A Generalized Utility Model with Binding Non-Negativity Constraints: Demand for Beer

Xiaonan Liu
Graduate Student
Washington State University

Jeffrey T. LaFrance
Professor
Monash University

Hayley H. Chouinard
Associate Professor
Washington State University

Thomas L. Marsh
Professor
Washington State University

November 2013

Abstract: We propose a Generalized Quadratic Utility (GQU) model for an incomplete demand system with binding non-negativity constraints which is flexible in income and price effects. The model accounts for zero consumption using choke prices identified by Kuhn-Tucker conditions. The GQU demand system can overcome the analytical and computational difficulties of the Kuhn-Tucker approach. Applying a homothetic AIDS specification, we estimate the demand for ale and lager beers and find that household income, age, presence of child, and prices play significant roles in beer consumption. The own-price elasticities are greater than unity. The cross-price elasticities suggest substitutes between ales and lagers.

KEYWORDS: Ale, Choke price, Generalized Quadratic Utility model, Lager, Kuhn-Tucker approach, Zero consumption

JEL Classification: D11, D12, C31, L66

1. Introduction

Some suggest correctly accounting for people that choose not to consume when characterizing demand functions represents one of the most pressing issues in applied demand analysis (Deaton 1986). Traditional demand analysis assumes interior solutions to consumers' utility-maximization problems. However, most households purchase only a small subset of the available commodities and those goods only sometimes. Others recognize the importance of considering zero consumption, but apply techniques which result in theoretical inconsistency and/or econometric concerns. Failing to account for zero expenditure or directly applying the ordinary least squares method to censored data generates biased and inconsistent estimates (Tobin 1958, Olsen 1980, Greene 1981). These results lead to an inaccurate assessment of consumer behavior and questionable policy recommendations.

We propose and estimate a new Generalized Quadratic Utility (GQU) model with binding non-negativity constraints which represents a general, consistent, and easy-to-estimate incomplete demand system. The model accounts for zero consumption using the choke price which denotes the highest price a consumer would pay for a product and is identified by the Kuhn-Tucker condition. We derive this demand system by extending a linear demand model to include the possibility that each good has a non-zero income elasticity and a non-linear price response. Applying the GQU model to beer data, we investigate consumers' preferences for ales and lagers.

Numerous studies use a two-step censored model to allow for zero consumption (Heckman 1979, Cragg 1971, Shonkwiler and Yen 1999). This relatively simple estimation method can accommodate a complicated demand system with a large number of goods, but often fails to accurately illustrate consumers' preferences. The two-step estimation strategy derives a system of demand equations without considering non-negativity restrictions directly, and it enforces these restrictions by employing an extension of Tobin's (1958) limited dependent variable model. The two-step approach does not account for the critical role of choke prices in a demand system. Arndt (1999) demonstrates the incompatibility of the two-step procedure with economic theory and suggests the potential for substantial bias.

Wales and Woodland (1983) and Lee and Pitt (1986) construct a unified and internally consistent framework of demand characterizing the occurrence of corner solutions using Kuhn-Tucker conditions. However, the severe analytical and computational difficulties of this approach hinder a wide application of the model. The calculation complexity occurs due to multiple integrals and cumulative joint distribution function evaluation which limits the applications to a small number of goods (Phaneuf 2000, Kim et al. 2002). Additionally, regularity conditions from demand theory (e.g. homogeneity and Slutsky symmetry and negative definiteness of the demand function) in Kuhn-Tucker models are extremely difficult to derive in general, and may result in conflicting and overly restrictive conditions (Van Soest et al. 1993). Failing to account for the regularity conditions in estimation generates inconsistent estimates. Previous studies only consider regularity conditions in simple models with limited scope (Ransom 1987, Lee and Pitt 1987, Van Soest and Kooreman 1990).

We develop a generalized and theoretically consistent demand system allowing for zero consumption using choke prices in which each demand equation consists of a linear function plus a ratio of univariate integrals. We avoid solving high-dimensional integration and obtain closed-form solutions using change of variables, conditioning steps, and various approximations. This estimation strategy applies to complex demand systems with a large number of products.

We can impose parameter restrictions on the demand functions during estimation to guarantee coherency with economic theory. We specify the well-defined reduced forms of demand equations for a generalized model while keeping the regularity conditions satisfied. Our method uses parameter constraints corresponding to regularity conditions and choke prices accounting for zero consumption to obtain accurate estimates.

Different from previous work on demand systems using Kuhn-Tucker conditions, we derive a generalized structural form of income elasticity which could be constant or non-constant in a "quasi-linear" demand system. The previous Kuhn-Tucker models normalize prices by income, which makes it difficult to distinguish the effects of price and income on demand. No substitution exists between any of the goods unless the income effects equal zero in a system of linear demand functions (LaFrance 1985). However, we would expect at least some income effect on demand.

Applying the GQU model to beer data, we estimate the demand of ales and lagers and examine what demographic characteristics shift consumers' demand between these two goods with a homothetic Almost Ideal Demand System (AIDS) specification. The demand for ales and lagers provides an appropriate example for the GQU model since a large number of consumers do not purchase beer of any kind and often do not purchase ales (Protz 1995, Choi and Stack 2005, Connelly 2013).

According to Tremblay and Tremblay (2005), U.S. consumers drink about 22 gallons of beer per capita on average. Ales yield a beer with higher alcohol content and tend to be dark, and lager is noted for its mild flavor and light color. Lager beer commands a market share of nearly 83%.[1] Among all categories of beer, light beer is the most popular with a 44% market share.

A number of studies examine consumer demand for beer. Gallet (2007) uses a meta-analysis to find that beer elasticities tend to be more inelastic compared to other alcoholic beverages. Previous research shows that demographic factors and prices of beer and its substitutes play major roles in determining beer demand (Lee and Tremblay 1992, Ornstein and Hassens 1985). Bray et al. (2009) conclude that individuals prefer to buy a higher-volume package of the same brand of beer than to switch brands. However, no prior studies have investigated the consumers' preferences for ales and lagers, and none has considered the issue of zero consumption.

The GQU model of demand has considerable advantages over the existing strategies allowing for binding non-negativity constraints. The model exhibits coherency in theory, flexibility in functional forms, and dramatic simplification of empirical implementation. Our methodology has the ability to overcome both the analytical and computational difficulties which have limited the application of Kuhn-Tucker models for decades. We propose a generalized structural form of income elasticity which could be non-constant in an additive and quasi-linear demand function. The analysis of demand for ales and lagers demonstrates the usefulness of the proposed GQU model.

The remainder of this paper is organized as follows. In Section 2, we specify the GQU model of demand and the structural form of income elasticity. The econometric estimation with binding non-negativity constrains is presented in Section 3. Section 4

[1] We aggregate the market shares of ice beer (3.5%), light beer (44.0%), malt liquor (2.7%), popular-priced beer (10.4%), premium beer (20.6%), and super-premium beer (1.8%) into lagers.

contains a homothetic AIDS specification of the GQU demand model for ale and lager demand estimation. We discuss allowing closed-form solutions in Section 5 and Section 6 provides the data description. Results and conclusions are given in Section 7 and Section 8.

2. Specification of the GQU Demand Model

We derive a generalized transformed demand function by extending a linear demand function to include the possibility of a non-zero income elasticity and a non-linear price response. We begin with a linear demand model,

(1) $\boldsymbol{q} = \boldsymbol{\alpha} + \boldsymbol{A}\boldsymbol{z} + \boldsymbol{B}\boldsymbol{p} + \boldsymbol{\varepsilon}$,

where **q** represents an N-vector of quantities demanded for a subset of market goods that are of empirical interest, **z** is a K-vector of demographic variables and other preference shifters, **p** represents the N-vector of market prices for the goods **q**, **ε** denotes an N-vector of random preference parameters known to the consumer but unknown and unobservable by the econometrician, **α** represents an N-vector of parameters, **A** is an N×K matrix of parameters, and **B** denotes an N×N symmetric and negative definite matrix of parameters.

To develop the possibility of a non-zero income elasticity and a non-linear price response, we need to include income and allow for price to enter non-linearly in Equation (1) in a theoretically consistent way. We define a strictly increasing function in total expenditure $f(m, \boldsymbol{z})$, where $m = e(\boldsymbol{p}, \tilde{\boldsymbol{p}}, \boldsymbol{z}, u)$ represents the expenditure function, $\tilde{\boldsymbol{p}}$ denotes a vector of market prices of other goods, and u is a utility index.[2] We define

[2] The expenditure function $e(\boldsymbol{p}, \tilde{\boldsymbol{p}}, \boldsymbol{z}, u)$ is smooth in m and **z** (partial derivatives of all orders exist and are continuous).

$x_i = g_i(p_i)$ as a strictly increasing and smooth function of the market price of the i^{th} good, and let its inverse be denoted by $p_i(x_i)$. Let $\boldsymbol{p}(\boldsymbol{x}) = \boldsymbol{g}^{-1}(\boldsymbol{x})$ equal the vector-valued inverse of the vector-valued function $\boldsymbol{x} = \boldsymbol{g}(\boldsymbol{p})$. For degree zero homogeneity of the demands for **q**, we assume that all prices and total expenditure are normalized by a known, positive, linearly homogeneous, increasing, weakly concave, and smooth price index, $\pi(\tilde{\boldsymbol{p}})$. We abuse notation slightly and absorb this price deflator into the definitions of $(\boldsymbol{p}, \tilde{\boldsymbol{p}}, m)$.

We define the transformed expenditure function as $f(e(\boldsymbol{p}(\boldsymbol{x}), \tilde{\boldsymbol{p}}, \boldsymbol{z}, u), \boldsymbol{z})$ and apply the composite function theorem,

$$(2)\ \frac{\partial}{\partial \boldsymbol{x}} \partial f(e(\boldsymbol{p}(\boldsymbol{x}), \tilde{\boldsymbol{p}}, \boldsymbol{z}, u), \boldsymbol{z}) = \frac{\partial f(e(\boldsymbol{p}(\boldsymbol{x}), \tilde{\boldsymbol{p}}, \boldsymbol{z}, u), \boldsymbol{z})}{\partial m} \times \boldsymbol{\Delta}\left(p_i'(x_i)\right) \times \boldsymbol{q}^c(\boldsymbol{p}(\boldsymbol{x}), \tilde{\boldsymbol{p}}, \boldsymbol{z}, u),$$

where $\boldsymbol{\Delta}\left(p_i'(x_i)\right)$ is a diagonal matrix with $p_i'(x_i)$ for the i^{th} main diagonal element. By Hotelling's Lemma, $\boldsymbol{q}^c(\boldsymbol{p}(\boldsymbol{x}), \tilde{\boldsymbol{p}}, \boldsymbol{z}, u)$ represents the vector of Hicksian compensated demands for **q** with prices **p** written as functions of transformed prices **x**.

