Data

Estimating elasticities of consumer demand has been an important research issue in agricultural economics. These estimates are used to understand the behavior of consumers and to establish firm or industry-level marketing strategies and appropriate agricultural policies at the government level. Past studies have generally estimated these elasticities using inter-temporal price information from time-series data. Recently, however, there have been discussions in the literature on estimating demand elasticities from cross-sectional data as this type of data becomes increasingly available from either traditional survey or electronic scan. Important questions often posed in the literature include: can we estimate price elasticities from cross-sectional data?; if so, what is the theoretical background we should base on and what procedures should we establish to make these elasticities comparable to those from time-series data?

The objective of this study is to develop an empirical framework that can be used to estimate price elasticities in cross-sectional demand analysis, which are theoretically consistent and comparable to elasticities from time-series data. Traditionally, the demand analysis with cross-sectional data assumed that all households or individuals face the same prices. In early 1950s, Prais and Houthakker argued that the price variation could exist in the cross-sectional data due to various reasons such as region, price discrimination, services purchased with the commodity, seasonal effects, and quality effects. Since then, there has been a general consensus in the literature that of these factors, price variation from regional and seasonal differences is desirable for demand analysis (e.g., Cox and Wohlgenant; Friedman; and Deaton). In particular, Friedman suggests that constructing a demand curve from spatial data is essentially similar to that from time-series data when conditions of supply vary considerably while conditions of demand vary little, which is possible for products that have distinctive local markets with different supply conditions. Therefore, it is suggested that if one can recover systematic price variations due to regional and seasonal differences, it is possible to estimate price elasticities from cross-sectional data that are comparable to those estimated from time series data.

The task of recovering the usable price variation for the cross-sectional demand analysis raises at least three important issues. First, the price variation which is equivalent to the price variation used for the typical time-series demand analysis must be clearly separated from the rest of the price variations. In most cross-sectional data, it is common that expenditures and total physical quantities are reported for a certain aggregate commodity. Here, expenditures and quantities are not all from homogeneous products, but represent aggregates of more or less closely related substitutes. When prices of individual products change, consumers choose quality as well as quantity. In other words, consumers respond to a price change by altering the composition of their choice bundle. As a result, if one uses unit value--expenditure divided by quantity--in the place of market price for the demand analysis, this unit value series most likely represents both price and quality variations. Therefore, it is important to single the net price variations out from the unit value series to obtain appropriate price elasticities. Second, in addition to recovering quality-adjusted prices, it is also important to consider corresponding quality-adjusted quantities (Nelson). As discussed earlier, a change in price may lead to a change in the composition of purchase bundle, which ultimately has the quality effects on both price and quantity. For example, say Susan purchases 10 pounds of beef (5 pounds of sirloin and 5 pounds of ground beef) for her household every week. When sirloin prices increase sharply due to the shortage of the stock, Susan gives up purchasing sirloin and purchase 10 pounds of ground beef. In this case, Susan still purchases 10 pounds of beef, but the quality of this beef is quite different.

A simple aggregation of physical quantity does not account for this problem. Finally, a few econometric issues such as selectivity and simultaneity problems need to be addressed in order to obtain appropriate price elasticities from cross-sectional data. When researchers use cross-sectional data, in most cases expenditures (therefore, unit values) are not observed for non-consuming households or individuals. Ignoring information from the non-consuming households may lead to biased estimates. One can take care of this problem by using either Tobit or Heckman's two-step selection model. However, if one needs to estimate a two equation system (e.g., expenditure and unit value equations or quantity and unit value equations) to account for the endogeneity of the unit values, she can also encounter simultaneity problem because disturbance terms of two equations are most likely corrected. This problem becomes more apparent particularly when two equations are estimated from truncated data. This is because when expenditure (or quantity) is unobserved, unit value is also unobserved..[1] We address all of these three issues in this study.

The outline of this paper is as follows. The next section demonstrates the potential problem of ignoring quality effects from the cross-sectional demand analysis and develops the relationship between biased price elasticities and corresponding true elasticities. It is then shown that the biased elasticities can be corrected via appropriate analytical and econometric procedures. We first discuss previous attempts to estimate price elasticities from cross-sectional data and then propose a simple procedure that produces unbiased and theory consistent price elasticities. Unlike previous studies (e.g., Cox and Wohlgenant; Goldman and Grossman; Cowling and Raynor), our approach directly derive quality-adjusted price elasticities based on traditional consumer demand theory. The last section provides an illustration of our procedure using the 1996 Mexican food purchase survey. This data is chosen because this type of research is more needed and useful in developing countries than developed countries. This is because few developing countries have good quality time-series data that can be used for demand analysis while these countries often have periodic cross-section survey data available for researchers.

What Would Happen If Quality Effects Are Ignored?

Most cross-sectional data provide records of expenditures on the commodity (with a certain degree of aggregation) and quantity purchased. In the case of household food consumption survey data, households are typically asked about their food expenditures and quantities for the previous week. In case of scanner data, this type of data contain more detailed purchase information including price and quantities by size, brand, and various product quality characteristics. However, researchers often aggregate these the purchase information in terms of total expenditures and the sum of physical quantities for some aggregate commodities due to difficulties in modeling the consumer behavior for individual products.

In consequence, a common practice to obtain price information is that researchers divide the total expenditures spent to purchase products in the commodity group by total quantities and use these "unit values" as corresponding commodity prices. However, the unit value is not the price of the commodity that can be readily used for the demand analysis (Deaton; Cox and Wohlgenent; Prais and Houthakker). This is because the unit value reflects quality as well as price variation. Let's revisit Susan's shopping record on beef. If initial prices of sirloin and ground beef were $6.5 and $1.5 per pound and later prices increased to $10 and $2, respectively, corresponding unit values decreased from $4 to $2 due to the change in Susan's food basket. In this case, the change in unit value does not correctly reflect the change in the true price of beef. While prices of both sirloin and ground beef went up, the unit value went down. In fact, the price elasticity calculated using unit values would erroneously indicate no consumer response to the change in beef price. Unit values do not take into account changes in the composition of the aggregate commodity. Consumers choose not only quantities but also qualities of the commodity. If consumers do not change the proportions of each product in the commodity group in response to changes in price, income, and other exogenous demand shifters, the quality effects would be zero and in turn, the use of unit values can be justified. However, this is less likely to happen in reality.

Suppose that for each aggregate commodity group A, there are N elementary products and prices of these products vary proportionally. Also, let $\mathbf{P}$ and $\mathbf{P}^*$ be vectors of observed product prices and relative prices of each product in the commodity group A, respectively. Further, let P_A be the price of aggregate commodity. Then, the relationship between observed prices of individual products and the corresponding aggregate commodity prices can be denoted by (Deaton):

(1) $$P = P_A P^*$$

where $\mathbf{P} = (P_{1,} ......P_i,.....P_N)$ and $\mathbf{P}^* = (P^*_{1,} ......P^*_i,.....P^*_N)$. Let x be a vector of product quantity. Then, the Hicks Composite Commodity Theorem yields:

(2) $$Q_A = P^{*\prime} x$$

Equation (2) indicates that an appropriate measure of the aggregate commodity quantity Q_A can be formulated by the relative price weighted sum of physical quantities of elementary products.

Equipped with definitions of P_A and Q_A, now we can explore the relationship between unit value V_A with P_A and Q_A, and then eventually between V_A, P_A, and the quality measure L_A. Following Nelson, the quality measure of a commodity L_A is defined by the ratio of the Hicksian composite good quantity to the sum of physical quantity, i.e. $L_A = Q_A/q_A$, where $q_A = \quad x_i$. Therefore, the quality level of the commodity increases with the proportion of higher-priced elementary products. The unit value V_A is then written as:

(3)
$$V_A = \frac{E_A}{q_A} = \frac{P'x}{q_A} = \frac{P_A Q_A}{q_A} = P_A L_A$$

where E_A is the total expenditure spent on commodity A. Taking the natural logarithm of equation (3) produces:

(4)
$$\ln V_A = \ln P_A + \ln L_A$$

Equation (4) clearly shows that the unit value includes quality factor as well as price variations. As a result, unless effects of the quality component are zero, i.e., $\ln L_A / \ \ln P_A = 0$, the use of V_A in the place of P_A would lead to the biased demand analysis.

To examine the effect of price changes on quality, we extend Deaton's derivation for our purpose (p. 422, Deaton). Assuming weak separability on the consumption of commodity A, the vector of product demand x can be written as a function of total expenditures on commodity A and the vector of prices of each product *i*. Since the demand functions of these elementary products are homogeneous of degree zero in total expenditures and prices, the vector of product demand x is represented by:

$$x = f(E_A, P) = f\left(\frac{E_A}{P_A}, P^*\right)$$

Then, from equations (2) and (3), L_A is only a function of **x** at a given $\mathbf{P}^*$ and in turn, is a function of E_A/P_A. That is:

(5) $$L_A = L_A(\frac{E_A}{P_A})$$

Differentiating equation (5) with respect to P_A gives:

(6) $$\frac{\partial \ln L_A}{\partial \ln P_A} = -\frac{\partial \ln L_A}{\partial \ln(E_A / P_A)} + \frac{\partial \ln L_A}{\partial \ln(E_A / P_A)} \frac{\partial \ln E_A}{\partial \ln P_A}$$

Also, taking the natural logarithm of the relationship, $E_A = P_A Q_A$ and differentiating it with respect to $\ln P_A$ produces:

(7) $$\frac{\partial \ln E_A}{\partial \ln P_A} = 1 + \frac{\partial \ln L_A}{\partial \ln P_A} + \hat{\eta}_P$$

where $\hat{}_P$ = $\ln q_A$/ $\ln P_A$, the elasticity of total physical quantity with respect to aggregate commodity price P_A. Substituting equation (7) for the last term of equation (6) yields:

(8) $$\frac{\partial \ln L_A}{\partial \ln P_A} = \frac{\hat{\eta}_P \, \partial \ln L_A / \partial \ln(E_A / P_A)}{1 - \partial \ln L_A / \partial \ln(E_A / P_A)}$$

Given P_A (i.e., P_A/ $M = 0$), differentiating equation (5) with respect to the total income M and rearranging it for $\ln L_A$/ $\ln(E_A/P_A)$, we have:

(9) $$\frac{\partial \ln L_A}{\partial \ln(E_A / P_A)} = \frac{\mu}{\mu + \hat{\eta}_M}$$

where = $\ln L_A$/ $\ln M$, the quality elasticity with respect to the income M, and $\hat{}_M$ =

$\ln q_A / \ln M$, the elasticity of total physical quantity with respect to the income M. Substituting equation (9) for the term $\ln L_A / \ln E_A$ in equation (8),we now have:

$$\frac{\partial \ln L_A}{\partial \ln P_A} = \frac{\hat{\eta}_P \wp}{\hat{\eta}_M} \tag{10}$$

Then, by equations (4) and (10), we have:

$$\frac{\partial \ln V_A}{\partial \ln P_A} = 1 + \frac{\partial \ln L_A}{\partial \ln P_A} = 1 + \frac{\hat{\eta}_P \wp}{\hat{\eta}_M} \tag{11}$$

Equation (11) suggests that quality effects cannot be ignored unless either $\hat{}_P$ or is perfectly inelastic, which seems quite consistent with our intuition. Therefore, if one erroneously estimates price elasticities using unit values, one would obtain:

$$\frac{d \ln q_A}{d \ln V_A} = \frac{\partial \ln q_A \,/\, \partial \ln P_A}{\partial \ln V_A \,/\, \partial \ln P_A} = \frac{\hat{\eta}_P}{1 + \frac{\hat{\eta}_P \wp}{\hat{\eta}_M}} \tag{12}$$

Equation (12) shows that the use of unit values for the estimation of price elasticity will overstate the price elasticity in absolute value because the price elasticity is negative for normal goods, and the product of price elasticity $\hat{}_P$ and quality elasticity with respect to income is normally smaller in absolute value than the income elasticity $\hat{}_M$(Deaton).

