WEB APPENDIX

A Simulated Maximum Likelihood Estimator for the Random Coefficient Logit Model Using Aggregate Data

Sungho Park

Sachin Gupta

In this Appendix we demonstrate some statistical properties of the simulated maximum likelihood estimator proposed in Park and Gupta (2008). As described in Park and Gupta (2008), our estimation approach involves drawing a new sample of size *R* (smaller than *H*) based on shares in the observed sample of size *H*, in order to overcome numerical difficulties associated with using the observed sample, if *H* is large. Thus, our sample of size *R* could be considered a "two-stage sample". The alternative is a sample of size *R* taken directly from the population (call this a "one-stage sample"). We know that standard results, such as consistency, apply to SML estimates obtained from a one-stage sample. We determine some of the properties of the two-stage sample by comparing a one-stage sample with a two-stage sample, both of size *R*. If the two samples are very similar, we expect standard results to apply to the two-stage sample as well.

In this Appendix we proceed as follows. First we define a two-stage sample. Next, we use simulations to examine the similarity between one-stage and two-stage samples. Finally, we directly compare SML estimates obtained from one-stage and two-stage samples for different values of *R* in simulated data.

<u>Definition of two-stage sample</u>

Define

$$P_{jt} = \int \frac{\exp(x_{jt}\beta_h + \xi_{jt})}{1+\sum \exp(x_{it}\beta_h + \xi_{it})}\phi(\beta)\, d\beta \,.$$

Let us assume that we observe aggregate shares $\{S_{jt}\}(= N_{jt}/H)$ based on a sample of size *H* which is finite, and thus observed shares contain sampling error. $\{N_{jt}\}$ ($\Sigma_j N_{jt} = H$) are outcomes of H multinomial draws from probabilities $\{P_{jt}\}$. When $\{N_{jt}\}$ are large, numerical problems prevent direct application of SML. In this case, we perform *R* multinomial draws (*R*<*H*) from shares { S_{jt} } and get outcomes { M_{jt} } ($\Sigma_j M_{jt} = R$, define $Q_{jt} = M_{jt} / R$). This is our two-stage sample. By using $\{M_{jt}\}$ instead of $\{N_{jt}\}$ to obtain SML estimates, we can circumvent the numerical difficulty of handling large exponents.

<u>Comparison of two-stage sample with one-stage sample</u>

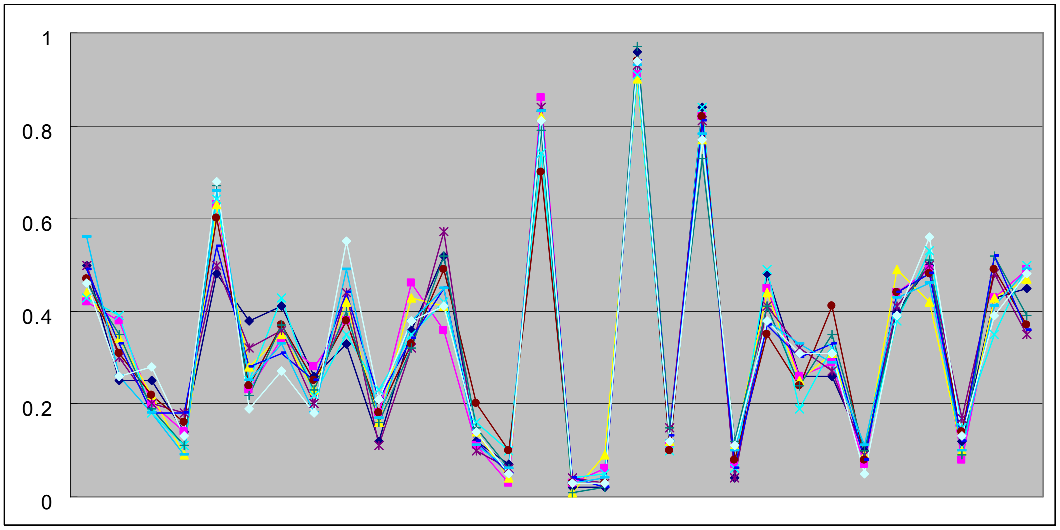
We use the DGP of case 1 of the simulation study reported in Park and Gupta (2008). We calculate $\{P_{jt}\}$ using 100,000 random draws and set *H*=1,000 and *R*=100. To generate a one-stage sample we perform *R*(=100) multinomial draws from $\{P_{jt}\}$. To generate a two-stage sample we perform *H*(=1,000) multinomial draws from $\{P_{jt}\}$, calculate $\{S_{jt}\}(= N_{jt} / H)$, and then perform *R*(=100) multinomial draws from $\{S_{jt}\}$ to obtain $\{M_{jt}\}$ or { $Q_{jt} = M_{jt} / R$ }. For each set of $\{P_{jt}\}$, we generated 100 sets of one-stage samples and 100 sets of two-stage samples. So, for P_{jt} with each *j* and *t*, we have $Q_{jt,d}^{1-stage}$ (*d*=1,…,100) from one-stage sample and $Q_{jt,d}^{2-stage}$ (*d*=1,…,100) from two-stage sample.