We substitute m for the expenditure function and rewrite this incomplete transformed demand system in terms of the Marshallian ordinary demands,

$$(3)\ y(\boldsymbol{p}, m, \boldsymbol{z}) = \frac{\partial f(m, \boldsymbol{z})}{\partial m} \times \boldsymbol{\Delta}\left(p_i'\left(g_i(p_i)\right)\right) \times \boldsymbol{h}(\boldsymbol{p}, \tilde{\boldsymbol{p}}, \boldsymbol{z}, m) = \boldsymbol{\alpha} + \boldsymbol{A}\boldsymbol{z} + \boldsymbol{B}\boldsymbol{g}(\boldsymbol{p}) + \boldsymbol{\varepsilon},$$

where $y(\boldsymbol{p}, m, \boldsymbol{z})$ represents an N-vector of transformed demand and $\boldsymbol{h}(\boldsymbol{p}, \tilde{\boldsymbol{p}}, \boldsymbol{z}, m)$ the N-vector of ordinary Marshallian quantities demanded.

Equation (3) constitutes the structural model, prior to accounting for binding non-negativity constraints on quantities demanded. It has the same form as a quadratic utility model's structural demand equation, but with considerably more generality and flexibility in terms of the income and price effects for the demands of interest. The goods **q** can have any income elasticities, which may or may not be constant. The demand functions may be non-linear in **p** and the own- and cross-price effects can have a variety of shapes. For this class of models, the quantity demanded for q_i equals zero if and only if its market price p_i lies on or above the conditional choke price, $\hat{p}_i$.

We derive the indirect utility function associated with this demand function to verify consistency with economic theory in Appendix 1 which is expressed as,

$$
\begin{aligned}
v(\boldsymbol{p}, \tilde{\boldsymbol{p}}, \boldsymbol{z}, m) &\equiv \upsilon\left(\varphi(\boldsymbol{p}, \tilde{\boldsymbol{p}}, \boldsymbol{z}, m), \tilde{\boldsymbol{p}}, \boldsymbol{z}\right) \\
&= \upsilon\left[f(m, \boldsymbol{z}) - (\boldsymbol{\alpha} + \boldsymbol{Az} + \boldsymbol{\varepsilon})^{\top} \boldsymbol{g}(\boldsymbol{p}) - \tfrac{1}{2} \boldsymbol{g}(\boldsymbol{p})^{\top} \boldsymbol{Bg}(\boldsymbol{p}), \tilde{\boldsymbol{p}}, \boldsymbol{z} \right].
\end{aligned} \tag{4}
$$

Equations (3) - (4) provide a complete description of the GQU demand system.

We solve for the income elasticity using the Marshallian demand function derived from Equation (3),

$$
\boldsymbol{h}(\boldsymbol{p}, \tilde{\boldsymbol{p}}, \boldsymbol{z}, m) = \left(\frac{\Delta\left(g_i'(p_i)\right)}{\partial f(m, \boldsymbol{z}) / \partial m} \right) \left(\boldsymbol{\alpha} + \boldsymbol{Az} + \boldsymbol{Bg}(\boldsymbol{p}) + \boldsymbol{\varepsilon} \right). \tag{5}
$$

The income elasticities are equal for the homothetically separable subset of demands for ***q*** (Dixit and Weller 1979, LaFrance and Hanemann 1989),

$$
\varepsilon_{q_i}^{m} = \frac{\partial h_i(\boldsymbol{p}, \tilde{\boldsymbol{p}}, \boldsymbol{z}, m)}{\partial m} \frac{m}{h_i(\boldsymbol{p}, \tilde{\boldsymbol{p}}, \boldsymbol{z}, m)} = -m \frac{\partial^2 f(m, \boldsymbol{z}) / \partial m^2}{\partial f(m, \boldsymbol{z}) / \partial m}, \ \forall\ i = 1, \cdots, n. \tag{6}
$$

If $\partial^2 f(m, \boldsymbol{z}) / \partial m^2 \neq 0$ and $\partial^2 f(m, \boldsymbol{z}) / \partial m \partial \boldsymbol{z} \neq \boldsymbol{0}$, then the income elasticity varies with $(m, \boldsymbol{z})$ reflecting homothetic separability of a subset of consumers' demands, which is more general than the constant income elasticity in a homothetic demand system.

If we consider the Box-Cox specifications of f and g_i , the income elasticities could vary,

$$
(7)\quad
\begin{aligned}
&f(m,\mathbf{z}) = \left(\left[a(\mathbf{z})m + b(\mathbf{z})\right]^{\kappa} - 1\right)\Big/\kappa,\ \lim_{\kappa\to 0} f(m,\mathbf{z}) = \ln\left[a(\mathbf{z})m + b(\mathbf{z})\right],\ \text{for } \kappa \in [0,1],\\
&g_i(p_i) = (p_i^{\lambda} - 1)/\lambda,\ \lim_{\lambda\to 0} g_i(p_i) = \ln p_i,\ \text{for } \lambda \in [0,1], \forall\ i = 1,\cdots,n.
\end{aligned}
$$

This model includes a linear incomplete demand system if $\kappa = \lambda = 1,\ \ a(\mathbf{z}) \equiv 1,$ and $b(\mathbf{z}) \equiv 0,$ a homothetic incomplete AIDS (and LA-AIDS, LaFrance 2004) when $\kappa = \lambda = 0,\quad a(\mathbf{z}) \equiv 1,$ and $b(\mathbf{z}) \equiv 0,$ and a large class of homothetic Price Independent Generalized Linearity (PIGL) and Price Independent Generalized Log Linearity (PIGLOG) models for other values of $\kappa,\ \lambda,\ \ a(\mathbf{z}),$ and/or $b(\mathbf{z}).$

We derive the common income elasticity of demand for the homothetically separable goods **q** given as

$$
(8)\quad \varepsilon_{q_i}^{m} = \frac{(1-\kappa)a(\mathbf{z})m}{\left[a(\mathbf{z})m + b(\mathbf{z})\right]},\ i = 1,\cdots,n.
$$
[3]

3. Econometric Estimation with Binding Non-Negativity Constraints

We account for the zero quantities demanded by replacing the observed market prices for the non-purchased goods with the conditional choke prices. The choke price presents the lowest price at which quantity demanded of good equals zero. Consumers themselves know their own conditional choke prices but econometricians cannot observe them. We replace the market price of a non-consumed good with the choke price because

[3] We derive the income elasticity using the following equations,
$\frac{\partial f(m,\mathbf{z})}{\partial m} = a(\mathbf{z})\left[a(\mathbf{z})m + b(\mathbf{z})\right]^{\kappa-1}$ and $\frac{\partial^2 f(m,\mathbf{z})}{\partial m^2} = (\kappa-1)a(\mathbf{z})^2\left[a(\mathbf{z})m + b(\mathbf{z})\right]^{\kappa-2}.$

it reflects the true preference information of a consumer. This inequality follows directly from the Kuhn-Tucker first-order conditions with respect to quantities demanded since $\partial u(q_{1h},\cdots,q_{i-1,h},0,q_{i+1,h}\cdots,q_{Nh},y_{0h})/\partial q_i \le \lambda p_{ih}$, where y_{0h} represents the expenditure on all other goods, $y_{0h} = m_h - \sum_{i=1}^{N} p_{ih}q_{ih}$, and h indexes households. The conditional choke prices are defined such that $\partial u(q_{1h},\cdots,q_{i-1,h},0,q_{i+1,h}\cdots,q_{Nh},y_{0h})/\partial q_i = \lambda \hat{p}_{ih}$, and $\lambda \ge 0$. In the more general model, this property applies to transformed prices, $x_{ih} = g_i(p_{ih})$, without change beyond the coordinate system for transformed prices since $g_i'(p_i) > 0$ and $\partial f/\partial m > 0$. That is, $\hat{x}_{ih} = g_i(\hat{p}_{ih}) \le x_{ih} = g_i(p_{ih})$ if and only if $\hat{p}_{ih} \le p_{ih}$, for all $i = 1,\cdots,N$ and $h = 1,\cdots,H$, where we define $\hat{x}_{ih}$ by

(9) $\partial u(q_{1h},\cdots,q_{i-1,h},0,q_{i+1,h}\cdots,q_{Nh},y_{0h})/\partial q_i = \lambda p_i(\hat{x}_{ih})$.

With N goods of interest, there are 2^N possible consumption regimes. We write the demand system for each of these possible regimes generically and in a compact way to derive the conditional means of transformed demand using inequality constraints of the error terms. From the theoretical model's properties, we can use the transformed demand functions to derive the inequality constraints associated with the error terms because the non-negative and negative conditions of demand and transformed demand are equivalent. For an arbitrary purchase regime of positive and zero consumption levels, we partition the vector of transformed demand functions $\boldsymbol{y}_h$ into two exhaustive sub-vectors, $\boldsymbol{y}_{(1)h} > \mathbf{0}_{(1)h}$ and $\boldsymbol{y}_{(2)h} = \mathbf{0}_{(2)h}$, and partition **A**, **B**, and **ε** conformably, then we obtain the structural models of demand as,

(10) $\boldsymbol{y}_{(1)h} = \boldsymbol{\alpha}_{(1)} + \boldsymbol{A}_{(1)}\boldsymbol{z}_h + \boldsymbol{B}_{(1,1)}\boldsymbol{x}_{(1)h} + \boldsymbol{B}_{(1,2)}\hat{\boldsymbol{x}}_{(2)h} + \boldsymbol{\varepsilon}_{(1)h} > \mathbf{0}_{(1)h}$,

(11) $\mathbf{0}_{(2)h} \geq \boldsymbol{\alpha}_{(2)} + \boldsymbol{A}_{(2)}\boldsymbol{z}_h + \boldsymbol{B}'_{(1,2)}\boldsymbol{x}_{(1)h} + \boldsymbol{B}_{(2,2)}\hat{\boldsymbol{x}}_{(2)h} + \boldsymbol{\varepsilon}_{(2)h}$.

We solve for the conditional choke prices by rearranging Equation (11) to yield,

(12) $\hat{\boldsymbol{x}}_{(2)h} = -\boldsymbol{B}_{(2,2)}^{-1}\left(\boldsymbol{\alpha}_{(2)} + \boldsymbol{A}_{(2)}\boldsymbol{z}_h + \boldsymbol{B}'_{(1,2)}\boldsymbol{x}_{(1)h} + \boldsymbol{\varepsilon}_{(2)h}\right)$.

Choke prices depend on an intercept term, demographic variables, market prices of other goods, and individual random preferences. Intuitively, consumers with different demographic characteristics, random preferences, or face different market prices of other goods have different choke prices.