However, note that the price elasticity employed in equation (12) is the price elasticity of the sum of physical quantity (q_A). As Nelson pointed out, the assumption of homogenous separability does not justify the use of a simple sum of physical quantities for the commodity demand. To address this problem, it has been recommended in the literature that researchers use proper quantity indices (Nelson; Deaton and Muellbauer). To this end, we use the Hicksian composite commodity index (Q_A) in this study. The use of q_A in the place of Q_A may also lead to the biased results.

Similar to equation (11), differentiating $\ln q_A$ with respect to $\ln Q_A$ provides the analysis for this problem as:

$$\frac{\partial \ln q_A}{\partial \ln Q_A} = \frac{\partial \ln(Q_A / L_A)}{\partial \ln Q_A} = 1 - \frac{\partial \ln L_A}{\partial \ln E_A} = 1 - \frac{\varphi}{\varphi + \hat{\eta}_M} \tag{13}$$

The last term of equation (13) follows from the assumption of fixed P_A. Hence, equation (13) indicates that unless = 0, the use of q_A is not justified for the estimation of the price elasticity. To examine the effects of using q_A in the place of Q_A , we first derive the relationship between price elasticities with q_A and Q_A, i.e., ^ $_P$ = $d\ln q_A/d\ln P_A$ and $_P$ = $d\ln Q_A/d\ln P_A$, as:

$$\hat{\eta}_p = \frac{d\ln q_A}{d\ln P_A} = \frac{\partial \ln Q_A}{\partial \ln P_A} - \frac{\partial \ln L_A}{\partial \ln P_A} = \eta_P - \frac{\hat{\eta}_P \varphi}{\hat{\eta}_M} \tag{14}$$

$$\rightarrow \hat{\eta}_p = \frac{\eta_P}{1 + \frac{\varphi}{\hat{\eta}_M}}$$

Then, rewriting equation (12) using the relationship derived in equation (14) gives a new analytical result as:

$$\frac{d\ln q_A}{d\ln V_A} = \frac{\hat{\eta}_P}{1 + \frac{\hat{\eta}_P \rho}{\hat{\eta}_M}} = \frac{\eta_P}{1 + (1 + \eta_P)\frac{\rho}{\hat{\eta}_M}} \tag{15}$$

Unlike Deaton's result described in equation (12), equation (15) indicates that the use of the unit value and the sum of physical quantities, V_A and q_A , in places of quality adjusted price and quantity, P_A and Q_A, does not necessarily overstate the true price elasticity $_P$. Assuming the typical negative price elasticity, the direction of "biasness" depends on the level of the true price elasticity $_P$. When $_P$ is unit elastic, there would be no bias. However, if, as expected, dlnq$_A$/ dlnqV$_A$ < 0, $_P$ < 0, > 0, $^\wedge{}_M$ > 0, and | | < |$^\wedge{}_M$|, an inelastic $_P$ would be understated while an elastic $_P$ would be overstated.

How to Correct for Quality Differences?

Analytical results from the previous section suggest that if one does not take into account quality effects in both price and quantity, a biased price elasticity is estimated. To avoid biased estimates of price elasticities, it is necessary to construct price series that reflect true price variation and corresponding quantity demanded for the demand analysis. We will first provide a brief review of previous approaches (Theil; Houthakker; Deaton; and Cox and Wohlgenant), then develop a simple procedure that can be used to estimate quality-adjusted price elasticities.

Theil and Houthakker provide theoretical frameworks that recognize quality effects in estimating demand functions. Both studies pose the utility maximization problem, where the level of consumer's utility is a function of both quantities and qualities. Theil assumes that quality factors of aggregate commodity are reflected in its composite price and quantity is defined as the sum of the physical quantities of elementary goods in the commodity group. A similar framework was presented in Houthakker, except that both quality and quantity factors are reflected in the composite price of the commodity. Armed with this conceptual background, both studies thoroughly analyzed the behavior of unit values, but their efforts fall short of estimating appropriate price elasticities.

Estimating quality adjusted price elasticities from cross-sectional data requires one to develop proper econometric procedures that account for quality effects included in price and quantity data due to the product aggregation. The most common approach to this end is to employ a hedonic pricing approach. The basic idea of this approach is that the quality component of commodity price can be completely described by a vector of quality characteristics. One way is to estimate implicit prices of different quality characteristics across commodities and then to net them out except for those from sources of true price variations that can be used for the cross-sectional demand analysis. What are the sources of these price variations? Among the sources of price variations identified by Prais and Houthakker, many studies in the literature support the proposition that price variations induced by regional and seasonal differences are most desirable for the estimation of demand functions from cross-sectional data (e.g., Cox and Wohlgenant; Deaton; Friedman). In this case, the proposition assumes that the relative price structure, particularly the supply side, is fixed within each geographical area (e.g., region or market) at a given time period (e.g., season, month, etc). Hence, individuals or households in the same geographical area are supposed to face the same market price and its variations, which are attributed to differences in supply conditions of different regions and seasons.

