GMM

Timothy

July 18, 2018

Contents

1 Estimation

It is known that in each time period, the distribution of supply and demand is binomial. However the difference between two binomial distributions that are not independent is difficult to estimate using likelihood methods. As a result, the generalized method of moments will be utilized. Since the price is uniquely defined in each time period, as is the quantity supplied, the question of estimation is feasible.

In each time period T, there exists two moment conditions specified, one for price, and one for quantity. Under the specification for the model, for each time period T: $F_V(p_T^*) = \prod_{t=1}^{T}(1-\xi_t)$ and $q_T^* = \xi_T \prod_{t=1}^{T}(1-\xi_t)$. This provides us with 2T moment restrictions on the model, and allows for estimation of up to 2T parameters.

A distinction must be made between observations and time periods. The data are divided into the median price and quantity sold in each day, and the question of how many data points are in a time period exists. For the purposes of the estimation in this paper, I will use 5 observations per time period. If there are N observations, then there are $T = \frac{N}{5}$ time periods.

For the model specified with T time periods, and for a distribution of prices of log-normal, there are 2 parameters for the distribution, and T parameters for the ξ. There are 2T moment restrictions, so the model is in fact over-identified. This allows us to test the specification for our model using the Sargan-Hansen J-test.

1.1 Complications

One important complication is that there exists a price-floor in the market. No item is able to be sold at less than \$ 0.03, this means that for all data points where the price is at this floor, the equilibrium condition is not binding. Since a price floor leads to excess supply at the binding price, the only condition that remains binding is that quantity demanded at the given price is equal to the quantity sold. Denote K as the number of time periods in which the price floor is binding.

This condition is written as: $q_d^T = N \prod_{t=1}^{T} \left[1 - \frac{F_V(p_T^*)}{\prod_{t=1}^{T-1}(1-\xi_t)}\right]$. For each time period where the price is at the floor, there is only one moment condition. For this model, this implies that there must be at least two time periods where the price is above the floor in order to identify the model. This condition is upheld in all the data sets examined in this paper, and effectively reduces the number of moments. In more complicated settings with more primitives in the model, this could become an important problem, as the current specification has the equilibrium price converging to zero in time.

1.2 Implementation

Consider a function g(Y_t ,μ,σ,ξ) which gives the moment condition for each time period, evaluated at the t^{th} element in that time period. Under the Null Hypothesis that this model fits the data, then the expected value of this function is zero.

$$\mathbb{E}[g(Y_t, \mu, \sigma, \xi)] = 0$$

We seek to estimate the parameters μ, σ, ξ by minimizing the sample analog of this with respect to a weighting matrix W. The sample analog is formed by averaging the data found contained in each time period. $\hat{m}(\mu, \sigma, \xi) = \frac{1}{M}\sum_{m=1}^{M} g(Y_m, \mu, \sigma, \xi)$. Let us combine the parameters of the model into a vector θ. Our goal then becomes to estimate a value of $\hat{\theta}$ by minimizing the quadratic form of $\hat{(m)}$ with respect to matrix W.

$$\hat{\theta} = \arg\min_{\theta} \hat{m}(\theta)' W \hat{m}(\theta)$$

The choice of W is selected by first choosing a positive definite matrix W, and estimating the model, and then estimating the matrix by the following method:

$$\hat{W}_i = \left[\frac{1}{M} \sum_{m=1}^{M} g(Y_m, \hat{\theta_{i-1}}) g(Y_m, \hat{\theta_{i-1}})' \right]^{-1}$$
$$\hat{\theta}_i = \arg\min_{\theta} \hat{m}(\theta_i)' \hat{W}_i \hat{m}(\theta_i)$$

This process is then continued until the value of θ_{i-1} is a minimizer for $\mathrm{W_i}$. This iterated GMM estimator is invariant to the scale of the data, which is important in this model, as the price and the quantity data are of wildly different magnitudes. (Cite Hamilton 1994) This method is also asymptotically equivalent to the Continuous Updating Efficient GMM, but does not have as many numerical instabilities. This process is complicated by $\hat{W}_i$ being of rank $\min\{M, 2T - K\}$. If the matrix is not of full rank, then it is not invertible, and we cannot estimate the model. In order to ensure that it has full rank, we add a positive number times the identity matrix to ensure that $\hat{W}_i$ is both positive definite and invertible.

1.3 Testing

1.3.1 Model Fit

Since our model is over-identified, we are able to test for model-fit using the J-test for model fit. Formally, we are testing the hypothesis that $M\hat{m}(\hat{\theta})'\hat{W}\hat{m}(\hat{\theta}) = 0$. Since there are 2T - K moments in the model, and T + 3 primitives in the model, the J-statistic is distributed $\chi^2_{\text{T-K-3}}$. For several of the cases examined, a table breaking down the model fit is shown.

Case	Sargan Test p-Value
Glove Case	1.0
Huntsman Case	0
Chroma Case	.89

1.3.2 Distribution Rate

Of interest is the question of whether or not there has been a constant drop rate of an item to users in the game over time. This can be written in the form of: