Estimating Demand Systems with Corner Solutions: The Performance of Tobit-Based Approaches

Kyunghoon Ban and Sergio H. Lence
Iowa State University - Department of Economics
khban@iastate.edu and shlence@iastate.edu

Selected Paper prepared for presentation at the 2022 Agricultural & Applied Economics Association Annual Meeting, Anaheim, CA; July 31-August 2

Estimating Demand Systems with Corner Solutions: The Performance of Tobit-Based Approaches

Abstract

Since Tobin's (1958) innovative research on a censored regression equation, a great number of frameworks have been proposed for modeling consumers' preferences in the presence of corner solutions from zero consumptions. However, practitioners have been facing a clear trade-off between flexibility and theoretical plausibility; the virtual price approach (or the Kuhn-Tucker approach) is based on rigorous choice theory but cannot be applied to complex and flexible demand systems because of the need for closed-form solutions to virtual prices, whereas the Tobit-based approach can be applied to any class of demand systems but is deficient in the theoretical foundations on the underlying preferences underlying the observed choices. Hence, we assess the performance of Tobit-based approaches and explore the extent of possible biases in elasticity estimates to provide reasonable criteria for model selection. Our analysis concludes that theoretical restrictions implied by the theory of choice are essential to the Tobit model and improve the ability to capture the true underlying elasticities. Moreover, we also find that the performance of the Tobit models reaches its peak when the datasets with moderate portions of zeros than a relatively lower or higher number of corners. However, even the most restrictive Amemiya-Tobin approach performing better than the other Tobit models has a relative weakness in recovering income elasticities compared to price elasticities under our setup.

Key words: Monte Carlo Simulation, Censored Regression, Almost Ideal Demand System

JEL classification: C34, Q11

1. Introduction

From the time of Tobin's (1958) pioneering work on a regression equation where the dependent variable is censored to be a single value from a certain threshold, a myriad of papers have proposed frameworks for modeling consumers' preferences in the presence of corner solutions, arising from zero consumption(s) for one or more good(s). Based on Tobin's approach, Nelson and Olson (1978) and Amemiya (1979) developed and examined a simultaneous-equation model with both completely observable and truncated equations, which can be applied to a demand system with corner solutions. While the earlier frameworks focused on truncated consumption in observed data, the literature started to develop estimation structures more consistent with choice theory. Wales and Woodland (1983) not only formalized the usage of first-order Kuhn-Tucker conditions for constructing the likelihood function from a utility maximization problem with non-negativity constraints on consumption, but also proposed a specific mapping technique called the Amemiya-Tobin approach for the Tobit model, so that the model can be consistent with the adding-up condition. Shortly after, Lee and Pitt (1986) developed a virtual price approach, which is dual to the Kuhn-Tucker approach but does not require beginning with the primary utility, allowing for a more general class of demand systems possibly stemming from indirect utility functions (such as the Translog indirect utility function).

A review of the literature suggests that practitioners have been facing a clear trade-off between flexibility and theoretical plausibility. That is, the virtual price approach (or the Kuhn-Tucker approach) is rigorously based on the theory of choice, but it cannot be applied to complex and flexible demand systems because of the need for closed-form solutions to virtual prices, whereas the Tobit-based approach can be applied to any class of demand systems as long as they have share equations to estimate, but is deficient in the theoretical foundations of the underlying preferences behind the observed optimal choices. For example, the Quadratic Almost Ideal (QUAI) demand system (Banks, Blundell, and Lewbel 1997) and the Exact Affine Stone Index (EASI) demand system (Lewbel and Pendakur 2009) provide a flexible form of demand system in terms of the Engel curve, but the virtual price approach cannot be used with them because the share equations are highly nonlinear in prices (or the logarithm of prices). In contrast, despite the choice theory deficiency of the Tobit approach, the latter can be easily adopted for such demand systems by simply adding normal error terms to the share equations (e.g., Zhen et al. (2013)).

Given the aforementioned state of affairs, our paper aims to assess the performance of Tobit-based approaches proposed in the literature and explore the extent of possible biases in elasticity estimates to provide reasonable criteria for model selection. We implement a series of Monte Carlo simulations to evaluate three different Tobit-based frameworks using the Linear Approximate Almost Ideal (LA-AI) demand system, namely, (a) simple equation-by-equation Tobit, (b) Tobit with a system of correlated equations, and (c) Amemiya-Tobin with the adding-up condition. The LA-AI demand system is chosen despite its well-documented problems, because not only it has been one of the most popular demand systems for the past few decades, but also it is simple enough to generate optimal consumption vectors with potential corner solutions utilizing the virtual price technique. Specifically, we first generate 1,000 samples, each consisting of 500 heterogeneous individuals who have different preferences over three goods in terms of the (indirect) utility parameters, where the heterogeneity parameters are drawn from the multivariate normal distribution. Then, based on the virtual prices derived from the LA-AI demand system, we generate optimal consumption vectors with non-negativity consumption constraints. Next, we evaluate the performance of the three aforementioned Tobit-based approaches by estimating them and computing the elasticities of interest from each model's parameter estimates. Given the heavy computational requirements of the proposed analysis, we adopt Geweke's (1988) antithetic sampling method to accelerate the simulations.

This exercise provides the following contributions. First, because we generate the hypothetical consumption data based on consumer choice theory, we can properly examine the performance of the Tobit-based approaches when they are applied to real-world datasets with rational consumers. Note that we minimize the other possibilities of biases that might arise from the misspecification of the functional form of the demand system, by matching the functional specification in the data-generating process (DGP) to that in the estimation step. Hence, either the Tobit-based approaches could be justified in empirical applications, or they should be reconsidered depending on the simulation results; the linearly additive error structure or the lack of theory foundation could negatively affect the latter's performance. Secondly, we can evaluate the impact of additional theoretical constraints within the Tobit-based methods. Because the Tobit methods analyzed here are nested within each other in terms of the theoretical constraints (with the simple equation-by-equation Tobit approach being the least restrictive, and the Amemiya-Tobin approach the most restrictive), it is interesting to investigate how each of the frameworks performs in capturing the true elasticity or the censoring thresholds for goods implied by the true DGP. In particular, given the relative complexity of implementing the Amemiya-Tobin method even with the simplification provided by Dong, Davis, and Stewart (2014), the results of our study should also provide useful information regarding the selection of Tobit-based techniques.

2. AI Demand System

Let $i \in \{1, \cdots, M\}$ be the index for M goods. For a given price vector $\mathbf{p} = (p_1, \cdots, p_M)'$ and income m, the optimal consumption share w_i^* for $i \in \{1, \cdots, M\}$ is equal to

$$w_i^* = \alpha_i + \sum_j \beta_{ij} \ln p_j + \gamma_i \ln\left(\frac{m}{g(\mathbf{p})}\right) + \varepsilon_i, \tag{1}$$

where

$$\ln g(\mathbf{p}) = \sum_j w_j^* \ln p_j,$$

and $\beta_{ij} = \beta_{ji}$, $\sum_i \beta_{ij} = 0$ for all j, $\sum_i \alpha_i = 1$, $\sum_i \gamma_i = 0$, and $\sum_i \varepsilon_i = 0$. Since optimal share w_i^* appears on both sides of the equality in (1), an issue that arises regarding the data generating process (DGP) is *the proper way to generate data using the linearized AI demand system.* In the next subsection, we first discuss the DGP for the typical case without corner solutions. Then, in the following subsection, we focus on the DGP for the case of interest to us, namely, the presence of corner solutions.