A hedonic price function, typically used in the literature, can be written for each commodity i as:

(16) $$V_i = \alpha_i + \sum_j \gamma_{ij} C_{ij}$$

where $_i$ can be interpreted as the quantity price and the second term represents the sum of values of quality per each unit of commodity i (see equation (4) for the justification of this interpretation). Here, the intercept term $_i$ can also be interpreted as the regional/seasonal mean price because if one can control for all the differences in quality characteristics (including regional and seasonal factors), $_i$ should be the same across regions and seasons. From equation (16), the deviations from regional/seasonal mean prices ($RDMP_i$) is written as:

(17) $$RDMP_i = V_i - \alpha_i = \sum_j \gamma_{ij} C_{ij} + e_i$$

Equation (17) indicates that the variation of RDMP can be explained by the variation of quality characteristics. Since e_i's are residuals from the regression equation, if one leaves regional and seasonal factors out while completely controlling for all other quality factors, e_i's reflect price variations solely induced by the regional/seasonal differences. Then, the quality adjusted prices P_{Ai} for commodity i can be generated as (Cox and Wohlgenant; Cowling and Raynor; Goldman and Grossman):

(18) $$P_{Ai} = \alpha_i + e_i$$

One of advantages of this type of hedonic pricing approach is that this approach does not need to identify the boundaries of regional markets to obtain the price variation due to the price difference in regional markets. However, a key limitation is that the price variation produced by this approach may not be solely from regional and seasonal differences. If equation (17) includes all of the major quality characteristics of commodity, except for regional/seasonal indicators, the residual term e_i's reflect price variations from these left-out factors. However, this is almost impossible because it is quite difficult to identify major quality factors that may affect variations of unit values and, more importantly, researchers do not have access to this type of data. Hence, researchers frequently rely on proxy variables such as household demographic characteristics. As a result, the residual term e_i in equations (17) and (18) may not represent price variations solely due to regional and seasonal differences, but still contain some of the quality effects. Alternative way of constructing hedonic pricing equation may be to directly estimate region- and season-based price variations by adapting equations (17) and (18) as:

(19) $$RDMP_i = V_i - \alpha_i = \sum_j \alpha_{ij} D_{ij} + \sum_k \alpha_{ik} S_{ik} + \sum_l \gamma_{il} C_{il} + e_i$$

(20) $$P_{Aijk} = \alpha_i + \alpha_{ij} + \alpha_{ik}$$

where D_{ij} and S_{ik} are regional and seasonal indicator variables and P_{Aijk} is quality adjusted price for commodity *i* in region *j* in season *k*. The new price series P_{Aijk} completely satisfies the theoretical argument in the literature (Deaton; Friedman). Since this new price series reflects price variations caused by different supply conditions of each region and season, everyone in the same region for a given time period faces the same quality adjusted price. A major limitation of this approach is that researchers need to identify correct market boundaries for regional markets that have different supply conditions and consequently different prices. Not only this task is difficult to achieve, but also many data sets simply do not contain enough information to construct sufficient level of price variations across regions (for example, the 1994 USDA food consumption survey include only four regions).

Our study develops a new procedure to overcome these limitations in previous approaches. We analytically derive the quality-adjusted price elasticity using relationships between unit value and true commodity price. In equations (4) and (11) of the previous section, we were able to decompose the variation of unit values into two parts: variation due to price and quality. These relationships are all conceptually based on the Hick's Composite Commodity Theorem, which will lead to quality adjusted price elasticities. By the definition of quality adjusted price elasticity with the total physical quantity q_A ($\hat{\eta}_P$) and equation (11), we have:

(21)
$$\hat{\eta}_P = \frac{d\ln q_A}{d\ln P_A} = \frac{\partial \ln q_A}{\partial \ln V_A} \cdot \frac{\partial \ln V_A}{\partial \ln P_A} = \hat{\eta}_V \cdot \left(1 + \frac{\hat{\eta}_P \mu}{\hat{\eta}_M}\right)$$

$$\rightarrow \hat{\eta}_P = \frac{\hat{\eta}_V}{1 - \frac{\mu}{\hat{\eta}_M} \hat{\eta}_V}$$

where $\hat{\eta}_V = d\ln q_A/d\ln V_A$. Equation (21) indicates that the quality adjusted price elasticity $\hat{\eta}_P$ can be derived from the unit value elasticity $\hat{\eta}_V$, the income elasticity $\hat{\eta}_M$, and the quality elasticity . Elasticities, $\hat{\eta}_V$ and $\hat{\eta}_M$, can be obtained from traditional quantity-dependent demand equation with unit value and income as explanatory variables, and the quality elasticity is typically estimated from unit value equation with an assumption of $d\ln P_A/d\ln M = 0$ (see

Houthakker and Prais; and Deaton for similar applications).

For the quality-adjusted price elasticity with the aggregate commodity quantity index Q_A, $_P$ = $dlnQ_A/dlnP_A$, we need to account for quality effects in aggregating quantities as well. The issue of quality effects in quantity aggregation has been mostly left out in the literature while most studies have sought to control for quality effects in the measurement of price variation. However, the quality adjustment in the measurement of quantity variation may well be important particularly when elementary products are quite heterogenous. Although the data for Q_A is not usually available from most data sets, this index can be easily recovered from the relationship, $E_A = Q_A P_A$. However, the estimation of demand equation, $Q_A = Q_A (P_A \mid Z)$, where Z represents a vector of demand shifting factors, is flawed due to the perfect correlation between Q_A and P_A . To avoid this problem, we propose to estimate the Engel equation, $E_A = E_A (P_A \mid Z)$, instead of a quantity-dependent demand function. Then, we can recover the price elasticity $_P$ from estimates of the Engel equation. To show this procedure, taking the natural logarithm of equation, $E_A = Q_A P_A$, and differentiating it with respect to lnP_A, gives:

(22)
$$\eta_P = \eta_{EG} - 1 = \frac{\partial \ln E_A}{\partial \ln V_A} \cdot \frac{\partial \ln V_A}{\partial \ln P_A} - 1 = \hat{\eta}_{EV} \cdot \left(1 + \hat{\eta}_P \frac{\mu^{\omega}}{\hat{\eta}_M}\right) - 1$$

$$= \hat{\eta}_{EV} \cdot \left[1 + \frac{\hat{\eta}_V}{1 - \frac{\mu^{\omega}}{\hat{\eta}_M}\hat{\eta}_V} \cdot \frac{\mu^{\omega}}{\hat{\eta}_M}\right] - 1$$

where $_{EG} = dlnE_A/dlnP_A$ and $_{EV} = dlnE_A/dlnV_A$. Equation (22) seems to indicate that in addition to the estimation of quantity (q_A) and unit value (V_A) equations, researchers also need to estimate Engel equation (E_A) to get $\hat{}_{EV}$. We suggest to estimate a two equation system with Engel and unit value equations. A conventional specification of Engel equation with unit value as one of explanatory variables results in $\hat{}_{EV}$ while the uinit value equation provides .