We obtain the empirical model of the goods with positive quantities demanded by plugging choke prices in Equation (12) into Equation (10),

$$\begin{aligned}(13)\quad \boldsymbol{y}_{(1)h} = &\left[\boldsymbol{\alpha}_{(1)} - \boldsymbol{B}_{(1,2)}\boldsymbol{B}_{(2,2)}^{-1}\boldsymbol{\alpha}_{(2)}\right] + \left[\boldsymbol{A}_{(1)} - \boldsymbol{B}_{(1,2)}\boldsymbol{B}_{(2,2)}^{-1}\boldsymbol{A}_{(2)}\right]\boldsymbol{z}_h + \\ &\left[\boldsymbol{B}_{(1,1)} - \boldsymbol{B}_{(1,2)}\boldsymbol{B}_{(2,2)}^{-1}\boldsymbol{B}'_{(1,2)}\right]\boldsymbol{x}_{(1)h} + \left[\boldsymbol{\varepsilon}_{(1)h} - \boldsymbol{B}_{(1,2)}\boldsymbol{B}_{(2,2)}^{-1}\boldsymbol{\varepsilon}_{(2)h}\right] > 0.\end{aligned}$$

We solve for the error terms from Equations (11) and (13) to consider the two sets of inequality conditions,

$$\begin{aligned}(14)\quad \boldsymbol{y}_{(1)h} > \mathbf{0}_{(1)h} \Leftrightarrow \hat{\boldsymbol{\varepsilon}}_{(1)h} \equiv \boldsymbol{\varepsilon}_{(1)h} - \boldsymbol{B}_{(1,2)}\boldsymbol{B}_{(2,2)}^{-1}\boldsymbol{\varepsilon}_{(2)h} > -\left[\boldsymbol{\alpha}_{(1)} - \boldsymbol{B}_{(1,2)}\boldsymbol{B}_{(2,2)}^{-1}\boldsymbol{\alpha}_{(2)}\right] \\ -\left[\boldsymbol{A}_{(1)} - \boldsymbol{B}_{(1,2)}\boldsymbol{B}_{(2,2)}^{-1}\boldsymbol{A}_{(2)}\right]\boldsymbol{z}_h - \left[\boldsymbol{B}_{(1,1)} - \boldsymbol{B}_{(1,2)}\boldsymbol{B}_{(2,2)}^{-1}\boldsymbol{B}'_{(1,2)}\right]\boldsymbol{x}_{(1)h},\end{aligned}$$

(15) $\boldsymbol{y}_{(2)h} = \mathbf{0}_{(2)h} \Leftrightarrow \boldsymbol{\varepsilon}_{(2)h} \leq -\left(\boldsymbol{\alpha}_{(2)} + \boldsymbol{A}_{(2)}\boldsymbol{z}_h + \boldsymbol{B}'_{(1,2)}\boldsymbol{x}_{(1)h} + \boldsymbol{B}_{(2,2)}\boldsymbol{x}_{(2)h}\right)$.

We use market prices of the goods with zero consumption in Equation (15) since the quantities demanded are less or equal to zero under market prices.

To derive the regime-specific joint density, cumulative distribution, and likelihood functions for the stochastic components of these two sets of inequalities, we need to make a change of variables from $\left(\boldsymbol{\varepsilon}_{(1)h}, \boldsymbol{\varepsilon}_{(2)h}\right)$ to $\left(\hat{\boldsymbol{\varepsilon}}_{(1)h}, \boldsymbol{\varepsilon}_{(2)h}\right)$. We accomplish this by applying a non-singular linear system of N equations in N variables,

$\boldsymbol{\varepsilon}_{(1)h} \equiv \hat{\boldsymbol{\varepsilon}}_{(1)h} + \boldsymbol{B}_{(1,2)}\boldsymbol{B}_{(2,2)}^{-1}\boldsymbol{\varepsilon}_{(2)h}$ and $\boldsymbol{\varepsilon}_{(2)h} = \boldsymbol{\varepsilon}_{(2)h}$, so that the determinant of the Jacobian matrix equals 1 and

(16) $f_{\hat{E}_{(1)},E_{(2)}}(\hat{\boldsymbol{\varepsilon}}_{(1)h}, \boldsymbol{\varepsilon}_{(2)h}) = f_{E_{(1)},E_{(2)}}\left(\hat{\boldsymbol{\varepsilon}}_{(1)h} + \boldsymbol{B}_{(1,2)}\boldsymbol{B}_{(2,2)}^{-1}\boldsymbol{\varepsilon}_{(2)h}, \boldsymbol{\varepsilon}_{(2)h}\right).$

To avoid multiple integrals, we factor the joint probability density function (pdf) for $(\hat{\boldsymbol{\varepsilon}}_{(1)h}, \boldsymbol{\varepsilon}_{(2)h})$ into the product of the marginal pdf for $\hat{\boldsymbol{\varepsilon}}_{(1)h}$ and the conditional pdf for $\boldsymbol{\varepsilon}_{(2)h}$ given $\hat{\boldsymbol{\varepsilon}}_{(1)h}$,

(17) $f_{\hat{E}_{(1)},E_{(2)}}(\hat{\boldsymbol{\varepsilon}}_{(1)h}, \boldsymbol{\varepsilon}_{(2)h}) = f_{\hat{E}_{(1)}}(\hat{\boldsymbol{\varepsilon}}_{(1)h}) \times f_{E_{(2)}|\hat{E}_{(1)}}(\boldsymbol{\varepsilon}_{(2)h} \mid \hat{\boldsymbol{\varepsilon}}_{(1)h}).$

We assume that $\boldsymbol{\varepsilon}_h$ are independently and identically distributed as multivariate normal, $\boldsymbol{\varepsilon}_h \sim N(\mathbf{0}, \boldsymbol{\Sigma})$, in order to use approximations of pdf and cumulative density function to simplify the integral for closed-form solutions. Partitioning $\boldsymbol{\varepsilon}_h$ commensurably with the regime $(\boldsymbol{y}_{(1)h}, \boldsymbol{y}_{(2)h})$, we have

(18)
$$\begin{aligned} f_{E_{(1)},E_{(2)}}(\boldsymbol{\varepsilon}_{(1)h}, \boldsymbol{\varepsilon}_{(2)h}) = (2\pi)^{-N/2} &\begin{vmatrix} \boldsymbol{\Sigma}_{(1,1)} & \boldsymbol{\Sigma}_{(1,2)} \\ \boldsymbol{\Sigma}'_{(1,2)} & \boldsymbol{\Sigma}_{(2,2)} \end{vmatrix}^{-1/2} \\ &\times \exp\left\{ -\tfrac{1}{2} \begin{bmatrix} \boldsymbol{\varepsilon}_{(1)h} \\ \boldsymbol{\varepsilon}_{(2)h} \end{bmatrix}' \begin{bmatrix} \boldsymbol{\Sigma}_{(1,1)} & \boldsymbol{\Sigma}_{(1,2)} \\ \boldsymbol{\Sigma}'_{(1,2)} & \boldsymbol{\Sigma}_{(2,2)} \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{\varepsilon}_{(1)h} \\ \boldsymbol{\varepsilon}_{(2)h} \end{bmatrix}' \right\}. \end{aligned}$$

We derive the means and variance-covariance matrices for $\hat{\boldsymbol{\varepsilon}}_{(1)h}$ and $\boldsymbol{\varepsilon}_{(2)h} \mid \hat{\boldsymbol{\varepsilon}}_{(1)h}$ to obtain their pdfs. Because $\hat{\boldsymbol{\varepsilon}}_{(1)h}$ represents a linear transformation of $(\boldsymbol{\varepsilon}_{(1)h}, \boldsymbol{\varepsilon}_{(2)h})$, this is a multivariate normal pdf. The unconditional mean vector for $\hat{\boldsymbol{\varepsilon}}_{(1)h}$ is $\mathbf{0}_{(1)}$. The unconditional variance-covariance matrix for $\hat{\boldsymbol{\varepsilon}}_{(1)h}$ is

$$(19)\quad \begin{aligned}\hat{\boldsymbol{\Sigma}}_{(1,1)} &= E\left(\hat{\boldsymbol{\varepsilon}}'_{(1)h}\hat{\boldsymbol{\varepsilon}}'_{(1)h}\right)\\ &= \boldsymbol{\Sigma}_{(1,1)} - \boldsymbol{B}_{(1,2)}\boldsymbol{B}^{-1}_{(2,2)}\boldsymbol{\Sigma}_{(2,1)} - \boldsymbol{\Sigma}_{(1,2)}\boldsymbol{B}^{-1}_{(2,2)}\boldsymbol{B}_{(2,1)} + \boldsymbol{B}_{(1,2)}\boldsymbol{B}^{-1}_{(2,2)}\boldsymbol{\Sigma}_{(2,2)}\boldsymbol{B}^{-1}_{(2,2)}\boldsymbol{B}_{(2,1)}.\end{aligned}$$

We write the covariance matrix between $\hat{\boldsymbol{\varepsilon}}_{(1)h}$ and $\boldsymbol{\varepsilon}_{(2)h}$ as

$$(20)\quad \hat{\boldsymbol{\Sigma}}_{(1,2)} = E\left(\hat{\boldsymbol{\varepsilon}}'_{(1)h}\boldsymbol{\varepsilon}'_{(2)h}\right) = \boldsymbol{\Sigma}_{(1,2)} - \boldsymbol{B}_{(1,2)}\boldsymbol{B}^{-1}_{(2,2)}\boldsymbol{\Sigma}_{(2,2)},$$

and the conditional mean for $\boldsymbol{\varepsilon}_{(2)h}$ given $\hat{\boldsymbol{\varepsilon}}_{(1)h}$ as,

$$(21)\quad E\left(\boldsymbol{\varepsilon}_{(2)h} \mid \hat{\boldsymbol{\varepsilon}}_{(1)h}\right) = \hat{\boldsymbol{\Sigma}}_{(2,1)}\hat{\boldsymbol{\Sigma}}^{-1}_{(1,1)}\hat{\boldsymbol{\varepsilon}}_{(1)h}.$$

The conditional variance-covariance matrix for $\boldsymbol{\varepsilon}_{(2)h}$ given $\hat{\boldsymbol{\varepsilon}}_{(1)h}$ is

$$(22)\quad \hat{\boldsymbol{\Sigma}}_{(2,2)} = E\left(\boldsymbol{\varepsilon}_{(2)h}\boldsymbol{\varepsilon}'_{(2)h} \mid \hat{\boldsymbol{\varepsilon}}_{(1)h}\right) = \boldsymbol{\Sigma}_{(2,2)} - \hat{\boldsymbol{\Sigma}}_{(2,1)}\hat{\boldsymbol{\Sigma}}^{-1}_{(1,1)}\hat{\boldsymbol{\Sigma}}_{(1,2)}.$$

We could derive the probability of inequality constraint for $\hat{\boldsymbol{\varepsilon}}_{(1)h}$ from the marginal joint normal pdf for $\hat{\boldsymbol{\varepsilon}}_{(1)h}$,

$$(23)\quad \Pr\left[\hat{\boldsymbol{\varepsilon}}_{(1)h} > -\left(\boldsymbol{\alpha}_{(1)} - \boldsymbol{B}_{(1,2)}\boldsymbol{B}^{-1}_{(2,2)}\boldsymbol{\alpha}_{(2)}\right) - \left(\boldsymbol{A}_{(1)} - \boldsymbol{B}_{(1,2)}\boldsymbol{B}^{-1}_{(2,2)}\boldsymbol{A}_{(2)}\right)\boldsymbol{z}_h - \left(\boldsymbol{B}_{(1,1)} - \boldsymbol{B}_{(1,2)}\boldsymbol{B}^{-1}_{(2,2)}\boldsymbol{B}'_{(1,2)}\right)\boldsymbol{x}_{(1)h}\right],$$

and the probability of inequality constraint for $\boldsymbol{\varepsilon}_{(2)h}$ given $\hat{\boldsymbol{\varepsilon}}_{(1)h}$ from the conditional joint normal pdf for $\boldsymbol{\varepsilon}_{(2)h}$ given $\hat{\boldsymbol{\varepsilon}}_{(1)h}$,

$$(24)\quad \Pr\left[\boldsymbol{\varepsilon}_{(2)h} \le -\left(\boldsymbol{\alpha}_{(2)} + \boldsymbol{A}_{(2)}\boldsymbol{z}_h + \boldsymbol{B}'_{(1,2)}\boldsymbol{x}_{(1)h} + \boldsymbol{B}_{(2,2)}\boldsymbol{x}_{(2)h}\right) \mid \hat{\boldsymbol{\varepsilon}}_{(1)h}\right].$$