2.1. DGP Without Corners

The share equations (1) with $i = 1, \cdots, M$ can be represented in matrix form:

$$\mathbf{w}^* = \boldsymbol{\alpha} + \mathbf{B}\overline{\mathbf{p}} + \boldsymbol{\gamma}[(\ln m) - \overline{\mathbf{p}}'\mathbf{w}^*] + \boldsymbol{\varepsilon}, \tag{2}$$

where $\mathbf{w}^* = (w_1^*, \cdots, w_M^*)$, $\boldsymbol{\alpha} = (\alpha_1, \cdots, \alpha_M)$, $\boldsymbol{\gamma} = (\gamma_1, \cdots, \gamma_M)$, $\boldsymbol{\varepsilon} = (\varepsilon_1, \cdots, \varepsilon_M)$, $\overline{\mathbf{p}} = (\ln p_1, \cdots, \ln p_M)$, and

$$\mathbf{B} = \begin{bmatrix} \beta_{11} & \beta_{12} & \cdots & \beta_{1M} \\ \beta_{21} & \beta_{22} & \cdots & \beta_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{M1} & \beta_{M2} & \cdots & \beta_{MM} \end{bmatrix}.$$

The right hand side (RHS) of (2) can be expressed as

$$\begin{aligned} \mathbf{w}^* &= \boldsymbol{\alpha} + \mathbf{B}\overline{\mathbf{p}} + \boldsymbol{\gamma}[(\ln m) - \overline{\mathbf{p}}'\mathbf{w}^*] + \boldsymbol{\varepsilon}, \\ &= \boldsymbol{\alpha} + \mathbf{B}\overline{\mathbf{p}} + (\ln m)\boldsymbol{\gamma} - \boldsymbol{\gamma}\overline{\mathbf{p}}'\mathbf{w}^* + \boldsymbol{\varepsilon}. \end{aligned}$$

Thus, we have

$$\mathbf{w}^* = [I_M + \boldsymbol{\gamma}\overline{\mathbf{p}}']^{-1}[\boldsymbol{\alpha} + \mathbf{B}\overline{\mathbf{p}} + (\ln m)\boldsymbol{\gamma} + \boldsymbol{\varepsilon}], \tag{3}$$

where I_M is an $M \times M$ identity matrix. Vector $\mathbf{w}^*$ is not recursively defined anymore in (3), and henceforth we call it a *complete closed-from solution* for $\mathbf{w}^*$. Using this complete closed-from of $\mathbf{w}^*$, we can generate optimal consumption data for a given $(\mathbf{p}, m)$ and the parameter values.

2.2. DGP With Corners

The virtual price approach is dual to the Kuhn Tucker method. Hence, we can find the optimal consumption vector *if there are non-negative constraints for* $\mathbf{w}^*$ *for sufficiently small* m, without solving the corresponding utility maximization problem (UMP). Recall that (3) does not restrict $\mathbf{w}^*$ to be non-negative, and w_i^* for some i can be negative for some realization of $\boldsymbol{\varepsilon}$; hence, $\mathbf{w}^*$ can be thought as comprising latent shares.

To consider the virtual price approach, without loss of generality, let $\{w_1^*, \cdots, w_K^*\}$ be a set of positive (latent) shares and $\{w_{K+1}^*, \cdots, w_M^*\}$ be a set of non-positive shares. Then, the virtual prices $\{\pi_{K+1}, \cdots, \pi_M\}$ and the constrained (observed) shares $\{w_1, \cdots, w_K\}$ can be jointly defined by letting $w_i^* = 0$ for $i = K+1, \cdots, M$ in (1): [1]

$$w_i = \alpha_i + \sum_{j=1}^{K} \beta_{ij} \ln p_j + \sum_{j=K+1}^{M} \beta_{ij} \ln \pi_j + \gamma_i \left(\ln m - \sum_{j=1}^{K} w_j \ln p_j \right) + \varepsilon_i, \tag{4}$$

for $i = 1, \cdots, K$, and

$$0 = \alpha_i + \sum_{j=1}^{K} \beta_{ij} \ln p_j + \sum_{j=K+1}^{M} \beta_{ij} \ln \pi_j + \gamma_i \left(\ln m - \sum_{j=1}^{K} w_j \ln p_j \right) + \varepsilon_i, \tag{5}$$

for $i = K+1, \cdots, M$. Note that (4) and (5) can be represented in matrix notation as:

$$\begin{aligned} \mathbf{w_0} &= \boldsymbol{\alpha}_\mathbf{0} + \mathbf{B_{00}}\overline{\mathbf{p}}_\mathbf{0} + \mathbf{B_{01}}\overline{\pi}_\mathbf{1} + (\ln m)\boldsymbol{\gamma}_\mathbf{0} - \boldsymbol{\gamma}_\mathbf{0}\overline{\mathbf{p}}_\mathbf{0}{}'\mathbf{w_0} + \boldsymbol{\varepsilon}_\mathbf{0}, \\ \mathbf{0} &= \boldsymbol{\alpha}_\mathbf{1} + \mathbf{B_{10}}\overline{\mathbf{p}}_\mathbf{0} + \mathbf{B_{11}}\overline{\pi}_\mathbf{1} + (\ln m)\boldsymbol{\gamma}_\mathbf{1} - \boldsymbol{\gamma}_\mathbf{1}\overline{\mathbf{p}}_\mathbf{0}{}'\mathbf{w_0} + \boldsymbol{\varepsilon}_\mathbf{1}, \end{aligned}$$

where the subscripts 0 and 1 represent the corresponding partition of first K rows/columns and the other $(M-K)$ rows/columns, respectively, and $\overline{\pi}_\mathbf{1} = (\ln \pi_{K+1}, \cdots, \ln \pi_M)$. Manipulating them as

$$\begin{aligned} \left(I_K + \boldsymbol{\gamma}_\mathbf{0}\overline{\mathbf{p}}_\mathbf{0}{}'\right)\mathbf{w_0} &= \boldsymbol{\alpha}_\mathbf{0} + \mathbf{B_{00}}\overline{\mathbf{p}}_\mathbf{0} + \mathbf{B_{01}}\overline{\pi}_\mathbf{1} + (\ln m)\boldsymbol{\gamma}_\mathbf{0} + \boldsymbol{\varepsilon}_\mathbf{0}, \\ \mathbf{B_{11}}\overline{\pi}_\mathbf{1} &= -\left[\boldsymbol{\alpha}_\mathbf{1} + \mathbf{B_{10}}\overline{\mathbf{p}}_\mathbf{0} + (\ln m)\boldsymbol{\gamma}_\mathbf{1} - \boldsymbol{\gamma}_\mathbf{1}\overline{\mathbf{p}}_\mathbf{0}{}'\mathbf{w_0} + \boldsymbol{\varepsilon}_\mathbf{1}\right], \end{aligned}$$

we can obtain

$$(6)\qquad \mathbf{w_0} = \left(I_K + \boldsymbol{\gamma}_0 \overline{\mathbf{p}}_0{}'\right)^{-1} \left[\boldsymbol{\alpha}_0 + \mathbf{B}_{00}\overline{\mathbf{p}}_0 + \mathbf{B}_{01}\overline{\pi}_1 + (\ln m)\boldsymbol{\gamma}_0 + \boldsymbol{\varepsilon}_0\right],$$

$$(7)\qquad \overline{\pi}_1 = -\mathbf{B}_{11}{}^{-1}\left[\boldsymbol{\alpha}_1 + \mathbf{B}_{10}\overline{\mathbf{p}}_0 + (\ln m)\boldsymbol{\gamma}_1 - \boldsymbol{\gamma}_1\overline{\mathbf{p}}_0{}'\mathbf{w}_0 + \boldsymbol{\varepsilon}_1\right].$$

We can now substitute each of Equations (6) and (7) into the other to derive complete closed-form solutions.

Firstly, substituting (7) into (6), we have

$$\begin{aligned}
\mathbf{w}_0 &= \left(I_K + \boldsymbol{\gamma}_0 \overline{\mathbf{p}}_0{}'\right)^{-1} \Big\{\boldsymbol{\alpha}_0 + \mathbf{B}_{00}\overline{\mathbf{p}}_0 - \mathbf{B}_{01}\mathbf{B}_{11}{}^{-1}\big[\boldsymbol{\alpha}_1 + \mathbf{B}_{10}\overline{\mathbf{p}}_0 \\
&\qquad + (\ln m)\boldsymbol{\gamma}_1 - \boldsymbol{\gamma}_1\overline{\mathbf{p}}_0{}'\mathbf{w}_0 + \boldsymbol{\varepsilon}_1\big] + (\ln m)\boldsymbol{\gamma}_0 + \boldsymbol{\varepsilon}_0\Big\}, \\
&= \left(I_K + \boldsymbol{\gamma}_0 \overline{\mathbf{p}}_0{}'\right)^{-1} \Big\{\boldsymbol{\alpha}_0 + \mathbf{B}_{00}\overline{\mathbf{p}}_0 - \mathbf{B}_{01}\mathbf{B}_{11}{}^{-1}\big[\boldsymbol{\alpha}_1 + \mathbf{B}_{10}\overline{\mathbf{p}}_0 \\
&\qquad + (\ln m)\boldsymbol{\gamma}_1 + \boldsymbol{\varepsilon}_1\big] + (\ln m)\boldsymbol{\gamma}_0 + \boldsymbol{\varepsilon}_0\Big\} \\
&\qquad + \left(I_K + \boldsymbol{\gamma}_0 \overline{\mathbf{p}}_0{}'\right)^{-1}\mathbf{B}_{01}\mathbf{B}_{11}{}^{-1}\boldsymbol{\gamma}_1\overline{\mathbf{p}}_0{}'\mathbf{w}_0.
\end{aligned}$$