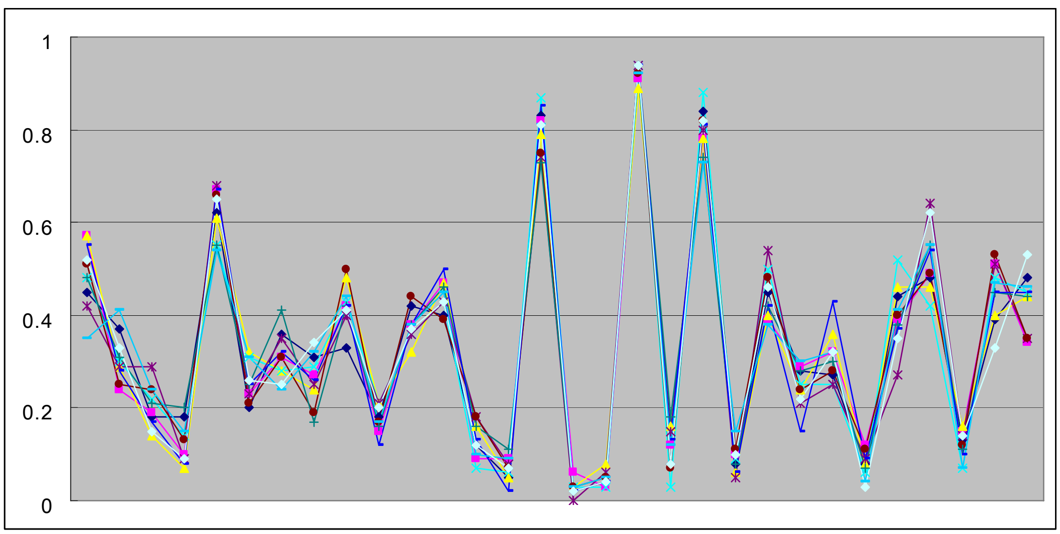
In Figure W1 we show $Q_{jt,d}^{1-stage}$ and $Q_{jt,d}^{2-stage}$ for the first ten samples, (*d*=1,…,10). (We do not show all 100 samples to allow easier reading of the figure.) We find that both $Q_{jt,d}^{1-stage}$ and $Q_{jt,d}^{2-stage}$ are centered around $\{P_{jt}\}$ and their distributions look quite similar.

Figure W1. One-stage samples (top panel) vs. Two-stage samples (bottom panel)

x-axis: (t=1; j=1), (t=1; j=2), (t=1; j=3), (t=2; j=1), (t=2; j=2), (t=2; j=3),…

y-axis: $Q_{jt,d}^{1-stage}$ (top panel) and $Q_{jt,d}^{2-stage}$ (bottom panel) for d=1,…,10.