Combining Equations (23) and (24), we can obtain the unconditional probability of being in this regime,

$$(25)\quad \begin{aligned}&\Pr\left(\boldsymbol{y}_{(1)h} > \boldsymbol{0}_{(1)h}, \boldsymbol{y}_{(2)h} = \boldsymbol{0}_{(2)h}\right)\\ &= \Pr\left\{\begin{array}{l}\left[\hat{\boldsymbol{\varepsilon}}_{(1)h} > -\left(\boldsymbol{\alpha}_{(1)} - \boldsymbol{B}_{(1,2)}\boldsymbol{B}^{-1}_{(2,2)}\boldsymbol{\alpha}_{(2)}\right) - \left(\boldsymbol{A}_{(1)} - \boldsymbol{B}_{(1,2)}\boldsymbol{B}^{-1}_{(2,2)}\boldsymbol{A}_{(2)}\right)\boldsymbol{z}_h - \left(\boldsymbol{B}_{(1,1)} - \boldsymbol{B}_{(1,2)}\boldsymbol{B}^{-1}_{(2,2)}\boldsymbol{B}'_{(1,2)}\right)\boldsymbol{x}_{(1)h}\right],\\ \boldsymbol{\varepsilon}_{(2)h} \le -\left(\boldsymbol{\alpha}_{(2)} + \boldsymbol{A}_{(2)}\boldsymbol{z}_h + \boldsymbol{B}'_{(1,2)}\boldsymbol{x}_{(1)h} + \boldsymbol{B}_{(2,2)}\boldsymbol{x}_{(2)h}\right)\end{array}\right\}.\end{aligned}$$

We could use the unconditional probability in Equation (25) to derive the conditional mean for $\boldsymbol{y}_{(1)h}$ given $\boldsymbol{y}_{(1)h} > \mathbf{0}_{(1)h}$ and $\boldsymbol{y}_{(2)h} = \mathbf{0}_{(2)h}$,

$$\begin{aligned} E\left(\boldsymbol{y}_{(1)h} \mid \boldsymbol{y}_{(1)h} > \mathbf{0}_{(1)h}, \boldsymbol{y}_{(2)h} = \mathbf{0}_{(2)h}\right) = {} & \left[\boldsymbol{\alpha}_{(1)} - \boldsymbol{B}_{(1,2)}\boldsymbol{B}_{(2,2)}^{-1}\boldsymbol{\alpha}_{(2)}\right] + \left[\boldsymbol{A}_{(1)} - \boldsymbol{B}_{(1,2)}\boldsymbol{B}_{(2,2)}^{-1}\boldsymbol{A}_{(2)}\right]\boldsymbol{z}_h + \left[\boldsymbol{B}_{(1,1)} - \boldsymbol{B}_{(1,2)}\boldsymbol{B}_{(2,2)}^{-1}\boldsymbol{B}_{(1,2)}'\right]\boldsymbol{x}_{(1)h} + \\ & + E\left(\begin{array}{l} \hat{\boldsymbol{\varepsilon}}_{(1)h} \mid \hat{\boldsymbol{\varepsilon}}_{(1)h} > -\left[\boldsymbol{\alpha}_{(1)} - \boldsymbol{B}_{(1,2)}\boldsymbol{B}_{(2,2)}^{-1}\boldsymbol{\alpha}_{(2)}\right] \\ \qquad -\left[\boldsymbol{A}_{(1)} - \boldsymbol{B}_{(1,2)}\boldsymbol{B}_{(2,2)}^{-1}\boldsymbol{A}_{(2)}\right]\boldsymbol{z}_h - \left[\boldsymbol{B}_{(1,1)} - \boldsymbol{B}_{(1,2)}\boldsymbol{B}_{(2,2)}^{-1}\boldsymbol{B}_{(1,2)}'\right]\boldsymbol{x}_{(1)h}, \\ \boldsymbol{\varepsilon}_{(2)h} \le -\left(\boldsymbol{\alpha}_{(2)} + \boldsymbol{A}_{(2)}\boldsymbol{z}_h + \boldsymbol{B}_{(1,2)}'\boldsymbol{x}_{(1)h} + \boldsymbol{B}_{(2,2)}\boldsymbol{x}_{(2)h}\right) \end{array} \right). \end{aligned} \tag{26}$$

We need to impose parameter restrictions to satisfy the regularity conditions of the demand system. Parameter restrictions required for Slutsky symmetry and negative definiteness, i.e., downward sloping transformed demands with finite conditional choke prices in every possible regime require $\boldsymbol{B} = \boldsymbol{B}'$ and $\boldsymbol{x}'\boldsymbol{B}\boldsymbol{x} < 0 \ \forall \ \boldsymbol{x} \neq \mathbf{0}$. We can impose this during estimation through the exactly identified Cholesky factorization $\boldsymbol{B} = -\boldsymbol{L}\boldsymbol{L}'$, where $\mathbf{L}$ represents a lower triangular matrix. If we observe positive quantities demanded for any vector of positive prices, then $\boldsymbol{\alpha} + \min_{h \in \{1,\cdots,H\}}\left(\boldsymbol{A}\boldsymbol{z}_h\right) > \mathbf{0}_N$. We impose this data-dependent set of parameter restrictions iteratively during estimation (LaFrance 1989, 1991).

4. An Application to Ale and Lager Demand with Homothetic AIDS Specification

We apply the GQU model to estimate the demand of ales and lagers using a homothetic incomplete AIDS specification, which is a popular demand system and relatively easy to implement with binding non-negativity constraints. We use the transformed demand as $y_i = p_i q_i / m$ and log form of price considering $f(m, \boldsymbol{z}) = \ln\left[a(\boldsymbol{z})m + b(\boldsymbol{z})\right]$ and $g_i(p_i) = \ln p_i, i = 1, 2$ for the homothetic AIDS model in

Equation (7).[4] The income elasticity equals 1 based on Equation (11). The demand for each good depends on households' characteristics, prices, and consumers' random preferences.

To account for consumers' zero consumption decisions, we divide consumers into four groups: those who only purchase good 1, those who only purchase good 2, those who purchase both goods, and those who purchase neither goods. Assume there are N consumers, with N_1 in regime 1, N_2 in regime 2, N_3 in regime 3, and N_4 in regime 4, so that $N = N_1 + N_2 + N_3 + N_4$.

Regime 1

In this regime, a consumer purchases good one, but not good two, $y_{1h} > 0$ and $y_{2h} = 0$,

(27) $y_{1h} = \alpha_1 + \boldsymbol{A}_1 \boldsymbol{s}_h + c_{11} \ln p_{1h} + c_{12} \ln \hat{p}_{2h} + \varepsilon_{1h} > 0$,

(28) $0 \le \alpha_2 + \boldsymbol{A}_2 \boldsymbol{s}_h + c_{12} \ln p_{1h} + c_{22} \ln \hat{p}_{2h} + \varepsilon_{2h}$,

where the intercepts α_1 and α_2 explain the sum of the effects of all other factors that are not included in the model but might influence consumers' purchasing decisions, $\boldsymbol{s}_h$ represents demographic information including household size and its square term, annual income and squared income, education level, age and squared age, presence of child, ethnicity, and region, $\boldsymbol{A}_1$, $\boldsymbol{A}_2$, c_{11} c_{12}, and c_{22} denote the parameters for estimation, and ε_{1h}

[4] Based on Equation (3), we derive the transformed demand using Equation (7) and the values of κ, λ, $a(\boldsymbol{z})$, and $b(\boldsymbol{z})$ for the homothetic AIDS model.

and ε_{2h} capture a vector of consumers' random preferences for good 1 and good 2, respectively.

As an illustration of Equation (12), we show the conditional choke price of good 2 derived from Equation (28) as,

(29) $\ln \hat{p}_{2h} = -(\alpha_2 + \boldsymbol{A}_2 \boldsymbol{s}_h + c_{12} \ln p_{1h} + \varepsilon_{2h}) / c_{22}$,

We derive the inequality constraint of $\hat{\varepsilon}_{1h}$ by substituting the choke price of good 2 in Equation (29) into the demand for good 1 in Equation (27),

$$\hat{\varepsilon}_{1h} \equiv \frac{c_{22}\varepsilon_{1h} - c_{12}\varepsilon_{2h}}{c_{22}}$$

(30) $$> -\left[\left(\frac{c_{22}\alpha_1 - c_{12}\alpha_2}{c_{22}}\right) + \left(\frac{c_{22}\boldsymbol{A}_1 - c_{12}\boldsymbol{A}_2}{c_{22}}\right)\boldsymbol{s}_h + \left(\frac{c_{11}c_{22} - c_{12}^2}{c_{22}}\right)\ln p_{1h}\right]$$

$$\equiv -\left(\hat{\alpha}_1 + \hat{\boldsymbol{A}}_1 \boldsymbol{s}_h + \hat{c}_{11} \ln p_{1h}\right) \equiv -\hat{\ell}_{1h},$$

where $\hat{\alpha}_1 = (c_{22}\alpha_1 - c_{12}\alpha_2)/c_{22}$, $\hat{\boldsymbol{A}}_1 = (c_{22}\boldsymbol{A}_1 - c_{12}\boldsymbol{A}_2)/c_{22}$, and $\hat{c}_{11} = (c_{11}c_{22} - c_{12}^2)/c_{22}$. We show the inequality constraint of ε_{2h} using Equation (31) with market price of good 2,

(31) $\varepsilon_{2h} \geq -\left(\alpha_2 + \boldsymbol{A}_2 \boldsymbol{s}_h + c_{12} \ln p_{1h} + c_{22} \ln p_{2h}\right) \equiv -\ell_{2h}$.

Equations (30) and (31) are specifications of Equations (14) and (15), respectively.

In a large sample set, we reasonably assume that ε_h are normally distributed with zero means and variance-covariance matrix $\begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}$, $\begin{bmatrix} \varepsilon_{1h} \\ \varepsilon_{2h} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}\right)$.

We derive the conditional means of the goods for Regimes 1-3 in Appendix 2. The conditional mean for y_1 given $y_1 > 0$ and $y_2 = 0$ is expressed as,

(32) $$E\left(y_{1h}\middle|y_{1h}>0, y_{2h}=0\right)=\hat{\alpha}_1+\hat{\boldsymbol{A}}_1\boldsymbol{s}_h+\hat{c}_{11}\ln p_{1h}+\frac{\sigma_{\hat{\varepsilon}_1}\displaystyle\int_{-\hat{\ell}_{1h}/\sigma_{\hat{\varepsilon}_1}}^{\infty} w\varphi(w)\Phi\left(\frac{-\sigma_{\hat{\varepsilon}_1}\ell_{2h}-\sigma_{\hat{\varepsilon}_1\varepsilon_2}w}{\sqrt{\sigma_1^2\sigma_2^2-\sigma_{12}^2}}\right)dw}{\displaystyle\int_{-\hat{\ell}_{1h}/\sigma_{\hat{\varepsilon}_1}}^{\infty} \varphi(w)\Phi\left(\frac{-\sigma_{\hat{\varepsilon}_1}\ell_{2h}-\sigma_{\hat{\varepsilon}_1\varepsilon_2}w}{\sqrt{\sigma_1^2\sigma_2^2-\sigma_{12}^2}}\right)dw}.$$

This is a specific expression of Equation (26). The last term of Equation (32) represents the bivariate analogue to the inverse Mill's Ratio, common in censored regression and sample selection problems.