Hence, we have

$$\begin{aligned}
&\left[I_K - \left(I_K + \boldsymbol{\gamma}_0 \overline{\mathbf{p}}_0{}'\right)^{-1}\mathbf{B}_{01}\mathbf{B}_{11}{}^{-1}\boldsymbol{\gamma}_1\overline{\mathbf{p}}_0{}'\right]\mathbf{w}_0 \\
&\quad = \left(I_K + \boldsymbol{\gamma}_0 \overline{\mathbf{p}}_0{}'\right)^{-1} \Big\{\boldsymbol{\alpha}_0 + \mathbf{B}_{00}\overline{\mathbf{p}}_0 - \mathbf{B}_{01}\mathbf{B}_{11}{}^{-1}\big[\boldsymbol{\alpha}_1 + \mathbf{B}_{10}\overline{\mathbf{p}}_0 \\
&\qquad\qquad + (\ln m)\boldsymbol{\gamma}_1 + \boldsymbol{\varepsilon}_1\big] + (\ln m)\boldsymbol{\gamma}_0 + \boldsymbol{\varepsilon}_0\Big\},
\end{aligned}$$

or

$$\begin{aligned}
(8)\quad \mathbf{w}_0 &= \left[I_K - \left(I_K + \boldsymbol{\gamma}_0 \overline{\mathbf{p}}_0{}'\right)^{-1}\mathbf{B}_{01}\mathbf{B}_{11}{}^{-1}\boldsymbol{\gamma}_1\overline{\mathbf{p}}_0{}'\right]^{-1}\left(I_K + \boldsymbol{\gamma}_0 \overline{\mathbf{p}}_0{}'\right)^{-1}\Big\{\boldsymbol{\alpha}_0 \\
&\qquad + \mathbf{B}_{00}\overline{\mathbf{p}}_0 - \mathbf{B}_{01}\mathbf{B}_{11}{}^{-1}\big[\boldsymbol{\alpha}_1 + \mathbf{B}_{10}\overline{\mathbf{p}}_0 + (\ln m)\boldsymbol{\gamma}_1 + \boldsymbol{\varepsilon}_1\big] + (\ln m)\boldsymbol{\gamma}_0 + \boldsymbol{\varepsilon}_0\Big\}, \\
&= \left(I_K + \boldsymbol{\gamma}_0 \overline{\mathbf{p}}_0{}' - \mathbf{B}_{01}\mathbf{B}_{11}{}^{-1}\boldsymbol{\gamma}_1\overline{\mathbf{p}}_0{}'\right)^{-1}\Big\{\boldsymbol{\alpha}_0 + \mathbf{B}_{00}\overline{\mathbf{p}}_0 \\
&\qquad - \mathbf{B}_{01}\mathbf{B}_{11}{}^{-1}\big[\boldsymbol{\alpha}_1 + \mathbf{B}_{10}\overline{\mathbf{p}}_0 + (\ln m)\boldsymbol{\gamma}_1 + \boldsymbol{\varepsilon}_1\big] + (\ln m)\boldsymbol{\gamma}_0 + \boldsymbol{\varepsilon}_0\Big\}.
\end{aligned}$$

Note that $\mathbf{w_0}$ is linear in the error terms $\boldsymbol{\varepsilon}$.

On the other hand, substituting (6) into (7), we have

$$\begin{aligned}
\overline{\pi}_{\mathbf{1}} &= -\mathbf{B_{11}}^{-1}\Big\{\boldsymbol{\alpha}_1+\mathbf{B_{10}}\overline{\mathbf{p}}_{\mathbf{0}}+(\ln m)\boldsymbol{\gamma}_{\mathbf{1}}-\boldsymbol{\gamma}_1\overline{\mathbf{p}}_{\mathbf{0}}{}'\left(I_K+\boldsymbol{\gamma}_0\overline{\mathbf{p}}_{\mathbf{0}}{}'\right)^{-1} \\
&\qquad [\boldsymbol{\alpha}_{\mathbf{0}}+\mathbf{B_{00}}\overline{\mathbf{p}}_{\mathbf{0}}+\mathbf{B_{01}}\overline{\pi}_{\mathbf{1}}+(\ln m)\boldsymbol{\gamma}_{\mathbf{0}}+\boldsymbol{\varepsilon}_{\mathbf{0}}]+\boldsymbol{\varepsilon}_{\mathbf{1}}\Big\}, \\
&= -\mathbf{B_{11}}^{-1}\Big\{\boldsymbol{\alpha}_1+\mathbf{B_{10}}\overline{\mathbf{p}}_{\mathbf{0}}+(\ln m)\boldsymbol{\gamma}_{\mathbf{1}}-\boldsymbol{\gamma}_1\overline{\mathbf{p}}_{\mathbf{0}}{}'\left(I_K+\boldsymbol{\gamma}_0\overline{\mathbf{p}}_{\mathbf{0}}{}'\right)^{-1} \\
&\qquad [\boldsymbol{\alpha}_{\mathbf{0}}+\mathbf{B_{00}}\overline{\mathbf{p}}_{\mathbf{0}}+(\ln m)\boldsymbol{\gamma}_{\mathbf{0}}+\boldsymbol{\varepsilon}_{\mathbf{0}}]+\boldsymbol{\varepsilon}_{\mathbf{1}}\Big\} \\
&\qquad +\mathbf{B_{11}}^{-1}\boldsymbol{\gamma}_1\overline{\mathbf{p}}_{\mathbf{0}}{}'\left(I_K+\boldsymbol{\gamma}_0\overline{\mathbf{p}}_{\mathbf{0}}{}'\right)^{-1}\mathbf{B_{01}}\overline{\pi}_1.
\end{aligned}$$

Thus, we have

$$\begin{aligned}
&\left[I_{M-K}-\mathbf{B_{11}}^{-1}\boldsymbol{\gamma}_1\overline{\mathbf{p}}_{\mathbf{0}}{}'\left(I_K+\boldsymbol{\gamma}_0\overline{\mathbf{p}}_{\mathbf{0}}{}'\right)^{-1}\mathbf{B_{01}}\right]\overline{\pi}_{\mathbf{1}} \\
&\quad = -\mathbf{B_{11}}^{-1}\Big\{\boldsymbol{\alpha}_1+\mathbf{B_{10}}\overline{\mathbf{p}}_{\mathbf{0}}+(\ln m)\boldsymbol{\gamma}_{\mathbf{1}}-\boldsymbol{\gamma}_1\overline{\mathbf{p}}_{\mathbf{0}}{}'\left(I_K+\boldsymbol{\gamma}_0\overline{\mathbf{p}}_{\mathbf{0}}{}'\right)^{-1} \\
&\qquad\qquad [\boldsymbol{\alpha}_{\mathbf{0}}+\mathbf{B_{00}}\overline{\mathbf{p}}_{\mathbf{0}}+(\ln m)\boldsymbol{\gamma}_{\mathbf{0}}+\boldsymbol{\varepsilon}_{\mathbf{0}}]+\boldsymbol{\varepsilon}_{\mathbf{1}}\Big\},
\end{aligned}$$