Next we calculate the average absolute distances of $Q_{jt,d}^{1-stage}$ and $Q_{jt,d}^{2-stage}$ respectively from P_{jt}. We define

$$dist\_1-stage = \frac{\Sigma_{t=1\sim T}\Sigma_{j=1\sim J}(\Sigma_{d=1\sim 100}\,|\,P_{jt}-Q_{jt,d}^{1-stage}\,|\,/100)}{J*T}$$

$$dist\_2-stage = \frac{\Sigma_{t=1\sim T}\Sigma_{j=1\sim J}(\Sigma_{d=1\sim 100}\,|\,P_{jt}-Q_{jt,d}^{2-stage}\,|\,/100)}{J*T}$$

In our data $dist\_1-stage$ =0.033 and $dist\_2-stage$ =0.034. Table W1 summarizes the results from simulations with different values of *H* and *R*. In general, $dist\_1-stage$ is very close to $dist\_2-stage$. As *R* increases, both distances decrease. Comparing results for *H*=10,000 and *H*=1,000, we observe that $dist\_2-stage$ decreases only marginally as *H* increases. From all these results, we conclude that a one-stage sample is empirically equivalent to a two-stage sample and thus we expect that the standard results of SML apply to a two-stage sample.

Table W1. Results of Simulation Study

H	*R*	*dist_2-stage*	*dist_1-stage*
10,000	150	0.027	0.027
10,000	100	0.033	0.033
10,000	50	0.046	0.046
1,000	150	0.028	0.027
1,000	100	0.034	0.033
1,000	50	0.047	0.046

Statistical properties of the proposed SML estimator

We consider how SML estimates behave as *R* changes in a simulation study. We use the same DGP as in case 1 of the paper. We generate $\{N_{jt}\}$ with *H*=1,000 and then draw three sets of two-stage samples with *R*=(50, 100, 150) as well as one-stage samples of the same three sizes. We generated 200 data sets for this simulation study.

We expect that standard errors decrease as *R* increases. The rate of decrease is expected to be $\sqrt{N}$ if there is no numerical integration for heterogeneity and unmeasured product characteristics. In the presence of numerical integration, the rate is expected to be slower than $\sqrt{N}$.

Table W2. Results of Simulation Study to Compare SML estimates from One-stage versus Two-stage Samples

Two-stage Sample							
Parameters	True values	H=1000, R=50		H=1000, R=100		H=1000, R=150	
		Mean	SE	Mean	SE	Mean	SE
$\bar{\beta}_1$	0.2	0.211	0.100	0.190	0.100	0.185	0.088
$\bar{\beta}_2$	0.5	0.515	0.102	0.512	0.095	0.492	0.090
$\bar{\beta}_3$	-1	-1.013	0.114	-0.976	0.105	-0.985	0.090
$\bar{\beta}_4$	1	0.988	0.144	0.980	0.122	0.991	0.093
$\theta_1 = \theta_2$	1	1.006	0.101	0.996	0.111	1.028	0.111
σ_{33}	1	1.011	0.133	1.000	0.122	1.005	0.109
$SD(\omega_{2jt})$	0.707	0.710	0.069	0.736	0.063	0.755	0.058

One-stage Sample							
Parameters	True values	R=50		R=100		R=150	
		Mean	SE	Mean	SE	Mean	SE
$\bar{\beta}_1$	0.2	0.206	0.095	0.191	0.092	0.196	0.089
$\bar{\beta}_2$	0.5	0.509	0.100	0.507	0.090	0.503	0.088
$\bar{\beta}_3$	-1	-0.984	0.110	-0.976	0.103	-0.980	0.101
$\bar{\beta}_4$	1	0.979	0.138	0.980	0.138	0.984	0.132
$\theta_1 = \theta_2$	1	0.994	0.106	0.990	0.104	0.994	0.102
σ_{33}	1	0.984	0.145	0.986	0.131	0.991	0.126
$SD(\omega_{2jt})$	0.707	0.687	0.063	0.699	0.059	0.714	0.058

The top panel in Table W2 summarizes the first two moments of estimates from two-stage samples. All estimates are tightly distributed around the true values. When we compare (*H*=1000, R=50) to (*H*=1000, *R*=100) and (*H*=1000, *R*=150), we observe that for each parameter (with the exception of) SE's decrease as *R* increases. The rate of decrease is slower than $\sqrt{N}$. The lower panel reports the first two moments of estimates from one-stage samples. Here we confirm all the same results as two-stage samples. By comparing upper and lower panels, we find that in all cases (*R*=50, *R*=100, and *R*=150), we obtain comparable results from one-stage and two-stage samples.

References

Park, Sungho, and Sachin Gupta (2008), "A Simulated Maximum Likelihood Estimator for the Random Coefficient Logit Model Using Aggregate Data," working paper, Johnson Graduate School of Management, Cornell University.