Regime 2

In this regime, a consumer purchases good two, but not good one, $y_{1h}=0$ and $y_{2h}>0$. Similar to Regime 1, we obtain the corresponding results with the roles of the two goods, prices, and error terms reversed. We express the conditional choke price for the first good as,

(33) $\ln\hat{p}_{1h}=-(\alpha_1+\boldsymbol{A}_1\boldsymbol{s}_h+c_{12}\ln p_{2h}+\varepsilon_{1h})/c_{11}$.

We substitute the choke price of good 1 into the demand equation for the second good,

(34) $y_{2h}=\alpha_2+\boldsymbol{A}_2\boldsymbol{s}_h+c_{12}\ln\hat{p}_{1h}+c_{22}\ln p_{2h}+\varepsilon_{2h}>0$

to obtain the inequality constraint of $\hat{\varepsilon}_{2h}$,

(35) $\hat{\varepsilon}_{2h}\equiv(-c_{12}\varepsilon_{1h}+c_{11}\varepsilon_{2h})/c_{11}>-(\hat{\alpha}_2+\hat{\boldsymbol{A}}_2\boldsymbol{s}_h+\hat{c}_{22}\ln p_{2h})\equiv-\hat{\ell}_{2h}$,

where $\hat{\alpha}_2=(-c_{12}\alpha_1+c_{11}\alpha_2)/c_{11}$, $\hat{\boldsymbol{A}}_2=(-c_{12}\boldsymbol{A}_1+c_{11}\boldsymbol{A}_2)/c_{11}$, and $\hat{c}_{22}=(c_{11}c_{22}-c_{12}^2)/c_{11}$.

Using zero consumption of good 1, we show the inequality constraint of ε_{1h} as,

(36) $\varepsilon_{1h}\geq\alpha_1+\boldsymbol{A}_1\boldsymbol{s}_h+c_{11}\ln p_{1h}+c_{12}\ln p_{2h}\equiv\ell_{1h}$.

We obtain the conditional mean for y_2 given $y_1=0$ and $y_2>0$,

$$(37)\quad E\left(y_{2h} \middle| y_{1h}=0, y_{2h}>0\right) = \hat{\alpha}_2 + \hat{\boldsymbol{A}}_2 \boldsymbol{s}_h + \hat{c}_{22} \ln p_{2h} + \frac{\sigma_{\hat{\varepsilon}_2} \int_{-\hat{\ell}_{2h}/\sigma_{\hat{\varepsilon}_2}}^{\infty} w\varphi(w)\Phi\left(\frac{-\sigma_{\hat{\varepsilon}_2}\ell_{1h} - \sigma_{\varepsilon_1\hat{\varepsilon}_2} w}{\sqrt{\sigma_1^2\sigma_2^2 - \sigma_{12}^2}}\right)dw}{\int_{-\hat{\ell}_{2h}/\sigma_{\hat{\varepsilon}_2}}^{\infty} \varphi(w)\Phi\left(\frac{-\sigma_{\hat{\varepsilon}_2}\ell_{1h} - \sigma_{\varepsilon_1\hat{\varepsilon}_2} w}{\sqrt{\sigma_1^2\sigma_2^2 - \sigma_{12}^2}}\right)dw}.$$

Regime 3

In this regime, the consumer purchases both goods, $y_{1h} > 0$ and $y_{2h} > 0$. Hence, we do not use choke prices. We write the demand functions as,

$$(38)\quad \begin{aligned} y_{1h} &= \alpha_1 + \boldsymbol{A}_1\boldsymbol{s}_h + c_{11} \ln p_{1h} + c_{12} \ln p_{2h} + \varepsilon_{1h} > 0, \\ y_{2h} &= \alpha_2 + \boldsymbol{A}_2\boldsymbol{s}_h + c_{12} \ln p_{1h} + c_{22} \ln p_{2h} + \varepsilon_{2h} > 0. \end{aligned}$$

The inequality constraints are derived from Equation (38) as,

$$(39)\quad \begin{aligned} \varepsilon_{1h} &> -(\alpha_1 + \boldsymbol{A}_1\boldsymbol{s}_h + c_{11} \ln p_{1h} + c_{12} \ln p_{2h}) = -\ell_{1h}, \\ \varepsilon_{2h} &> -(\alpha_2 + \boldsymbol{A}_2\boldsymbol{s}_h + c_{12} \ln p_{1h} + c_{22} \ln p_{2h}) = -\ell_{2h}. \end{aligned}$$

We show the conditional mean of y_1 given $y_{1h} > 0$ and $y_{2h} > 0$ as,

$$(40)\quad E\left(y_{1h} \middle| y_{1h}>0, y_{2h}>0\right) = \alpha_1 + \boldsymbol{A}_1 \boldsymbol{s}_h + c_{11} \ln p_{1h} + c_{12} \ln p_{2h} + \frac{\sigma_1 \int_{-\ell_{1h}/\sigma_1}^{\infty} w\varphi(w)\Phi\left(\frac{\sigma_1\ell_{2h} + \sigma_{12} w}{\sqrt{\sigma_1^2\sigma_2^2 - \sigma_{12}^2}}\right)dw}{\int_{-\ell_{1h}/\sigma_1}^{\infty} \varphi(w)\Phi\left(\frac{\sigma_1\ell_{2h} + \sigma_{12} w}{\sqrt{\sigma_1^2\sigma_2^2 - \sigma_{12}^2}}\right)dw},$$

and the conditional mean of y_2 given $y_{1h} > 0$ and $y_{2h} > 0$ is,

$$(41)\quad E\left(y_{2h} \middle| y_{1h}>0, y_{2h}>0\right) = \alpha_2 + \boldsymbol{A}_2 \boldsymbol{s}_h + c_{12} \ln p_{1h} + c_{22} \ln p_{2h} + \frac{\sigma_2 \int_{-\ell_{2h}/\sigma_2}^{\infty} w\varphi(w)\Phi\left(\frac{\sigma_2\ell_{1h} + \sigma_{12} w}{\sqrt{\sigma_1^2\sigma_2^2 - \sigma_{12}^2}}\right)dw}{\int_{-\ell_{2h}/\sigma_2}^{\infty} \varphi(w)\Phi\left(\frac{\sigma_2\ell_{1h} + \sigma_{12} w}{\sqrt{\sigma_1^2\sigma_2^2 - \sigma_{12}^2}}\right)dw}.$$

Regime 4

In this regime, a consumer does not consume either good, $y_{1h} = y_{2h} = 0$. The unconditional probability of being in this regime equals to one minus the sum of the probabilities of being in Regimes 1-3.

5. Closed-Form Solutions and Estimation

In order to avoid solving high-dimensional integrals and obtain closed-form solutions in each demand equation, we need transformations of univariate integrals and approximations to standard normal cumulative and probability density functions. We transform the integral to a compact interval on the real line,

$$(42)\ \int_a^\infty f(x)dx \approx \int_0^1 f\left(a + \frac{t}{1-t}\right)\frac{1}{(1-t)^2}dt \approx \int_0^{\frac{999}{1000}} f\left(a + \frac{t}{1-t}\right)\frac{1}{(1-t)^2}dt\,.$$

To avoid dividing by zero, we use 999/1000 as an approximation of 1 for the upper bound in Equation (42). Then we use Simpson’s rule as a third order approximation for numerical integration,

$$(43)\ \int_a^b f(x)dx \approx \frac{b-a}{6} f(a) + \frac{2(b-a)}{3} f\left(\frac{a+b}{2}\right) + \frac{b-a}{6} f(b)\,.$$

We use a closed-form third-order logistic approximation to the standard normal to calculate the integral,

$$(44)\quad \begin{aligned} \Phi(x) &\approx \frac{e^{1.6033914x\ +\ .067671036x^3}}{1+e^{1.6033914x\ +\ .067671036x^3}}, \\ \varphi(x) &\approx (1.6033914\ +\ 0.20301309x^2)\Phi(x)\left[1-\Phi(x)\right]. \end{aligned}$$

This distribution matches the first four moments to a standard normal exactly and gives a very close approximation to both the cumulative and probability density functions. Using these approximations, we convert the integral to a non-linear function of demographics, prices, and parameters.

We can identify all the unknown parameters $\alpha_1, \alpha_2, \boldsymbol{A}_1, \boldsymbol{A}_2, c_{11}, c_{12}, c_{22}, \sigma_1, \sigma_{12}$, and σ_2 by estimating the demand equations in which the conditional means expressed as Equations (40) and (41) in Regime 3 using seemingly unrelated regression and imposing the following cross-equation parameter restrictions,

$$(45)\quad \begin{bmatrix} c_{11} & c_{12} \\ c_{12} & c_{22} \end{bmatrix} < 0,\ \alpha_1 + \min_{h\in\{1,\cdots,H\}} \boldsymbol{A}_1 \boldsymbol{s}_h > \boldsymbol{0},\ \alpha_2 + \min_{h\in\{1,\cdots,H\}} \boldsymbol{A}_2 \boldsymbol{s}_h > \boldsymbol{0}.$$

We calculate the parameters $\hat{\alpha}_1, \hat{\alpha}_2, \hat{\boldsymbol{A}}_1, \hat{\boldsymbol{A}}_2, \hat{c}_{11}, \hat{c}_{22}, \sigma_{\hat{\varepsilon}_1}, \sigma_{\hat{\varepsilon}_1\varepsilon_2}, \sigma_{\hat{\varepsilon}_2}$, and $\sigma_{\varepsilon_1\hat{\varepsilon}_2}$ in Equation (32) and (37) for Regimes 1-2 using the following recursive relationships,

$$\hat{\alpha}_1 = \alpha_1 - (c_{12}/c_{22})\alpha_2,\ \hat{\alpha}_2 = \alpha_2 - (c_{12}/c_{11})\alpha_1,$$

$$\hat{\boldsymbol{A}}_1 = \boldsymbol{A}_1 - c_{12}\boldsymbol{A}_2/c_{22},\ \hat{\boldsymbol{A}}_2 = \boldsymbol{A}_2 - c_{12}\boldsymbol{A}_1/c_{11},$$

$$\hat{c}_{11} = c_{11} - c_{12}^2/c_{22},\quad \hat{c}_{22} = c_{22} - c_{12}^2/c_{11},$$

$$(46)\quad \sigma_{\hat{\varepsilon}_1}^2 = \sigma_1^2 - 2(c_{12}/c_{22})\sigma_{12} + (c_{12}/c_{22})^2\sigma_2^2,$$

$$\sigma_{\hat{\varepsilon}_1\varepsilon_2} = \sigma_{12} - (c_{12}/c_{22})\sigma_2^2,$$

$$\sigma_{\hat{\varepsilon}_2}^2 = \sigma_2^2 - 2(c_{12}/c_{11})\sigma_{12} + (c_{12}/c_{11})^2\sigma_1^2,$$

$$\sigma_{\varepsilon_1\hat{\varepsilon}_2} = \sigma_{12} - (c_{12}/c_{11})\sigma_1^2.$$

To take heteroscedasticity into account, we make inferences using White's consistent covariance estimates (White 1980). Then we can get consistent parameter estimates. We use parameter estimates of linear demand functions without integral terms in Regime 3 as the starting values for estimation. For statistical inferences, we use the delta method to calculate the approximate standard errors of marginal effects of demographics and prices.