or

$$\begin{aligned}
(9)\qquad \overline{\pi}_{\mathbf{1}} &= -\left[I_{M-K}-\mathbf{B_{11}}^{-1}\boldsymbol{\gamma}_1\overline{\mathbf{p}}_{\mathbf{0}}{}'\left(I_K+\boldsymbol{\gamma}_0\overline{\mathbf{p}}_{\mathbf{0}}{}'\right)^{-1}\mathbf{B_{01}}\right]^{-1} \\
&\qquad \mathbf{B_{11}}^{-1}\Big\{\boldsymbol{\alpha}_1+\mathbf{B_{10}}\overline{\mathbf{p}}_{\mathbf{0}}+(\ln m)\boldsymbol{\gamma}_{\mathbf{1}}-\boldsymbol{\gamma}_1\overline{\mathbf{p}}_{\mathbf{0}}{}'\left(I_K+\boldsymbol{\gamma}_0\overline{\mathbf{p}}_{\mathbf{0}}{}'\right)^{-1} \\
&\qquad [\boldsymbol{\alpha}_{\mathbf{0}}+\mathbf{B_{00}}\overline{\mathbf{p}}_{\mathbf{0}}+(\ln m)\boldsymbol{\gamma}_{\mathbf{0}}+\boldsymbol{\varepsilon}_{\mathbf{0}}]+\boldsymbol{\varepsilon}_{\mathbf{1}}\Big\}, \\
&= -\left[\mathbf{B_{11}}-\boldsymbol{\gamma}_1\overline{\mathbf{p}}_{\mathbf{0}}{}'\left(I_K+\boldsymbol{\gamma}_0\overline{\mathbf{p}}_{\mathbf{0}}{}'\right)^{-1}\mathbf{B_{01}}\right]^{-1} \\
&\qquad \Big\{\boldsymbol{\alpha}_1+\mathbf{B_{10}}\overline{\mathbf{p}}_{\mathbf{0}}+(\ln m)\boldsymbol{\gamma}_{\mathbf{1}}-\boldsymbol{\gamma}_1\overline{\mathbf{p}}_{\mathbf{0}}{}'\left(I_K+\boldsymbol{\gamma}_0\overline{\mathbf{p}}_{\mathbf{0}}{}'\right)^{-1} \\
&\qquad [\boldsymbol{\alpha}_{\mathbf{0}}+\mathbf{B_{00}}\overline{\mathbf{p}}_{\mathbf{0}}+(\ln m)\boldsymbol{\gamma}_{\mathbf{0}}+\boldsymbol{\varepsilon}_{\mathbf{0}}]+\boldsymbol{\varepsilon}_{\mathbf{1}}\Big\}.
\end{aligned}$$

Note that $\overline{\pi}_{\mathbf{1}}$ is also linear in the error terms $\boldsymbol{\varepsilon}$. [2]

Based on the preceding derivations, given a vector of prices and income for an individual $(\boldsymbol{p}, m)$, we can generate the constrained optimal shares $\tilde{\boldsymbol{w}}$ for that individual as follows:

1. Draw $\boldsymbol{\varepsilon}$ from the multivariate normal distribution with a mean $\mathbf{0}_M$ and a positive semi-definite variance-covariance matrix Σ with $\boldsymbol{\iota}_M'\Sigma = \mathbf{0}_M$.[3]

2. Obtain the latent shares from (3).

3. Rearrange the index so that we have positive latent shares w_i^* for $i = 1, \cdots, K$ and negative for $i = K+1, \cdots, M$.

4. Obtain the constrained consumption vector w_i for $i = 1, \cdots, K$ using (8).

5. Denote the final consumption vector by $\tilde{\boldsymbol{w}}$ obtained from $(w_1, \cdots, w_K)$ together with zeros by rearranging the index back.

Steps 2 - 5 may have to be repetated until the non-negativity constraint is satisfied for all goods, depending on the curvature of the (indirect) utility function.

3. Parameter Values and Sample Distribution

3.1. Benchmark Setup

We consider $M = 3$ as the simplest possible case to minimize further difficulties in numerical optimization, so that we can focus on the differences in performance across econometric models. Regarding the number of observations for each sample and the proportion of zero consumption, we refer to eight empirical research papers that use either Tobit-based models or the virtual price (or the Kuhn-Tucker) approach for the constrained demand system estimation.[4] In these papers, the number of observations divided by the number of parameters varies from 9.2 to 168.1 with the mean of 59.7, and the proportion of zero consumption observations varies from 13% to 65% with the mean of 41%. Considering the fact that the number of parameters to estimate is 10 in our DGP with $M = 3$, we generate $N = 500$ individuals for each sample targeting around 40% of zero corner solutions for our benchmark setup. We label the benchmark setup as (**N500**, **C40**) henceforth in reference to the $N = 500$ individuals per sample and about 40% of zero corner solutions.

The proportion of zero consumption in a sample depends on the curvature of the (indirect) utility function and the distribution of the individual characteristics; income, prices of goods, and the error terms $(m, \boldsymbol{p}, \boldsymbol{\varepsilon})$. Thus, the following parameter values are chosen with the multivariate normality assumption on $\boldsymbol{\varepsilon}$:

$$
\begin{aligned}
(\alpha_1, \alpha_2) &= (0.5, 0.3) \\
(\beta_{11}, \beta_{21}, \beta_{22}) &= (-0.5, 0.3, -0.2) \\
(\gamma_1, \gamma_2) &= (-0.3, 0.2) \\
(\Sigma_{11}, \Sigma_{21}, \Sigma_{22}) &= \left(0.35^2, -0.4 \cdot (0.35 \cdot 0.45), 0.45^2\right),
\end{aligned}
$$

Moreover, the log-normal distribution is assumed for $(\boldsymbol{p}, m)$, where $\boldsymbol{p}$ has means and variances of 1 and m has a unit mean and a variance of 0.5^2.[5] Finally, we generate $B = 1,000$ samples that constitute a set of simulations. Figure 1 shows the generated sample distributions for each variable.

3.2. Other Setups for Sensitivity Analysis

We consider two additional dimensions to evaluate the performances of the econometric models more thoroughly: the number of observations (N) and the proportions of zero corner solutions. First, we select $N = 100$ (**N100**, **C40**) and $N = 2,000$ (**N2000**, **C40**) based on the literature cited in the previous subsection, and run the corresponding sets of simulations by fixing the other parameter values and sample distribution as the benchmark setup. Second, to analyze how performance is impacted by the proportion of zero observations, we consider about 10% (**N500**, **C10**) and 70% (**N500**, **C70**) of zero corner solutions by fixing the number of observations with $N = 500$.[6]

In order to obtain samples of 10% or 70% of zero corner solutions, we adjust the sample distribution of the prices $\boldsymbol{p}$ and the error terms $\boldsymbol{\varepsilon}$. Particularly, to obtain samples with 10% of zero observed consumption data points, we specify the variances of the prices and the error terms to be $(0.5, 0.5, 0.5)$ and $(0.1, 0.2)$,[7] respectively. Likewise, to obtain samples with 70% corners, we set the variances of the prices and the error terms to be $(2, 2, 2)$ and

$(0.9, 1)$, respectively. The following Figure 2 shows how the distribution of p_1 or u_1 differs according to the setups **C10**, **C40**, and **C70**.

4. Estimation

We consider three different Tobit-based econometric models to estimate the postulated LA-AI demand system.

1. **T1**: Simple equation-by-equation Tobit
2. **T2**: Tobit with a system of correlated equations
3. **T3**: Amemiya-Tobin with the adding-up condition

The Tobit methods are nested in terms of the theoretical constraints; **T1** is the least restrictive and **T3** is the most restrictive.

The estimation of the Tobit methods is implemented by the maximum likelihood method. As our sample consists of independently and identically drawn individuals ($n = 1, \cdots, N$), we can rewrite the log-likelihood function as a sum of each individual's log-likelihood function:

$$\begin{aligned} l(\boldsymbol{\theta}; \boldsymbol{x}_1, \cdots, \boldsymbol{x}_N) &\equiv \ln \mathscr{L}(\boldsymbol{\theta}; \boldsymbol{x}_1, \cdots, \boldsymbol{x}_N), \\ &= \sum_{n=1}^{N} \ln L(\boldsymbol{\theta}; \boldsymbol{x}_n), \end{aligned}$$

where $\boldsymbol{\theta}$ is a vector of unknown parameters, $\boldsymbol{x}_n \equiv (\tilde{\boldsymbol{w}}, \boldsymbol{p}, m)$ is a vector of observed data for individual n.

For each individual, there are $2^M - 1$ possible combinations of corner solutions for the optimal consumption shares.[8] Without loss of generality, we derive the results for the following general cases by re-arranging the index as $\tilde{\boldsymbol{w}} = (\tilde{w}_1, \cdots, \tilde{w}_K, \mathbf{0}_{M-K})$.