Elasticities, $\hat{}_{V}$ and $\hat{}_{M}$, are derived from the Engel equation. First, we take the natural logarithm of equation $E_A = q_A V_A$ and differentiate it with respect to $\ln V_A$, which leads to:

(23) $$\hat{\eta}_V = \hat{\eta}_{EV} - 1$$

Then, differentiating $E_A = q_A V_A$ with respect to M yields:

(24) $$\hat{\eta}_M = \hat{\eta}_{EM} - \varphi$$

where $\hat{}_{EM} = d\ln E_A/d\ln M$. Therefore, equations (21) - (24) suggest that the two equation system that includes Engel and unit value equations would directly estimate elasticities, $\hat{}_{EV}$, $\hat{}_{EM}$, and , and with derived elasticities, $\hat{}_{M}$ and $\hat{}_{V}$, conveniently produce both quality-adjusted price elasticities, $\hat{}_{P}$ and $_{P}$. These price elasticities are consistent with consumer theory and comparable to those estimated from time-series data because quality effects are explicitly eliminated through the derivation.

An Illustration

In this section, we apply the theoretical framework developed in the preceding sections to the 1996 Mexican food purchase data. The Mexican data is a nation-wide survey data on household expenditure and income and was collected by the Instituto Nacional de Estadistica, Geografia e Informatica in 1996. Surveyed households were asked to maintain diary for their daily expenses during seven consecutive days and the survey data were compiled with various demographic characteristics of each household. There were 12,580 households in the initial sample, but we selected only urban residents, reducing the sample to 6,394 households. The decision was made to avoid potential bias in estimates of price elasticities due to self-produced foods in rural areas. Table 1 provides the definition of exogenous variables used in our econometric model along with sample means and standard deviations. Our model hypothesizes that household decisions on the level of food expenditures and unit values are affected by household characteristics, the level of meal planners' education attainment, the age composition of household members, and region and location of residence. The third column of Table 1 identifies the variables used in unit value and expenditure equations. Table 2 presents purchase frequencies, means, and standard deviations of weekly food expenditures, quantities, and unit values. It appears that grain is the most frequently purchased while seafood is the least frequently purchased commodity. This indicates that the selectivity problem may affect estimates of seafood equation most significantly.

Since ignoring censored nature of the data and simultaneity issue between expenditure and unit value equations would produce biased coefficients, a two-equation system needs to be estimated. An empirical specification of the expenditure and unit value equation may be expressed as:

$$(25)\; E_{Ai} = \begin{cases} \alpha_0 + \alpha_1 \ln V_{Ai} + \alpha_2' Z_i + \varepsilon_{1i}, & if\, E_{Ai} > 0 \\ 0, & otherwise \end{cases}$$

$$(26)\; \ln V_{Ai} = \begin{cases} \beta_0 + \beta_1' X_i + + \varepsilon_{2i}, & if\, E_{Ai} > 0 \\ 0, & otherwise \end{cases}$$

where vectors Z_i and X_i represent household characteristics described in Table 1. Estimation of two censored equations simultaneously has been well documented in the literature (e.g., Wales and Woodland; Dong et al). Following Wales and Woodland, and Dong et al, we assume a joint normal distribution with mean vector of zero and variance-covariance matrix , where

$$\Omega = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}$$

Then, the combined likelihood function for both purchasing and non-purchasing households in the sample size N is:

$$(27)\ L(\alpha,\beta,\Omega \mid E_A, \ln V_A) = \prod_{i=1}^{M} n(\varepsilon_{1i}, \varepsilon_{2i}; 0, \Omega) \prod_{i=M+1}^{N} \Phi(-\kappa_i)$$

where M is the number of households which purchase the commodity; and n(.) and (.) denote the bivariate normal density function for ${}_{1i}$ and ${}_{2i}$, and the standard normal distribution, respectively, and ${}_{i}$ = [${}_{0}$ + ${}_{1}$(${}_{0}$ + ${}_{1}$'X_i) + ${}_{2}$'Z_i]/(${}_{1}^{2}$ + ${}_{1}^{2}$ ${}_{2}^{2}$ + 2 ${}_{1}$ ${}_{12}$)^{½} (see wales and Woodland for the derivation of ${}_{i}$). By taking logarithm of equation (27), we have:

$$(28)\ \ln L = \sum_{i=1}^{M} \{-0.5 \ln 2\pi - 0.5|\Omega| - 0.5\, \varepsilon_i' \Omega^{-1} \varepsilon_i\} + \sum_{i=M+1}^{N} \ln \Phi(-\kappa_i)$$

where ${}_{i}$ = [E_{Ai} - ${}_{0}$ - ${}_{1}$lnV_{Ai} - ${}_{2}$'Z_i, lnV_{Ai} - ${}_{0}$ - ${}_{1}$'X_i].