6. Data

We use IRI 2005 beer data which includes expenditures and quantities of beer consumption and consumers' demographic characteristics for 5,225 households in two cities, Pittsfield, MA and Eau Claire, WI. We aggregate from weekly to annual observations to reduce the observations of zero consumption due to inventory and other reasons except for economic factors. We divide the beer into ales and lagers as two general categories of goods.[5] The dataset includes 2,269 purchasing records for beer.

Table 1 shows demographic characteristics for households in the dataset. The average household contains 2-3 people. Household heads average 53 years old and the median age of the U.S. population is 36.2 years.[6] The average annual household income is $53,000, which is higher than the median of U.S. household income $46,326 in 2005. Approximately, 54% of household heads finished high school as a maximum education level, 35% of household heads graduated from college, and 5% gained a post college degree. The educational attainment is higher than the U.S. average level that 27.6% of population with Bachelors' degree or higher. Nearly 29% households have a child under the age of 18. White families account for nearly 98% in the sample which is 15% above the national level. Nearly 37% households reside in Pittsfield, MA.

We present the summary statistics of beer consumption in Table 2. Among beer purchasing households, less than 18% of households purchased ales and about 97% purchased lagers at least once in 2005. Both lager expenditure and volume of consumption outweigh those of ales. On average, consumers spent $69 on lagers annually

[5] For ales, we include ale, bitter ale, cream ale, cream stout, extra stout, hefeweizen, hefeweizen ale, lambic, oatmeal stout, porter, stock ale, stout, weisse, and wheat. For lagers, barley lager, bock, dopplebock, double bock, dunkel, lager, malt, malt liquor, oktoberfest, pilsner, and stout lager are incorporated. We drop the observations with missing or assorted type.
[6] The national level data in 2005 come from U.S. Census Bureau.

and purchased 123 bottles (12 ounces per bottle) per household in the whole sample. Households pay relatively higher unit prices for ales than lagers.

Table 3 reports the expenditures, quantities, and prices of ales and lagers and the percentage of households in different regimes. Most households consume lagers only, which accounts for about 82% among beer drinkers. Only about 3% of households purchase solely ales. The expenditure and quantity in Regime 3 are larger than other regimes, while we find similar prices across different regimes for ales and lagers.

7. Empirical Results

Table 4 provides the marginal effects of demographics and prices on ale and lager demand in different regimes. We find household annual income, age, presence of child, and prices play significant roles on quantities demanded for ales and lagers in all regimes. All the factors have greater effects on the demand of lager than ale and the impacts in Regime 3 where consumers purchase both products are generally greater than other regimes in absolute terms. Most consumers prefer lagers so the change of demand for lagers is more sensitive to the variations of demographics and prices compared with ales.

We observe a negative effect of education on beer consumption by comparing the marginal effects for high school and college graduates, which might be the case that more educated consumers have more exposure to the health information and are more likely to drink in moderation. Age has a positive impact on demand of ales and lagers. Families with children consume more beers. White households have higher demand for beer, but the effect is not significant which might be caused by the insufficient variations in ethnicity of the sample. Consumers living in Pittsfield purchase less beer compared with those who reside in Eau Claire.

Income positively affects the quantities demanded for both goods, which implies ales and lagers are normal goods. Household size has a negative and insignificant impact. Consistent with economic theory, an increase in ale (or lager) price would lead to a decreased quantity demanded and an increase in ale price would result in increased consumption on lagers, and vice versa.

We report uncompensated price elasticities of demand of ales and lagers in Table 5. We find both of the own-price elasticities negative and elastic. Lager beers have greater own-price elasticity than ales in absolute values. Ales and lagers act as gross substitutes considering the positive cross-price elasticities. We observe the elasticity of lager demand with respect to the price of ale greater than unity, but the cross-price elasticity of ale demand is inelastic. This suggests that the price of ale has a large effect on lager demand, while an increase in the price of lager cannot generate significant change of ale demand.

8. Conclusions

We introduce a new demand system with binding non-negativity constraints, the Generalized Quadratic Utility model, in which we make non-zero and non-constant income elasticity and non-linear price response possible. We specify the well-defined reduced forms of demand equations while keeping the regularity conditions satisfied. Using this generalized and theoretically consistent demand system, we could avoid solving a complicated integration and obtain closed-form solutions. This estimation strategy overcomes the analytical and computational difficulties that limited the application of Kuhn-Tucker models. We can apply this GQU model to complex demand systems with a large number of products.

We provide illustrations of the GQU structural model and its estimation with a homothetic AIDS specification. Using data of beer consumption, we examine the demand of ales and lagers and report the marginal effects of demographics and prices, as well as price elasticities. Household annual income, age, presence of child, and prices play significant roles in determining beer consumption. Consumers with higher income purchase more ales and lagers. Age positively affects the beer consumption. Households with children have more preferences for both types of beers. We find negative own-price elasticities which are greater than unity. The positive cross-price elasticities for ales and lagers imply that they serve as gross substitutes.

References

Arndt, C., “Demand for Herbicide in Corn: An Entropy Approach Using Micro-Level Data,” Journal of Agricultural and Resource Economics 24(1999): 204-221.

Bray, J.W., Loomis, B.R., and Engelen, M., “You Save Money When You Buy in Bulk: Does Volume-Based Pricing Cause People to Buy More Beer?” Health Economics 18 (2009): 607-618.

Choi, D.Y., Stack, M.H., “The All-American Beer: a Case of Inferior Standard (taste) Prevailing?” Business Horizons 48(2005):79-86.

Connelly, S., “Beer Style Review-Light Lager,” The Beer Connoisseur Online, URL: http://www.beerconnoisseur.com/light-lager, Data accessed: 8 October, 2013.

Cragg, J., “Some Statistical Models for Limited Dependent Variables with Application to the Demand for Durable Goods,” Econometrica 39 (1971): 829-844.

Deaton, A.S., Demand Analysis, in Handbook of Economics, Volume III, Edited by Griliches Z. and Intriligator M. D., North-Holland, 1986.

Dixit, A.K. and Weller, P.A., “The Three Consumer’s Surpluses,” Economica, London School of Economics and Political Science 46 (1979): 125-135.

Gallet, C.A., “The Demand for Alcohol: a Meta-Analysis of Elasticities,” Australian Journal of Agricultural and Resource Economics 51(2007): 121-135.

Greene, H.W., “On the Asymptotic Bias of the Ordinary Least-Squares Estimator of the Tobit Model,” Econometrica 49 (1981): 505-514.

Hausman, J., “Exact Consumers’ Surplus and Deadweight Loss,” American Economic Review 71 (1981): 662-676.

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

Kim, J., Allenby, G.M., and Rossi, P.E., “Modeling Consumer Demand for Variety,” Marketing Science 21(2002): 229-250.

LaFrance, J.T., “Linear Demand Functions in Theory and Practice,” Journal of Economic Theory 37(1985): 147-166.

____, “Weak Separability and Two-Stage Budgeting: Problems of Estimation and Welfare Management,” Paper presented at the Australasian meeting of the Econometric Society, 1989.

____, “When is Expenditure ‘Exogenous’ in Separable Demand Models?” Western Journal of Agricultural Economics 16 (1991): 49-62.

____, “Integrability of the Linear Approximate Almost Ideal Demand System,” Economics Letters 84 (2004): 297-303.

LaFrance, J.T. and Hanemann, W.M., “The Dual Structure of Incomplete Demand Systems,” American Journal of Agricultural Economics 71 (1989): 262-274.

Lee, L.F., and Pitt M.M., “Microeconometric Demand Systems with Binding Nonnegativity Constraints: The Dual Approach,” Econometrica 54 (1986):1237-1242.

____, “Microeconometric Models of Rationing, Imperfect Markets, and Non-Negativity Constraints,” Journal of Econometrics 36 (1987): 89-110.

Lee, B., and Tremblay, V.J., “Advertising and the Market Demand for Beer,” Applied Economics 24(1992):69-76.

Olsen, V., "Approximating a Truncated Normal Regression with the Method of Moments," Econometrica 48 (1980): 1099-1105.
Ornstein, S.I. and Hanssens, D.M., "Alcohol Control Laws and the Consumption of Distilled Spirits and Beer," Journal of Consumer Research 12(1985): 200-213.
Phaneuf, D., Kling, C., and Herriges, J., "Estimation and Welfare Calculations in a Generalized Corner Solution Model with an Application to Recreation Demand," Review of Economics and Statistics 82(2000): 83-92.
Protz, R., The Ultimate Encyclopedia of Beer, London: Carleton Books, 1995.
Ransom, M., "A Comment on Consumer Demand Systems with Binding Non-Negativity Constraints," Journal of Econometrics 34(1987): 355-359.
Shonkwiler, J.S. and Yen, S.T., "Two-Step Estimation of a Censored System of Equations," American Journal of Agricultural Economics 81(1999): 972-982.
Tobin, J., "Estimation of Relationships for Limited Dependent Variables," Econometrica 46 (1958): 24-36.
Tremblay, V.J. and Tremblay, C.H., The U.S. Brewing Industry, Cambridge: MIT Press, 2005.
Van Soest, A. and Kooreman, P., "Coherency of the Indirect Translog Demand System with Binding Non-Negativity Constraints," Journal of Econometrics 44(1990):391-400.
Van Soest, A., Kooreman, P., and Kapteyn, A., "Coherency and Regularity of Demand Systems with Equality and Inequality Constraints," Journal of Econometrics 57(1993):161-188.
Wales, T. J., and Woodland, D.A., "Estimation of Consumer Demand Systems with Binding Nonnegativity Constraints," Journal of Econometrics 21 (1983):263-285.
White, H., "A Heteroskedasticity-Consistent Covariance-Matrix Estimator and a Direct Test for Heteroskedasticity," Econometrica 48(1980): 817-838.