4.1. T1: Simple equation-by-equation Tobit

The econometric model for **T1** consists of

$$\tilde{w}_i = \max\{w_i^p + \varepsilon_i, 0\}, \quad i = 1, \cdots, M, \tag{10}$$

where the predicted share w_i^p with the observed-share Stone price index $\overline{\mathbf{p}}'\tilde{\mathbf{w}}$ is defined as

$$\mathbf{w}^{\mathbf{p}}(\boldsymbol{\theta}, \boldsymbol{x}_n) \equiv \boldsymbol{\alpha} + \mathbf{B}\overline{\mathbf{p}} + \boldsymbol{\gamma}[(\ln m) - \overline{\mathbf{p}}'\tilde{\mathbf{w}}],$$

and $\varepsilon_i \sim N(0, \sigma_i^2)$ for $i = 1, \cdots, M$. Note that the error terms do not sum up to 0, and they are not correlated across the share equations. Therefore, the log-likelihood function for an individual n can be obtained as

$$\begin{aligned}
\ln L(\boldsymbol{\theta}; \boldsymbol{x}_n) &= \sum_{i=1}^{M} \ln \left[\phi(\tilde{w}_i - w_i^p; 0, \sigma_i^2) \right]^{\mathbf{1}\{\tilde{w}_i > 0\}} \left[\int_{-\infty}^{-w_i^p} \phi(\varepsilon; 0, \sigma_i^2) \mathrm{d}\varepsilon \right]^{\mathbf{1}\{\tilde{w}_i = 0\}} \\
&= \sum_{i=1}^{K} \ln \left[\phi(\tilde{w}_i - w_i^p; 0, \sigma_i^2) \right] + \sum_{i=K+1}^{M} \ln \left[\int_{-\infty}^{-w_i^p} \phi(\varepsilon; 0, \sigma_i^2) \mathrm{d}\varepsilon \right],
\end{aligned}$$

where $\mathbf{1}\{x\}$ is the indicator function that takes 1 if the expression x is true and 0 otherwise.

4.2. T2: Tobit with a system of correlated equations

Similar to **T1**, we still maintain Equation (10), but now the error terms are assumed to follow a multivariate normal distribution. Particularly, we have $\boldsymbol{\varepsilon} \sim \boldsymbol{N}(\mathbf{0}_M, \Sigma)$ where Σ is a positive semi-definite variance-covariance matrix with $\boldsymbol{\iota}_M'\Sigma = \mathbf{0}_M$, which ensures that the latent shares $\mathbf{w}^{\mathbf{p}} + \boldsymbol{\varepsilon}$ sum up to 1. Hence, for the re-indexed shares $\tilde{\boldsymbol{w}} = (\tilde{w}_1, \cdots, \tilde{w}_K, \mathbf{0}_{M-K})$, the individual n's log-likelihood function can be obtained as

$$\ln L(\boldsymbol{\theta}; \boldsymbol{x}_n) = \ln \int_{-\infty}^{-w_{K+1}^p} \cdots \int_{-\infty}^{-w_M^p} \boldsymbol{\phi}\left((\tilde{\boldsymbol{w}}_{\mathbf{0}} - \boldsymbol{w}_{\mathbf{0}}^p, \varepsilon_{K+1}, \cdots, \varepsilon_M); \mathbf{0}_{M-1}, \Sigma_{-1,-1} \right) \mathrm{d}\varepsilon_M, \cdots, \mathrm{d}\varepsilon_{K+1}$$

Note that the formula does not require an integration if $K = M$:

$$\ln L(\boldsymbol{\theta}; \boldsymbol{x}_n) = \ln \boldsymbol{\phi}(\tilde{\boldsymbol{w}}_{-1} - \boldsymbol{w}_{-1}^p; \mathbf{0}_{M-1}, \Sigma_{-1,-1})$$

4.3. T3: Amemiya-Tobin with the adding-up condition

The adding-up restriction on **T2** only ensures that the *latent shares* ($\mathbf{w^p} + \boldsymbol{\varepsilon}$) sum up to 1. However, the predicted shares after the estimation procedure need not sum up to 1 in general.[9] In contrast, the Amemiya-Tobin approach guarantees that the observed shares sum up to 1, by augmenting the model with an additional mapping structure between the latent shares and the observed shares (Wales and Woodland 1983):[10]

$$\tilde{w}_i = \begin{cases} \frac{w_i^p + \varepsilon_i}{\sum_{j=1}^K w_j^p + \varepsilon_j} & \text{if } w_i^p + \varepsilon_i > 0 \quad (i = 1, \cdots, K) \\ 0 & \text{if } w_i^p + \varepsilon_i \leq 0 \quad (i = K+1, \cdots, M) \end{cases},$$

where $\boldsymbol{\varepsilon} \sim \boldsymbol{N}(\mathbf{0}_M, \Sigma)$ in which Σ is a positive semi-definite variance-covariance matrix with $\boldsymbol{\iota}_M' \Sigma = \mathbf{0}_M$.

It has to be noted that historically **T3** involved much higher computational burden than **T1** or **T2**, because such a mapping requires non-trivial integration over the multivariate normal density function for the likelihood function evaluation. However, Dong, Gould, and Kaiser (2004) proposed a novel approach that only requires a relatively simple integration over a rectangular region, which enables us to use the GHK (Geweke-Hajivassiliou-Keane) algorithm (Geweke 1991; Hajivassiliou and McFadden 1991; Keane 1994). We transplant the `Gauss` code provided by Dong, Davis, and Stewart (2014), where the same technique is applied, to `R`, but the detailed derivation of the likelihood function can be found in Dong, Gould, and Kaiser (2004).

5. Results

5.1. Parameter Estimates

Figures 3 and 4 show the parameter estimates across $B = 1{,}000$ samples for each econometric model. Each row represents **T1**, **T2**, and **T3**, respectively, and each column shows a different parameter. For instance, the histogram at the first column and the first row of Figure 3 shows the distribution of α_1 estimates for **T1**. For each panel, red lines represent the true parameter value from the DGP, and blue dotted lines represent the mean of all the estimates across samples. ‘Rej.’ on the left-upper corner of each panel reports the rejection rate with respect to the true parameter value under a 5% significance level obtained from the variance estimates of parameters for each sample. Note that **T3** performs the best in estimating (α_1, α_2), whereas **T2** shows the best performance for the estimation of other parameters $(\beta_{11}, \beta_{12}, \beta_{22})$ and (γ_1, γ_2).

To assess the performances of the models more rigorously, we evaluate the normalized root mean square error (*RMSE*) or the normalized RMSE (*NRMSE*):

$$\begin{aligned} RMSE &\equiv \sqrt{\frac{\sum_{b=1}^{B}(\hat{\theta}_b - \theta)^2}{B}} \\ NRMSE &\equiv \frac{RMSE}{\theta}, \end{aligned}$$

where the normalization is implemented by means of the true parameter value. We use *NRMSE* for parameter estimates and *RMSE* for elasticity estimates. We evaluate *NRMSE* for each parameter and each model (**T1**, **T2**, and **T3**), and we take the mean of *NRMSE* across all parameters for each model to evaluate the performances.

Tables 1 and 2 summarize the mean NRMSE obtained from the parameter estimates across different numbers of observations and corner proportions, respectively. The second row of Table 1 or 2 summarizes Figures 3 and 4 in three numbers; **T2** performs the best in recovering the true parameter values in the benchmark setup, followed by **T3**. From Table 1, it is clear that the number of observations helps to recover the true parameter values, as we observe a slight improvement in the mean RMSE values along with the larger number of observations. However, the performance ranking remains the same regardless of the number of observations. Contrastingly, when we have different proportions of corner solutions in the dataset (see Table 2), the hierarchy changes when we have a larger proportion of corner solutions. Although it seems that **T2** does a relatively good job in recovering the parameters and **T1** is always outperformed by **T2** or **T3**, this does not necessarily mean that **T2** would perform the best and **T1** would perform the worst in empirical works because each model assumes different error structures and different mechanisms for predicting shares. Hence, we analyze how each model performs in terms of capturing elasticities in the next subsection.

5.2. Elasticity Simulation

As pointed out in Dong, Gould, and Kaiser (2004), it is analytically demanding to derive elasticity values for a class of demand systems with corner solutions. Hence, we adopt a simulation-based method developed by Phaneuf, Kling, and Herriges (2000). In particular, for a given small amount of change in the price of good j ($\equiv \Delta p_j$),[11] and a given set of parameter estimates of a sample b, we can obtain Marshallian (uncompensated) price elasticities of demand with respect to the good j at a specific individual characteristic $(\boldsymbol{p}^0, m^0)$ from the following steps:

1. Generate N vectors of error terms: $\boldsymbol{\varepsilon}_1, \cdots, \boldsymbol{\varepsilon}_N$.

2. Obtain N vectors of optimal constrained shares $\tilde{\boldsymbol{w}}^0_{1,b}, \cdots, \tilde{\boldsymbol{w}}^0_{N,b}$ using the given set of parameter estimates and the error vectors $\boldsymbol{\varepsilon}_1, \cdots, \boldsymbol{\varepsilon}_N$, but fixing $(\boldsymbol{p}^0, m^0)$ across $n = 1, \cdots, N$.