Maximum likelihood estimates and corresponding standard errors were obtained by maximizing equation (28) using GAUSS program and were reported in Table 3. For the unit value equation, higher income households appear to pay higher prices (maybe to purchase higher quality products) while households with large family size purchase lower price products. The age of meal planner does not seem to be a significant factor in determining unit values except vegetables and non-alcoholic beverage equations. Overall, more educated meal planners tend to look for higher-quality foods with exceptions from vegetables, beans, and non-alcoholic beverages. It is also shown that regional location of residence is important determinant of unit values. With regard to the expenditure equation, household income and size are significantly positive factor on expenditure. The attainment of meal planners' education appears to have negative impact on expenditures on beans and grain while it affects positively expenditures on meat. Again, regional location of residence significantly impacts on the level of food expenditures. High correlation coefficient ( ) is reported for seafood and beans. Recall that these two commodities show the lowest purchase frequencies in Table 2. This result supports that the issue of selectivity should be addressed along with the simultaneity problem.

Computing price elasticities from censored regression of equations (25) and (26) requires unconditional expectations of E_{Ai} and $\ln V_{Ai}$ for both purchasing and no-purchasing households. They are:

$$
\begin{aligned}
E(E_{Ai}) &= E(E_{Ai} \mid E_{Ai} > 0) \cdot prob(E_{Ai} > 0) + E(E_{Ai} \mid E_{Ai} = 0) \cdot prob(E_{Ai} = 0) \\
&= E(E_{Ai} \mid E_{Ai} > 0) \cdot \Phi(\kappa_i) \qquad (29) \\
&= \Phi(\kappa_i)[\alpha_0 + \alpha_{10}(\beta_0 + \beta_{10}' X_i) + \alpha_2' Z_i] + \phi(\kappa_i)\sqrt{\sigma_1^2 + \alpha_1^2 \sigma_{21}^2 + 2\alpha_1 \sigma_{21}}
\end{aligned}
$$

$$
(30) \quad \begin{aligned} E(\ln V_{Ai}) &= E(\ln V_{Ai} | E_{Ai} > 0) \cdot \Phi(\kappa_i) + E(\ln V_{Ai} | E_{Ai} = 0) \cdot (1 - \Phi(\kappa_i)) \\ &= \beta_0 + \beta_{10}' X_i \end{aligned}
$$

In the previous section, it was suggested that elasticities, $\hat{}_{EV}$, $\hat{}_{EM}$, and be estimated from regression equations. Elasticities, $\hat{}_{EM}$, and , can be easily estimated from equations (29) and (30), but $\hat{}_{EV}$ can not be estimated from equation (29) because the unit value is not found in the right-hand-side of the equation. To obtain $\hat{}_{EV}$, we derive expectations for E_{Ai} given unit value as[2]:

$$
(31) \quad E(E_{Ai} | \ln V_{Ai}) = \Phi(\lambda)[\alpha_0 + \alpha_{10} \ln V_{Ai} + \alpha_2' Z_i] + \phi(\lambda) \sigma_1 \sqrt{1 - \rho^2}
$$

where $_i$ = [$_0$ + $_1$lnV_{Ai} + $_2$'Z_i + $_{12}$(lnV_{Ai} - $_0$ - $_1$'X_i)/ $_2^2$]/ $_1$(1- ^{2})^{½} and = $_{12}$/ $_1$ $_2$. From equations (30) and (31), we can now obtain all the elasticities $\hat{}_{EV}$, $\hat{}_{EM}$, , $\hat{}_V$, and $\hat{}_M$ to compute price elasticities, $\hat{}_P$ and $_P$ and results are summarized in Table 4.

Most elasticities show the correct sign and are significant. The elasticities reflect the unconditional response to a change in each independent variable and approximate standard errors were derived using the Delta method from the estimated parameter variance-covariance matrix (Green, p297). Our primary concern is with the comparison of quality adjusted ($\hat{}_P$ and $_P$) versus unadjusted price elasticities ($\hat{}_V$). Results in Table 4 are consistent with our findings from the analytical analysis in the preceding sections. Unadjusted unit values tend to overstate the quality-adjusted price elasticity $\hat{}_P$. With regard to the adjusted price elasticities in both price and quantity ($\hat{}_P$), it appears that unadjusted unit value understate the inelastic price elasticity while it overstate the elastic price elasticity. Although the differences in magnitude may be small for some product (e.g., non-alcoholic beverage), the differences between unadjusted and adjusted elasticities are overall significant. Our results suggest that erroneous use of unit values in the place of quality adjusted prices may lead to biased policy implications.

Conclusions

This study develops a framework that can be used to estimate quality-adjusted price elasticities from cross-sectional data, which are theoretically consistent and comparable to elasticities from time-series data. Our new approach shows the importance of properly adjusting for quality variation in both prices and quantities. The new framework improves previous studies at least in three ways. First, we explicitly eliminate quality effects from unit values based on microeconomic theory. Second, we account for both sample selection and simultaneity problems in estimating demand equations. Finally, we also consider quality effects in aggregating quantities for the commodity group. Our analytical results clearly show that ignoring quality adjustment in either prices or quantities could produce the biased demand analysis. If the unit value and the sum of physical quantity are used for the estimation of price elasticity, the price elasticity may be overstated. However, when quality adjusted price and quantity are used, the direction of 'biasness" depends on the level of the price elasticity. If the price elasticity is unit elastic, there would be no bias. However, an inelastic price elasticity would be understated while an elastic price elasticity would be overstated. Our analytical framework was applied to the estimation of the Mexican food demand, and the empirical results were consistent with findings from analytical derivations.

Footnotes

1. One might also raise a simultaneity issue for demand equations across products. However, as pointed out by Cox and Wohlgenant, this type of approach to modeling would lead to the potential of intractable model solutions.

2. See Dong et al and Wales and Woodland for the derivation in detail.

References

Cowling, K. and A. J. Raynor. “”Price, Quality, and Market Share.” *J. Polit. Econ.* 78(1970): 1292-1309.

Cox, T. L. and M. K. Wohlgenant. “Prices and Quality Effects in Cross-Sectional Demand Analysis.” *Amer. J. Agr. Econ.* 68(November 1986): 908-19.