Table 1: Descriptive Statistics for Household Demographics

Variable	Definition	Min	Max	Mean	Standard Deviation
Household Size	Number of household members	1.000	6.000	2.678	1.233
Household Income	Annual income in $10,000	0.000	13.000	5.314	3.412
Grade School	Max education of household head some/graduated grade school	0.000	1.000	0.060	0.254
High School	Max education of household head some/graduated high school	0.000	1.000	0.541	0.498
College	Max education of household head some/graduated college	0.000	1.000	0.347	0.476
Post College	Max education of household head post college	0.000	1.000	0.052	0.223
Age	Mean age of female and male heads	21.000	70.000	52.972	12.924
Presence of Child	Presence of child under 18	0.000	1.000	0.294	0.456
White	Household head race is white	0.000	1.000	0.981	0.138
Black	Household head race is black	0.000	1.000	0.005	0.073
Asian	Household head race is Asian	0.000	1.000	0.003	0.055
Hispanic	Household head race is Hispanic	0.000	1.000	0.001	0.030
Other Race	Household head race is other race	0.000	1.000	0.010	0.066
Pittsfield	Resides in Pittsfield, MA	0.000	1.000	0.369	0.483
Eau Claire	Resides in Eau Claire, WI	0.000	1.000	0.631	0.483

Table 2: Annual Expenditures, Quantities, and Prices for Ales and Lagers and the Percentage of Consuming Households in the Sample

Beer	**Expenditures ($)**		**Quantity (12 ounces)**		**Price ($/12 ounces)**	**Number of Consuming Households**	**Percentage of Consuming Households in the Sample**
	Whole Sample	Consuming Households	Whole Sample	Consuming Households	Consuming Households		
Ale	3.66 (13.22)	20.90 (25.24)	4.27 (16.05)	24.36 (31.33)	0.942 (0.30)	398	17.54%
Lager	68.91 (101.87)	71.10 (147.43)	122.89 (303.01)	126.80 (306.99)	0.731 (0.23)	2199	96.91%

Note: Standard errors in parentheses.

Table 3: Annual Expenditures, Quantities, and Prices for Ales and Lagers and the Percentage among Beer Drinkers

	Beer	Expenditures ($)	Quantity (12 ounces)	Price ($/12 ounces)	Number of Consuming Households	Percentage of Households among Beer Drinkers
Regime 1 $\begin{cases} q_{Ale} > 0 \\ q_{Lager} = 0 \end{cases}$	Ale	15.41 (24.55)	17.64 (29.53)	0.96 (0.23)	70	3.09%
	Lager	-	-	-		
Regime 2 $\begin{cases} q_{Ale} = 0 \\ q_{Lager} > 0 \end{cases}$	Ale	-	-	-	1871	82.46%
	Lager	67.08 (137.73)	120.74 (287.47)	0.72 (0.23)		
Regime 3 $\begin{cases} q_{Ale} > 0 \\ q_{Lager} > 0 \end{cases}$	Ale	22.07 (25.26)	25.79 (31.56)	0.93 (0.32)	328	14.46%
	Lager	94.02 (192.37)	161.37 (399.37)	0.77 (0.22)		

Note: Standard errors in parentheses.

Table 4: Marginal Effects on the Demand for Ales and Lagers

	Regime 1	Regime 2	Regime 3	
	$q_{Ale} > 0, q_{Lager} = 0$	$q_{Ale} = 0, q_{Lager} > 0$	$q_{Ale} > 0, q_{Lager} > 0$	
	Ale	**Lager**	**Ale**	**Lager**
Household Size	-9.32	-18.92	-13.05	-21.22
Household Income	1.22**	13.02*	3.19**	14.61**
High School	-27.05	-129.05	-37.16	-156.32
College	-28.35	-180.41	-42.01	-211.24
Age	3.84***	59.93**	7.36***	64.24***
Presence of Child	20.44***	191.23*	31.90***	210.01*
White	5.92	124.11	2.21	18.16
Pittsfield	-2.10	-3.94	-2.39	-5.34
Price of Ale	-50.89***	-	-53.43***	39.21**
Price of Lager	-	-577.64	39.21**	-606.41***

Notes: The marginal effects are calculated at the sample means. Significance levels: ***0.01; **0.05; *0.10.

Table 5: Uncompensated Price Elasticities of the Demand of Ales and Lagers

	The Elasticity of Product with respect to:					
	Regime 1 $q_{Ale} > 0, q_{Lager} = 0$		**Regime 2** $q_{Ale} = 0, q_{Lager} > 0$		**Regime 3** $q_{Ale} > 0, q_{Lager} > 0$	
	Ale	**Lager**	**Ale**	**Lager**	**Ale**	**Lager**
Ale	-2.88	-	-	-	-2.07	0.24
Lager	-	-	-	-4.78	1.52	-3.76

Note: The elasticities are calculated at the sample means.

Appendix 1: Derivation of the Indirect Utility Function in the GQU Demand System

We derive the utility function associated with this demand function to verify consistency with economic theory by rewriting Equation (3) as,

(A1) $\frac{\partial f}{\partial \boldsymbol{x}} = \boldsymbol{\alpha} + \boldsymbol{A}\boldsymbol{z} + \boldsymbol{B}\boldsymbol{x} + \boldsymbol{\varepsilon}$.

We integrate both sides of Equation (A1) to yield,

(A2) $f(e(\boldsymbol{p}(\boldsymbol{x}), \tilde{\boldsymbol{p}}, \boldsymbol{z}, u), \boldsymbol{z}) = (\boldsymbol{\alpha} + \boldsymbol{A}\boldsymbol{z} + \boldsymbol{\varepsilon})^{\top} \boldsymbol{x} + \frac{1}{2} \boldsymbol{x}^{\top} \boldsymbol{B}\boldsymbol{x} + \theta(\tilde{\boldsymbol{p}}, \boldsymbol{z}, u)$,

where $\theta(\tilde{\boldsymbol{p}}, \boldsymbol{z}, u)$ denotes the function of integration. We cannot identify this function from the demands for q, but know it depends on other goods' prices, demographics, and the utility index and exhibits zero degree homogeneity in $\tilde{\boldsymbol{p}}$ and increases in u.

We obtain the quasi-indirect utility function,

(A3) $\varphi(\boldsymbol{p}, \tilde{\boldsymbol{p}}, \boldsymbol{z}, m) = f(m, \boldsymbol{z}) - (\boldsymbol{\alpha} + \boldsymbol{A}\boldsymbol{z} + \boldsymbol{\varepsilon})^{\top} \boldsymbol{g}(\boldsymbol{p}) - \frac{1}{2} \boldsymbol{g}(\boldsymbol{p})^{\top} \boldsymbol{B}\boldsymbol{g}(\boldsymbol{p})$,

motived by the duality theory that $\theta(\tilde{\boldsymbol{p}}, \boldsymbol{z}, u^*) = \varphi(\boldsymbol{p}, \tilde{\boldsymbol{p}}, \boldsymbol{z}, m)$, where u^* is the maximized utility level (Hausman 1981, LaFrance 1985, 2004, LaFrance and Hanneman 1989).

Taking other goods' prices and consumers' demographics into consideration for a representative consumer, the class of indirect utility functions consistent with this incomplete demand system has the structure,

(A4)
$$\begin{aligned} v(\boldsymbol{p}, \tilde{\boldsymbol{p}}, \boldsymbol{z}, m) &\equiv \upsilon\left(\varphi(\boldsymbol{p}, \tilde{\boldsymbol{p}}, \boldsymbol{z}, m), \tilde{\boldsymbol{p}}, \boldsymbol{z}\right) \\ &= \upsilon\left[f(m, \boldsymbol{z}) - (\boldsymbol{\alpha} + \boldsymbol{A}\boldsymbol{z} + \boldsymbol{\varepsilon})^{\top} \boldsymbol{g}(\boldsymbol{p}) - \tfrac{1}{2} \boldsymbol{g}(\boldsymbol{p})^{\top} \boldsymbol{B}\boldsymbol{g}(\boldsymbol{p}), \tilde{\boldsymbol{p}}, \boldsymbol{z} \right], \end{aligned}$$

where $\upsilon\left(\varphi(\boldsymbol{p}, \tilde{\boldsymbol{p}}, \boldsymbol{z}, m), \tilde{\boldsymbol{p}}, \boldsymbol{z}\right)$ is the inverse of $\theta(\tilde{\boldsymbol{p}}, \boldsymbol{z}, u)$ with respect to u.

Appendix 2: Derivation of the Conditional Mean of Demand in Each Regime

Regime 1

We use change of variables and conditioning steps that minimize the number of integrals during model estimation. To do this, we need the marginal probability density function (pdf) for $\hat{\varepsilon}_{1h}$ and the conditional pdf for ε_{2h} given $\hat{\varepsilon}_{1h}$.

Note that $\hat{\varepsilon}_{1h} = (c_{22}\varepsilon_{1h} - c_{12}\varepsilon_{2h})/c_{22} \sim N\left(0, (c_{22}^2\sigma_1^2 - 2c_{12}c_{22}\sigma_{12} + c_{12}^2\sigma_2^2)/c_{22}^2\right)$ and $c_{22} < 0$, so that

(A5) $$f_{\hat{E}_1}(\hat{\varepsilon}_1) = \frac{|c_{22}|}{\sqrt{2\pi(c_{22}^2\sigma_1^2 - 2c_{12}c_{22}\sigma_{12} + c_{12}^2\sigma_2^2)}} \exp\left\{\frac{-c_{22}^2\hat{\varepsilon}_1^2}{2(c_{22}^2\sigma_1^2 - 2c_{12}c_{22}\sigma_{12} + c_{12}^2\sigma_2^2)}\right\}.$$

To find the conditional cumulative density function (cdf) for ε_2 given $\hat{\varepsilon}_1$, for calculation of the conditional probability that $y_2 = 0$ given $y_1 > 0$, we use $\begin{bmatrix}\varepsilon_1 \\ \varepsilon_2\end{bmatrix} \sim N\left(\begin{bmatrix}0 \\ 0\end{bmatrix}, \begin{bmatrix}\sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2\end{bmatrix}\right)$ and $c_{22} < 0$, so that

(A6) $$\sigma_{\hat{\varepsilon}_1}^2 = \frac{c_{22}^2\sigma_1^2 - 2c_{12}c_{22}\sigma_{12} + c_{12}^2\sigma_2^2}{c_{22}^2}, \quad \sigma_{\hat{\varepsilon}_1\varepsilon_2} = \frac{c_{22}\sigma_{12} - c_{12}\sigma_2^2}{c_{22}}.$$

For a bivariate normal, $\varepsilon_2 | \hat{\varepsilon}_1 \sim N\left(\frac{\sigma_{\hat{\varepsilon}_1\varepsilon_2}}{\sigma_{\hat{\varepsilon}_1}^2}\hat{\varepsilon}_1, \sigma_2^2\left(1 - \rho_{\hat{\varepsilon}_1\varepsilon_2}^2\right)\right)$. Since the change of variables from $\begin{bmatrix}\varepsilon_1 \\ \varepsilon_2\end{bmatrix}$ to $\begin{bmatrix}\hat{\varepsilon}_1 \\ \varepsilon_2\end{bmatrix}$ has a unit Jacobian determinant, $\sigma_{\hat{\varepsilon}_1}^2\sigma_2^2 - \sigma_{\hat{\varepsilon}_1\varepsilon_2}^2 = \sigma_1^2\sigma_2^2 - \sigma_{12}^2$, which follows from

$$\sigma_{\hat{\varepsilon}_1}^2\sigma_2^2-\sigma_{\hat{\varepsilon}_1\varepsilon_2}^2=\left(\frac{c_{22}^2\sigma_1^2-2c_{12}c_{22}\sigma_{12}+c_{12}^2\sigma_2^2}{c_{22}^2}\right)\sigma_2^2-\frac{(c_{22}\sigma_{12}-c_{12}\sigma_2^2)^2}{c_{22}^2}$$