3. Obtain another set of N optimal constrained share vectors $\tilde{\boldsymbol{w}}^1_{1,b}, \cdots, \tilde{\boldsymbol{w}}^1_{N,b}$ in a similar manner, but using $(\boldsymbol{p}^1, m^1)$ where $p^1_j \equiv p^0_j + \Delta p_j$.

4. Evaluate the mean share vectors for $k = 0, 1$:
$$\bar{\tilde{\boldsymbol{w}}}^k_b \equiv \frac{1}{N} \sum_{n=1}^{N} \tilde{\boldsymbol{w}}^k_{n,b}.$$

5. Obtain a set of demand elasticities using the mid-point formula:
$$\eta_{ij,b} = -\mathbf{1}\{i = j\} + \frac{(\bar{\tilde{\boldsymbol{w}}}^1_b - \bar{\tilde{\boldsymbol{w}}}^0_b)}{\Delta p_i} \cdot \frac{(p^0_j + p^1_j)}{(\bar{\tilde{\boldsymbol{w}}}^1_b + \bar{\tilde{\boldsymbol{w}}}^0_b)}, \qquad i = 1, \cdots, M.$$

Note that the steps can be repeated for each sample $b = 1, \cdots, B$, and for each model **T1**, **T2** and **T3**, but we maintain the same $N \times M$ array of uniform random variables to generate the vectors of error terms $\boldsymbol{\varepsilon}_1, \cdots, \boldsymbol{\varepsilon}_N$ across samples and models to produce as consistent results as possible.[12] Moreover, the antithetic sampling method (Geweke 1988) is applied to generate the array of uniform random variables to accelerate the simulation.

For each model, we obtain three income elasticities, three own-price Marshallian elasticities of demand, and six cross-price Marshallian elasticities of demand at the mean prices and income across all the samples. As elasticity is unit-free, mean RMSE is calculated for each model, and the true elasticity values are obtained from the true DGP using the same simulation-based method. The following tables 3 and 4 summarize the results by models and different setups.

Firstly, in the benchmark setup (the second row of Table 3 or 4), **T3** performs the best among all the three models, followed by **T2** and **T1**. We also notice that the number of observations increases the performance, but the ranking remains the same (see Table 3). Even though **T3** falls in second place when the samples have about 10% of corner solutions, **T3** keeps first place in the samples with a higher proportion of corner solutions (see Table 4). Hence, we can conclude that the theoretical restrictions have to be imposed to as precisely predict the overall behavior of the economy as possible in the presence of the corner solutions in the demand system. However, it should be noted that the general performance of Tobit-based models highly depends on the proportions of the corner solutions. In particular, for the model **T3**, the mean RMSE of the setup **C70** is about twice as high as that of the setup **C40**, and the mean RMSE of the setup **C10** is four times as high as that of the setup **C40**.

To achieve more fundamental understanding of the performance, we decompose the mean RMSE across different corner proportions (Table 4) by the types of elasticity in Table 5: income, own-price, and cross-price elasticities. First, the overall preferable performance of **T3** is mainly driven by its superiority in estimating cross-price elasticities, but **T3** still does a reasonable job for the other two types of elasticities as well. Second, it is noticeable that **T1** performs the best in the **C10** samples in every elasticity type, which shows that the restrictions are unnecessary and could be obstructive when the dataset has a relatively small number of corner solutions. Lastly, even though **T3** is recommended, especially for datasets with moderate proportions of corner solutions, it has to be noted that the income elasticity estimates are relatively biased compared to the price elasticity estimates.

6. Conclusions

In this paper, we examined three Tobit-based econometric models that are universally adopted for estimating demand systems in the presence of corner solutions: a simple equation-by-equation Tobit, a system of correlated truncated equations, and the Amemiya-Tobin with the adding-up condition. The three models are nested with each other in terms of the degree of the theoretical restrictions, and we evaluated their performances in capturing the true parameters and elasticities. For a series of Monte Carlo simulations, we used the LA-AI demand system for the true DGP because of its popularity and flexibility, and the virtual price technique was adapted to generate the constrained optimal consumption from the indirect utility function of the LA-AI demand system. Using 1,000 different samples of 500 individuals, we obtained 1,000 different parameter estimates for each Tobit-based model and analyzed the performances of the models in terms of capturing the true parameters and the true elasticity values.

From the examination of the samples and the simulation results, we have obtained the following findings. First, adding theoretical restrictions to the Tobit model following the theory of choice does not hurt the flexibility of the model but indeed improves the ability to capture the true underlying elasticities. Second, having more observations helps the performance of the Tobit-based models. Third, the proportion of the corner solutions in the dataset significantly affects the performance of the Tobit models, and datasets with moderate portions of corners yield the most reliable performance compared to the other datasets with a relatively lower or higher number of corner solutions. Fourth, even though the most restrictive Amemiya-Tobin approach performs better than the other two Tobit models, it still is relatively weak at recovering income elasticities compared to price elasticities.

Nevertheless, this study has limitations, and further investigations have to be conducted. First, apart from what we have controlled for in our research, there could be other factors that might affect the performance of the Tobit-based models. For instance, in empirical works concerning real-world datasets, the true underlying preferences could not be approximated by the LA-AI demand system despite its second-order flexibility, or the heterogeneity of individuals could not be represented by the way we assumed here with normally distributed errors. Second, since we solely focused on the Tobit-based models, other frameworks that have been developed in the literature to deal with the corner solution problem are not analyzed here and need to be objectively compared with the Tobit-based models. In particular, the multiple discrete-continuous extreme value model (MDCEV) by Bhat (2005, 2008) or even the virtual price models (Lee and Pitt 1986; Phaneuf 1999) could be examined by applying the same approach.

Notes

[1]Equation (3) cannot be used to derive the virtual prices because it is highly non-linear in $\mathbf{p}$.

[2]Equation (9) is not required for the data generation, but it is so for the virtual price estimation procedure.

[3]$\iota_M \equiv (1, \cdots, 1)'$

[4]These papers are Dong, Davis, and Stewart (2014); Zhen et al. (2013); Dong, Gould, and Kaiser (2004); Phaneuf, Kling, and Herriges (2000); Phaneuf (1999); Cornick, Cox, and Gould (1994); Heien and Wesseils (1990); and Wales and Woodland (1983).

[5]Precisely, $m \sim \ln N\left(\mu = \ln \frac{1^2}{\sqrt{1^2+0.5^2}}, \sigma^2 = \ln \frac{1^2+0.5^2}{1^2}\right)$ and $p_i \sim \ln N\left(\mu = \ln \frac{1^2}{\sqrt{1^2+1^2}}, \sigma^2 = \ln \frac{1^2+1^2}{1^2}\right)$ for $i = 1, \cdots, M$ are assumed to have distributions of $(\boldsymbol{p}, m)$ with the desired means and variances.

[6]For ease of reference, we call (**N100**, **C40**) or (**N2000**, **C40**) as **N100** or **N2000**, and (**N500**, **C10**) or (**N500**, **C70**) as **C10** or **C70**, respectively. Note that both **N500** and **C40** are reffering the same benchmark setup.

[7]i.e., $(\Sigma_{11}, \Sigma_{21}, \Sigma_{22}) = \left(0.1^2, -0.4 \cdot (0.1 \cdot 0.2), 0.2^2\right)$.

[8]The zero vector ($\mathbf{0}_M$) cannot be optimal for consumers with locally nonsatiated preferences.

[9]In other words, the model cannot be used for predictions, data generations, or simulations, but only the elasticity estimates are available.

[10]The mapping is defined for the re-indexed shares $\tilde{\boldsymbol{w}} = (\tilde{w}_1, \cdots, \tilde{w}_K, \mathbf{0}_{M-K})$ without loss of generality. See Equation (18) of Wales and Woodland (1983) for the standard mapping definition.