Deaton A. and J. Muellbauer. *Economics and Consumer Behavior*. New York: Cambridge University Press, 1993.

Deaton, A. “Quality, Quantity, and Spatial Variation of Price.” *Amer. Econ. Rev.* 78(June 1988): 418-503.

Friedman, M. *Price Theory*. New York: Aldine Publishing Co., 1976.

Goldman, F. and M. Grossman. “The Demand for Pediatric Care: A Hedonic Approach.” *J. Polit. Econ.* 86(1978): 259-80.

Green, W. H. *Econometric Analysis*. New York: Macmillan Publishing Company, 1993.

Houthakker, H. S. “Compensated Changes in Quantities and Qualities Consumed.” *Rev. Econ. Stud.* 19(1952): 155-63.

Nelson, J. A. “Quality Variation and Quantity Aggregation in Consumer Demand for Food.” *Amer. J. Agr. Econ.* 73(November 1991): 1204-12.

Prais, S. J. and H. J. Houthakker. *The Analysis of Family Budgets*. Cambridge: Cambridge University Press, 1955.

Suits, D. B. “Dummy Variables: Mechanics versus Interpretation.” *Rev. Econ. Statist.* 66(1984): 177-80.

Theil, H. “Qualities, Prices, and Budget Enquiries.” *Rev. Econ. Stud.* 19(1952): 129-147.

Table 1. Descriptive Statistics of Exogenous Variables Used in the Econometric Model.

Variable	Description	Equation	Mean	Standard Deviation
Household Characteristics				
HHINC	Monthly household income (1,000 pesos)	UV, E	9.51	9.21
HHSIZE	Number of household members (#)	UV, E	4.38	2.05
MPAGE	Age of meal planner (years)	UV, E	41.18	14.56
REFRIG	Household owns a refrigerator/freezer (0/1)	UV, E	0.80	----
EST1	Reside in major metropolitan area (0/1)	UV	0.42	----
EST2	Reside in area with population more than 100,000 (0/1)	UV	0.29	----
Meal Planner Education Attained				
PRIMED	Have completed grade school (0/1)	UV, E	0.22	----
SECED	Have completed secondary school (0/1)	UV, E	0.29	----
HIGHED	Have completed high school (0/1)	UV, E	0.29	----
COLLED	Have completed college education (0/1)	UV, E	0.10	----
Household Composition				
PERLT6	Proportion of members less than 6 years old (%)	E	0.12	0.16
PER6_15	Proportion of members between 6 and 15 years old (%)	E	0.19	0.20
PER16_24	Proportion of members between 16 and 24 years old (%)	E	0.18	0.22
PER45_65	Proportion of members between 45 and 65 years old (%)	E	0.15	0.24
PERGT65	Proportion of members older than 65 years old (%)	E	0.06	0.19
Region of Residence				
REGCEN	Aguascalientes, Hidalgo, Morelos, Puebla, Tlaxcala (0/1)	UV, E	0.12	----
REGNC	Durango, San Luis Potos, Queretaro, Zacatecas (0/1)	UV, E	0.07	----
REGNE	Coahuila, Chihuahua, Nuevo Leon, Tamaulipas (0/1)	UV, E	0.17	----
REGNO	Baj California, Baja California Sur, Sonora, Sinalo (0/1)	UV, E	0.10	----
REGOCC	Nayrit, Jalisco, Colima, Guanajuato, Michoacan (0/1)	UV, E	0.20	----
REGS	Guerrero, Oaxaca, Veracruz (0/1)	UV, E	0.06	----
REGSE	Yucatan, Tabasco, Quintana Roo, Chiapas, Campeche (0/1)	UV, E	0.17	----

Note: “UV” and “E” denote whether the variable is used in the unit value (UV) or expenditure (E) equation.

Table 2. Commodity Purchase Frequencies and Average Purchase Statistics.

Commodity	Percent	Mean Expenditure	Mean Quantity	Mean Unit Value
Meat*	89.24	57.21 (45.48)	2.80 (2.14)	20.98 (5.85)
Beef	67.91	34.54 (26.60)	1.40 (1.06)	25.63 (6.42)
Seafood	24.65	18.88 (22.26)	0.98 (1.03)	21.88 (11.50)
Vegetables	88.38	17.26 (13.82)	3.90 (3.14)	5.00 (2.95)
Grain	98.51	33.58 (23.84)	8.44 (6.16)	4.82 (3.33)
Beans	58.68	13.48 (9.96)	1.60 (1.13)	8.66 (2.59)
Nonalcoholic Beverages	70.46	19.40 (16.78)	13.90 (20.32)	5.88 (17.55)

Note: Percent is equal to the percent of households that purchased the respective commodities. Mean statistics represent means of purchase households only, standard deviations are given in the parentheses. All quantities are in kilograms, except the beverage category which is in liters.
* Meat includes beef, pork, and poultry.

Table 3. Estimated Coefficients from Econometric Model, by Commodity.