(A7)
$$=\frac{c_{22}^2\sigma_1^2\sigma_2^2-2c_{12}c_{22}\sigma_{12}\sigma_2^2+c_{12}^2\sigma_2^4-c_{22}^2\sigma_{12}^2+2c_{12}c_{22}\sigma_{12}\sigma_2^2-c_{12}^2\sigma_2^4}{c_{22}^2}$$

$$=\sigma_1^2\sigma_2^2-\sigma_{12}^2.$$

This property also applies to Regimes 2-3 below. Therefore, we write the conditional probability that $\varepsilon_{2h}\le-(\alpha_2+\boldsymbol{A}_2\boldsymbol{s}_h+c_{12}\ln p_1+c_{22}\ln p_2)\equiv-\ell_{2h}$ given $\hat{\varepsilon}_{1h}$ in terms of a $N(0,1)$ cdf as,

(A8) $$\Pr\left(\varepsilon_2\le-\ell_{2h}\big|\hat{\varepsilon}_{1h}\right)=\Phi\left(\frac{-\sigma_{\hat{\varepsilon}_1}\ell_{2h}-(\sigma_{\hat{\varepsilon}_1\varepsilon_2}/\sigma_{\hat{\varepsilon}_1})\hat{\varepsilon}_{1h}}{\sqrt{\Delta}}\right)=1-\Phi\left(\frac{\sigma_{\hat{\varepsilon}_1}\ell_{2h}+(\sigma_{\hat{\varepsilon}_1\varepsilon_2}/\sigma_{\hat{\varepsilon}_1})\hat{\varepsilon}_{1h}}{\sqrt{\Delta}}\right),$$

where $\Delta=\sigma_{\hat{\varepsilon}_1}^2\sigma_2^2-\sigma_{\hat{\varepsilon}_1\varepsilon_2}^2=\sigma_1^2\sigma_2^2-\sigma_{12}^2$ and the far right expression follows from symmetry of the normal distribution, so that $\Phi(-x)=1-\Phi(x)$ for all $x\in R$.

Given $\hat{\varepsilon}_{1h}\sim N(0,\sigma_{\hat{\varepsilon}_1}^2)$, we apply a change of variables from $\hat{\varepsilon}_{1h}$ to $\hat{\varepsilon}_{1h}=\sigma_{\hat{\varepsilon}_1}w,$ $w\sim N(0,1)$, to write the unconditional probability of being in Regime 1 in terms of a univariate integral of a standard normal random variable,

(A9) $$\Pr\left(y_{1h}>0,y_{2h}=0\right)=\Pr\left(\hat{\varepsilon}_{1h}>-\hat{\ell}_{1h},\varepsilon_{2h}\le-\ell_{2h}\right)=\int_{-\hat{\ell}_{1h}/\sigma_{\hat{\varepsilon}_1}}^{\infty}\varphi(w)\Phi\left(\frac{-\sigma_{\hat{\varepsilon}_1}\ell_{2h}-\sigma_{\hat{\varepsilon}_1\varepsilon_2}w}{\sqrt{\Delta}}\right)dw.$$

The conditional mean for y_1 given $y_1>0$ and $y_2=0$ is,

$$E\left(y_{1h}\big|y_{1h}>0,y_{2h}=0\right)=\hat{\alpha}_1+\hat{\boldsymbol{A}}_1\boldsymbol{s}_h+\hat{c}_{11}\ln p_1+E\left(\hat{\varepsilon}_{1h}\big|\hat{\varepsilon}_{1h}>-\hat{\ell}_{1h},\varepsilon_{2h}\le-\ell_{2h}\right)$$

(A10)
$$=\hat{\alpha}_1+\hat{\boldsymbol{A}}_1\boldsymbol{s}_h+\hat{c}_{11}\ln p_1+\frac{\sigma_{\hat{\varepsilon}_1}\displaystyle\int_{-\hat{\ell}_{1h}/\sigma_{\hat{\varepsilon}_1}}^{\infty}w\varphi(w)\Phi\left(\frac{-\sigma_{\hat{\varepsilon}_1}\ell_{2h}-\sigma_{\hat{\varepsilon}_1\varepsilon_2}w}{\sqrt{\Delta}}\right)dw}{\displaystyle\int_{-\hat{\ell}_{1h}/\sigma_{\hat{\varepsilon}_1}}^{\infty}\varphi(w)\Phi\left(\frac{-\sigma_{\hat{\varepsilon}_1}\ell_{2h}-\sigma_{\hat{\varepsilon}_1\varepsilon_2}w}{\sqrt{\Delta}}\right)dw}.$$

Regime 2

The above derivations and results apply to this case with the roles of the two goods, prices, and error terms reversed.

In this case, $\hat{\varepsilon}_{2h} = (-c_{12}\varepsilon_{1h} + c_{11}\varepsilon_{2h})/c_{11} \sim N\left(0, (c_{12}^2\sigma_1^2 - 2c_{12}c_{11}\sigma_{12} + c_{11}^2\sigma_2^2)/c_{11}^2\right)$, so that

(A11) $$\sigma_{\hat{\varepsilon}_2}^2 = \frac{c_{12}^2\sigma_1^2 - 2c_{12}c_{11}\sigma_{12} + c_{11}^2\sigma_2^2}{c_{11}^2}, \quad \sigma_{\varepsilon_1\hat{\varepsilon}_2} = \frac{-c_{12}\sigma_1^2 + c_{11}\sigma_{12}}{c_{11}}.$$

As above, we denote the conditional probability that

$\varepsilon_{1h} \le -(\alpha_1 + \boldsymbol{A}_1\boldsymbol{s}_h + c_{11}\ln p_1 + c_{12}\ln p_2) \equiv -\ell_{1h}$ given $\hat{\varepsilon}_{2h}$ in terms of a $N(0,1)$ cdf as,

(A12) $$\Pr\left(\varepsilon_{1h} \le -\ell_{1h} \middle| \hat{\varepsilon}_{2h}\right) = \Phi\left(\frac{-\sigma_{\hat{\varepsilon}_2}\ell_{1h} - (\sigma_{\varepsilon_1\hat{\varepsilon}_2}/\sigma_{\hat{\varepsilon}_2})\hat{\varepsilon}_{2h}}{\sqrt{\Delta}}\right),$$

where now $\Delta = \sigma_1^2\sigma_{\hat{\varepsilon}_2}^2 - \sigma_{\varepsilon_1\hat{\varepsilon}_2}^2 = \sigma_1^2\sigma_2^2 - \sigma_{12}^2$. The change of variables to $\hat{\varepsilon}_{2h} = \sigma_{\hat{\varepsilon}_2} w$, allows us to write the unconditional probability of being in Regime 2 in terms of the univariate integral of a single standard normal random variable,

(A13) $$\Pr\left(y_{1h} = 0, y_{2h} > 0\right) = \Pr\left(\varepsilon_{1h} \le -\ell_{1h}, \varepsilon_{2h} > -\hat{\ell}_{2h}\right) = \int_{-\hat{\ell}_{2h}/\sigma_{\hat{\varepsilon}_2}}^{\infty} \varphi(w)\Phi\left(\frac{-\sigma_{\hat{\varepsilon}_2}\ell_{1h} - \sigma_{\varepsilon_1\hat{\varepsilon}_2} w}{\sqrt{\Delta}}\right)dw.$$

We obtain the conditional mean for y_2 given $y_1 = 0$ and $y_2 > 0$ as,

(A14) $$E\left(y_{2h} \middle| y_{1h} = 0, y_{2h} > 0\right) = \hat{\alpha}_2 + \hat{\boldsymbol{A}}_2\boldsymbol{s}_h + \hat{c}_{22}\ln p_2 + \frac{\sigma_{\hat{\varepsilon}_2}\displaystyle\int_{-\hat{\ell}_{2h}/\sigma_{\hat{\varepsilon}_2}}^{\infty} w\varphi(w)\Phi\left(\frac{-\sigma_{\hat{\varepsilon}_2}\ell_{1h} - \sigma_{\varepsilon_1\hat{\varepsilon}_2} w}{\sqrt{\Delta}}\right)dw}{\displaystyle\int_{-\hat{\ell}_{2h}/\sigma_{\hat{\varepsilon}_2}}^{\infty} \varphi(w)\Phi\left(\frac{-\sigma_{\hat{\varepsilon}_2}\ell_{1h} - \sigma_{\varepsilon_1\hat{\varepsilon}_2} w}{\sqrt{\Delta}}\right)dw}.$$

Regime 3

The derivations and results for this case are much simpler because we do not have to calculate a choke price or a modified structural demand equation for either good. In this case, we can use $\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix} \right)$ directly. We can condition either on ε_{1h} or ε_{2h} to obtain two equivalent expressions for the probability of Regime 3,

(A15)
$$\begin{aligned} \Pr\left(y_{1h} > 0, y_{2h} > 0\right) &= \int_{-\ell_{1h}/\sigma_1}^{\infty} \varphi(w) \Phi\left(\frac{\sigma_1 \ell_{2h} + \sigma_{12} w}{\sqrt{\sigma_1^2 \sigma_2^2 - \sigma_{12}^2}} \right) dw \\ &= \int_{-\ell_{2h}/\sigma_2}^{\infty} \varphi(w) \Phi\left(\frac{\sigma_2 \ell_{1h} + \sigma_{12} w}{\sqrt{\sigma_1^2 \sigma_2^2 - \sigma_{12}^2}} \right) dw. \end{aligned}$$

Exploiting the equivalence of the two expressions to our advantage, we derive the conditional means of y_1 and y_2 given $y_{1h} > 0$ and $y_{2h} > 0$ as,

(A16)
$$E\left(y_{1h} \middle| y_{1h} > 0, y_{2h} > 0\right) = \alpha_1 + \boldsymbol{A}_1 \boldsymbol{s}_h + c_{11} \ln p_1 + c_{12} \ln p_2 + \frac{\sigma_1 \int_{-\ell_{1h}/\sigma_1}^{\infty} w \varphi(w) \Phi\left(\frac{\sigma_1 \ell_{2h} + \sigma_{12} w}{\sqrt{\sigma_1^2 \sigma_2^2 - \sigma_{12}^2}} \right) dw}{\int_{-\ell_{1h}/\sigma_1}^{\infty} \varphi(w) \Phi\left(\frac{\sigma_1 \ell_{2h} + \sigma_{12} w}{\sqrt{\sigma_1^2 \sigma_2^2 - \sigma_{12}^2}} \right) dw},$$

(A17)
$$E\left(y_{2h} \middle| y_{1h} > 0, y_{2h} > 0\right) = \alpha_2 + \boldsymbol{A}_2 \boldsymbol{s}_h + c_{12} \ln p_1 + c_{22} \ln p_2 + \frac{\sigma_2 \int_{-\ell_{2h}/\sigma_2}^{\infty} w \varphi(w) \Phi\left(\frac{\sigma_2 \ell_{1h} + \sigma_{12} w}{\sqrt{\sigma_1^2 \sigma_2^2 - \sigma_{12}^2}} \right) dw}{\int_{-\ell_{2h}/\sigma_2}^{\infty} \varphi(w) \Phi\left(\frac{\sigma_2 \ell_{1h} + \sigma_{12} w}{\sqrt{\sigma_1^2 \sigma_2^2 - \sigma_{12}^2}} \right) dw}.$$