[11]We use 10% for the small change, and the income elasticity of demand can be calculated in a similar manner.

[12]Note that **T1** uses all the three columns of the array of uniform random variables, whereas **T2**, **T3** and, the true DGP use the first two columns.

References

Amemiya, T. 1979. "The Estimation of a Simultaneous-Equation Tobit Model." *International Economic Review* 20:169–181.

Banks, J., R. Blundell, and A. Lewbel. 1997. "Quadratic Engel Curves and Consumer Demand." *Review of Economics and Statistics* 79:527–539.

Bhat, C.R. 2005. "A Multiple Discrete–Continuous Extreme Value Model: Formulation and Application to Discretionary Time-use Decisions." *Transportation Research Part B: Methodological* 39:679–707.

—. 2008. "The Multiple Discrete-Continuous Extreme Value (MDCEV) Model: Role of Utility Function Parameters, Identification Considerations, and Model Extensions." *Transportation Research Part B: Methodological* 42:274–303.

Cornick, J., T.L. Cox, and B.W. Gould. 1994. "Fluid Milk Purchases: A Multivariate Tobit Analysis." *American Journal of Agricultural Economics* 76:74–82.

Dong, D., C.G. Davis, and H. Stewart. 2014. "The Quantity and Variety of Households' Meat Purchases: A Censored Demand System Approach." *Agricultural Economics* 46:99–112.

Dong, D., B.W. Gould, and H.M. Kaiser. 2004. "Food Demand in Mexico: An Application of the Amemiya-Tobin Approach to the Estimation of a Censored Food System." *American Journal of Agricultural Economics* 86:1094–1107.

Geweke, J. 1988. "Antithetic Acceleration of Monte Carlo Integration in Bayesian Inference." *Journal of Econometrics* 38:73–89.

—. 1991. "Efficient Simulation from the Multivariate Normal and Student-t Distributions Subject to Linear Constraints and the Evaluation of Constraint Probabilities." In *Computing Science and Statistics: Proceedings of the 23rd Symposium on the Interface*. pp. 571–578.

Hajivassiliou, V.A., and D. McFadden. 1991. "The method of simulated scores for the estimation of LDV models with an application to external debt crisis.", pp. .

Heien, D., and C.R. Wesseils. 1990. "Demand Systems Estimation With Microdata: A Censored Regression Approach." *Journal of Business & Economic Statistics* 8:365–371.

Keane, M.P. 1994. "A Computationally Practical Simulation Estimator for Panel Data." *Econometrica* 62:95.

Lee, L.F., and M.M. Pitt. 1986. "Microeconometric Demand System with Binding Non-negativity Constraints: The Dual Approach." *Econometrica: Journal of the Econometric Society*, pp. 1237—1242.

Lewbel, A., and K. Pendakur. 2009. "Tricks with Hicks: The EASI Demand System." *American Economic Review* 99:827–863.

Nelson, F., and L. Olson. 1978. "Specification and Estimation of a Simultaneous-Equation Model with Limited Dependent Variables." *International Economic Review* 19:695.

Phaneuf, D.J. 1999. "A Dual Approach to Modeling Corner Solutions in Recreation Demand." *Journal of Environmental Economics and Management* 37(1):85–105.

Phaneuf, D.J., C.L. Kling, and J.A. Herriges. 2000. "Estimation and Welfare Calculations in a Generalized Corner Solution Model with an Application to Recreation Demand." *Review of Economics and Statistics* 82:83–92.

Tobin, J. 1958. "Estimation of Relationships for Limited Dependent Variables." *Econometrica* 26:24.

Wales, T.J., and A.D. Woodland. 1983. "Estimation of Consumer Demand Systems with Binding Non-negativity Constraints." *Journal of Econometrics* 21:263–285.

Zhen, C., E.A. Finkelstein, J.M. Nonnemaker, S.A. Karns, and J.E. Todd. 2013. "Predicting the Effects of Sugar-Sweetened Beverage Taxes on Food and Beverage Demand in a Large Demand System." *American Journal of Agricultural Economics* 96:1–25.

Figures

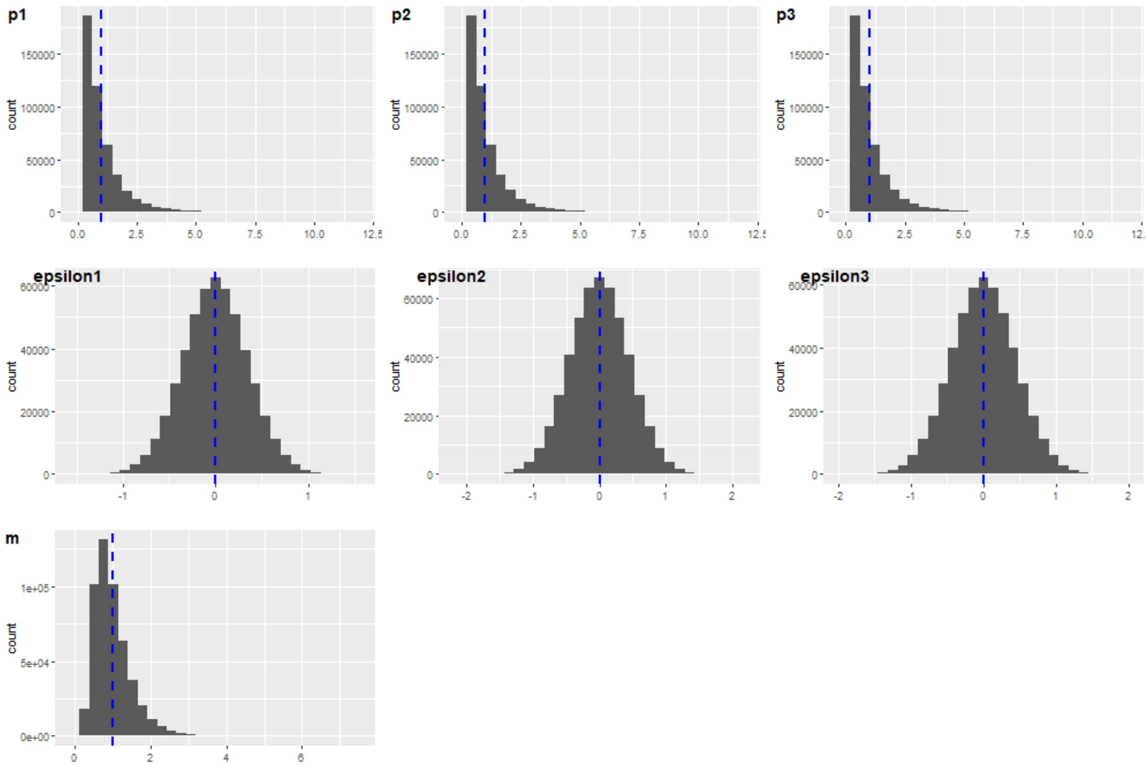

Figure 1. Histograms for $(\boldsymbol{p}, \boldsymbol{\varepsilon}, m)$ in $B = 1,000$ Samples