Equation	Meat		Beef		Seafood		Vegetable	
Variable	Unit Value	Expenditure	Unit Value	Expenditure	Unit Value	Expenditure	Unit Value	Expenditure
Intercept	2.940*	1.073	3.131*	1.152	2.652*	1.557	1.511*	0.190
ln(Unit Value)	----	-0.136	----	-0.269	----	-0.649		0.054
ln(HHINC)	0.088*	0.272*	0.070*	0.177*	0.148*	0.198*	0.064*	0.044*
HHSIZE^{-1}	0.210*	-0.519*	0.193*	-0.300*	0.210*	-0.031	0.163*	-0.213*
MPAGE	0.000	0.002*	0.000	0.000	0.001	0.000	-0.001*	0.000
REFRIG	0.032*	0.091*	0.025*	0.081*	0.047	0.060	-0.021	0.008
EST1	0.013	----	0.049*	----	0.065*	----	-0.044*	----
EST2	0.017	----	0.024*	----	-0.015	----	-0.005	----
PRIMED	0.025	0.037	0.016	0.034	0.048	0.001	-0.006	0.002
SECED	0.032*	0.078*	0.014	0.061*	0.078	0.083	-0.006	0.005
HIGHED	0.050*	0.082	0.030	0.064*	0.120*	0.092	-0.006	0.003
COLLED	0.043*	0.013	0.037	-0.000	0.152*	0.137	0.011	-0.018*
PERLT6	----	-0.069	----	-0.106*	----	0.005	----	-0.027
PER6_15	----	0.015	----	-0.028	----	0.027	----	-0.006
PER16_25	----	0.035	----	0.011	----	-0.050	----	0.001
PER45_65	----	0.061	----	0.050	----	0.071*	----	0.028*
PERGT65	----	0.066	----	0.064	----	0.005	----	0.017
REGCEN	-0.028	-0.183*	-0.010	-0.153*	0.015	-0.054	-0.054*	-0.034*
REGNC	-0.015	-0.290*	-0.078*	-0.154*	-0.191	-0.126	0.023	-0.071*
REGNE	-0.067*	-0.230*	-0.143*	-0.133	-0.092	-0.182*	0.076*	-0.131*
REGNO	-0.060*	-0.205*	-0.060*	-0.110*	-0.031	-0.001	0.193*	-0.081*
REGOCC	0.023	-0.191*	0.007	-0.079*	-0.060	-0.95	-0.092*	-0.082*
REGS	-0.038*	-0.116*	-0.019	-0.134*	-0.289*	-0.096	0.082	-0.065*
REGSE	-0.120*	0.037	-0.149*	-0.101*	-0.258*	-0.086	0.160*	-0.087*
$^{2}_{1}$, var. of Exp.	0.177*		0.118*		0.290*		0.019*	
$^{2}_{2}$, var. of UV.	0.072*		0.076*		0.232*		0.132*	
	0.229		0.391*		0.821*		-0.000	

* Indicates significance at the 5% level.

Table 3. Estimated Coefficients from Econometric Model, by Commodity (continued)

Equation	Beans		Grain		Nonalcoholic Beverages	
Variable	Unit Value	Expenditure	Unit Value	Expenditure	Unit Value	Expenditure
Intercept	2.096*	1.301	1.007*	0.357*	1.242*	-0.036
ln(Unit Value)	----	-0.565	----	0.037	----	0.078
ln(HHINC)	0.032*	0.006	0.140*	0.071*	0.015	0.086*
$HHSIZE^{-1}$	0.070*	-0.208*	0.532*	-0.353*	0.257*	-0.039
MPAGE	-0.000	0.001	0.000	0.002*	-0.005*	-0.000
REFRIG	0.032*	-0.027*	0.045*	-0.019*	-0.060	0.013
EST1	0.021*	----	0.083*	----	0.068*	----
EST2	0.012	----	-0.037*	----	0.100*	----
PRIMED	0.008	-0.017	0.065*	-0.008	0.009	0.006
SECED	-0.008	-0.050*	0.155*	-0.038*	-0.026	0.000
HIGHED	-0.014	-0.069*	0.244*	-0.055*	-0.081	-0.004
COLLED	0.024*	-0.065*	0.341*	-0.084*	-0.063	-0.031*
PERLT6	----	-0.017	----	0.009	----	0.001
PER6_15	----	0.052*	----	0.167*	----	0.012
PER16_24	----	0.019	----	0.057*	----	-0.002
PER45_65	----	0.010	----	0.001	----	-0.068*
PERGT65	----	-0.026	----	-0.007	----	-0.085*
REGCEN	0.005	-0.183*	0.125*	0.020	-0.212*	0.043*
REGNC	-0.041*	-0.290*	0.071*	-0.019*	-0.007	0.090*
REGNE	-0.078*	-0.230*	0.077*	-0.084*	-0.046	0.199*
REGNO	0.104*	-0.205*	-0.010	-0.084*	-0.161*	0.100*
REGOCC	0.169*	-0.191*	0.126*	0.018*	-0.519*	0.135*
REGS	-0.098*	-0.116*	0.070*	0.044*	-0.124	0.075*
REGSE	-0.181*	0.037	-0.015	-0.031*	-0.497*	0.194*
$^2_1$, var. of Exp.	0.046*		0.043*		0.045*	
$^2_2$, var. of UV. eqn	0.060*		0.244*		0.841*	
	0.776		0.254		-0.389	

* Indicates significance at the 5% level.

Table 4. Comparison of Price Elasticities.

	------------------------ Commodity ------------------------						
Elasticity	Meat	Beef	Seafood	Vegetables	Beans	Grain	Nonalcoholic Beverages
$^{\wedge}{}_{EV}$	-0.238	-0.871	-5.619	0.298	-5.215	0.105	0.402
	0.088*	0.070*	0.148*	0.064*	0.032*	0.140*	0.015
$^{\wedge}{}_{M}$	0.451*	0.489*	0.528*	0.261*	-0.099*	0.213*	0.445*
$^{\wedge}{}_{V}$	-1.238	-1.871	-6.619	-0.702	-6.215	-0.895	-0.598
$^{\wedge}{}_{P}$	-0.952	-1.423*	-1.853*	-0.571**	0.848	-0.327*	-0.585**
$_{P}$	-1.183	-1.662*	-2.573*	-0.757**	-0.289	-0.962*	-0.606**

Note: $^{\wedge}{}_{EV}$= $dlnE_A/dlnV_A$, = dln_A/dln_M , $^{\wedge}{}_{M}$= $dlnq_A/dln_M$, $^{\wedge}{}_{V}$= $dlnq_A/dlnV_A$,
$^{\wedge}{}_{P}$= $dlnq_A/dlnP_A$, $^{\wedge}{}_{P}$= $dlnQ_A/dlnP_A$

* and ** indicate significance at the 5% and 10% levels, respectively.