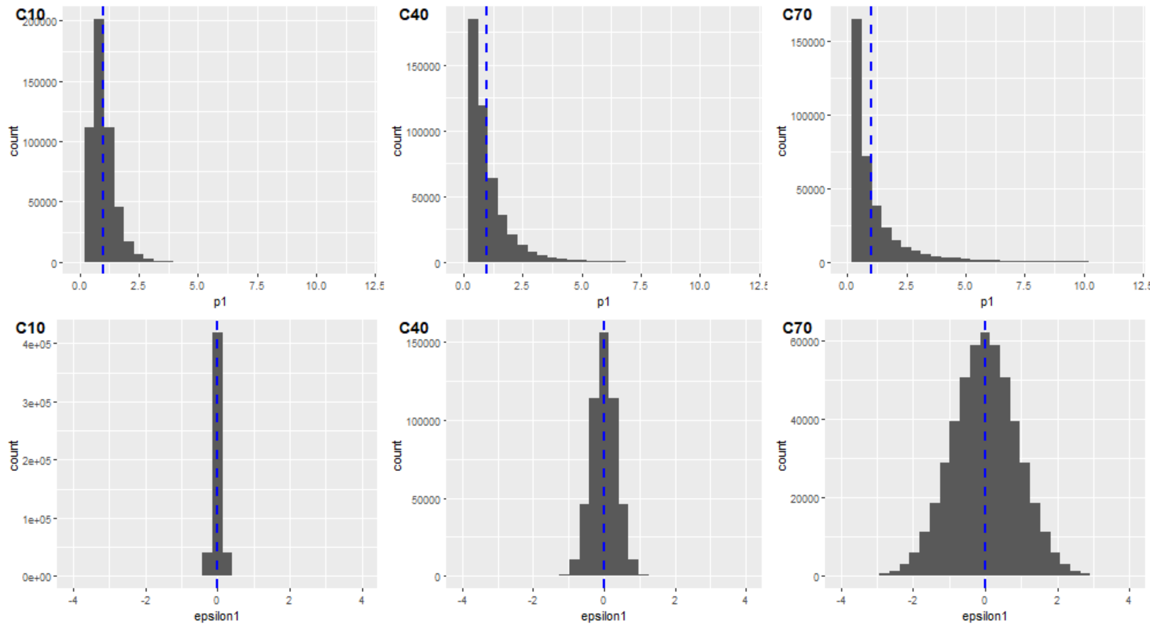

Figure 2. Histograms for p_1 and u_1 from the Setups C10, C40, and C70

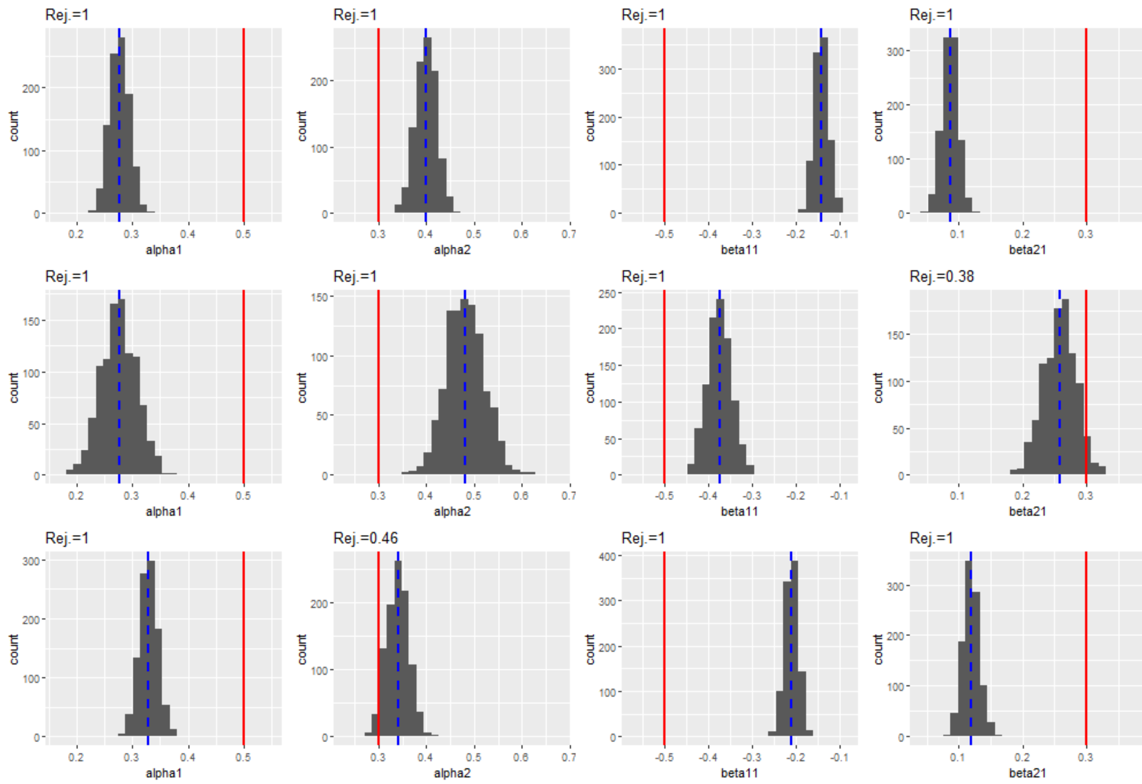

Figure 3. Parameter Estimates in $B = 1,000$ Samples (1/2)

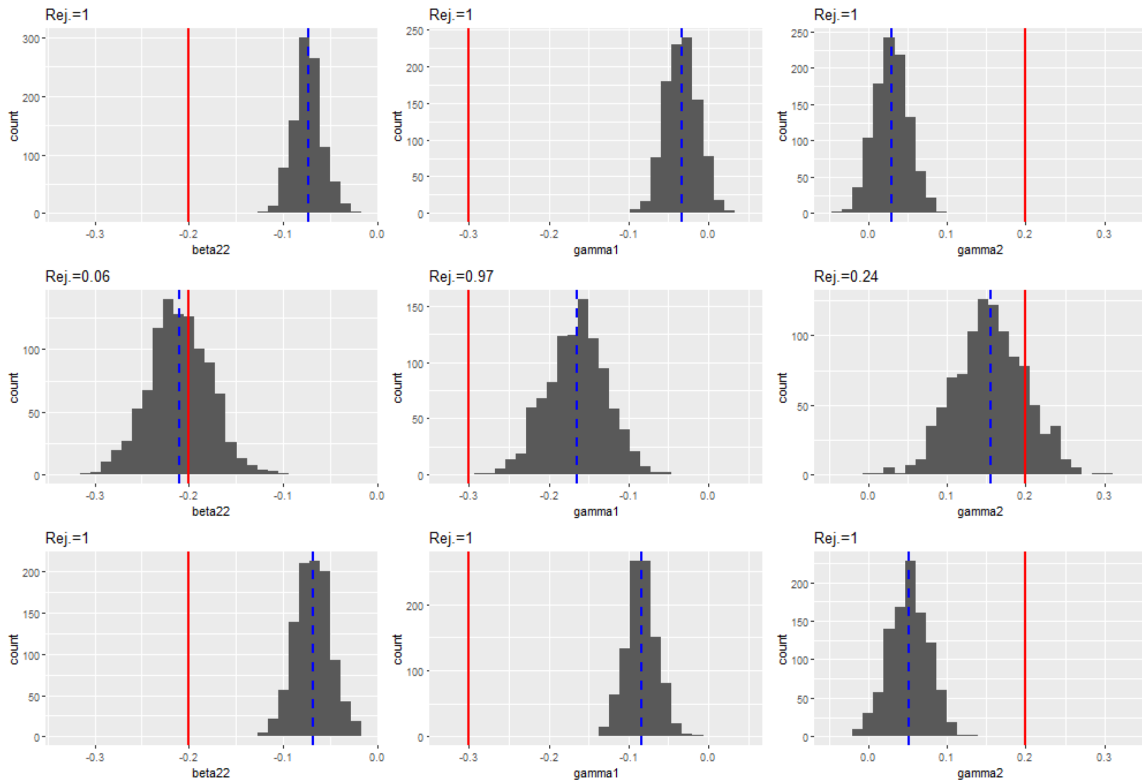

Figure 4. Coefficient Estimates in $B = 1,000$ Samples (2/2)

Tables

Table 1. Mean NRMSE by Models and Number of Observations: Parameters

	T1	T2	T3
N100	0.212	**0.143**	0.178
N500	0.210	**0.118**	0.170
N2000	0.209	**0.112**	0.169

Table 2. Mean NRMSE by Models and Corner Proportions: Parameters

	T1	T2	T3
C10	0.088	**0.020**	0.048
C40	0.210	**0.118**	0.170
C70	0.358	0.289	**0.239**

Table 3. Mean RMSE by Models and Number of Observations: Elasticity

	T1	T2	T3
N100	0.628	0.549	**0.477**
N500	0.424	0.309	**0.274**
N2000	0.371	0.249	**0.227**

Table 4. Mean RMSE by Models and Corner Proportions: Elasticity

	T1	T2	T3
C10	**0.747**	1.128	1.062
C40	0.424	0.309	**0.274**
C70	0.814	0.688	**0.633**

Table 5. Mean RMSE by Models and Corner Proportions: Elasticity by Types

	Income			Own-price			Cross-price		
	T1	T2	T3	T1	T2	T3	T1	T2	T3
C10	**0.883**	1.448	1.345	**0.157**	0.252	0.213	**0.973**	1.406	1.345
C40	0.843	**0.527**	0.540	**0.056**	0.081	0.067	0.399	0.315	**0.245**
C70	1.044	**0.755**	0.997	0.206	0.180	**0.171**	1.004	0.908	**0.683**