Supplement to "Estimating a model of excess demand for public housing"

(*Quantitative Economics,* Vol. 4, No. 3, November 2013, 483–513)

Judy Geyer
Abt Associates

Holger Sieg
University of Pennsylvania and NBER

S.1. The nested logit model

We assume that preference shocks, denoted by $\varepsilon_{i,t}$, are continuous random variables with support over the real line. Moreover, they are i.i.d. across unit i and time t in the baseline model. Of course, the rationing rules imply that the conditional choice probabilities generated by our model also depend on the offer probabilities and, hence, do not satisfy the "independence of irrelevant alternatives" property.

One concern with these assumptions is that it is plausible that households may have an overall preference for public versus private housing. Broadly speaking, households may view the different public housing options as closer substitutes than the private housing alternative. A concern with the simple logit model specification is that it does not capture this type of correlation in unobserved preferences among public housing communities. It is, therefore, desirable to explore alternatives. We focus on the nested logit specification as the main alternative. We create one additional nest that includes all potentially available public housing choices. Changing the error structure affects household sorting and, hence, the resulting equilibrium offer probabilities, which are nonlinear functions of the structural parameters of the model.

To implement our estimator, we explored different optimization algorithms including a simplex method with simulated annealing, a gradient-based approach, and a limited grid search over possible values for the correlation coefficient. Table S.1 summarizes these efforts. Our initial starting values for all other parameters are based on our estimates for the baseline model. For the model estimated with the full set of demographic characteristics, the likelihood function increases, but not substantially, and the estimated correlation parameter is very small. This result is robust to the different optimization methods that we explored and random perturbation of the starting values used in the optimization algorithm.

Judy Geyer: judy_geyer@abtassoc.com
Holger Sieg: holgers@sas.upenn.edu

DOI: 10.3982/QE148

TABLE S.1. Estimation of the excess demand model with a nested logit specification.

	No X Covariates	Full Model
LL, no nest	−689,897	−688,796
LL, nest	−689,865	−688,791
LL ratio test	($p < 0.001$)	($p < 0.05$)
Nest correlation*	0.108	0.014
Max iterations	80K	1.8M
Computation time	200	540 hours
Minimum range**	$(-40\%, +40\%)$	$(-32\%, +30\%)$

**This is the correlation of the ε_{ij} random preferences for all public housing choices, $j \in \{1, J\}$.

**Minimum range describes the smallest interval around a starting value that the optimization explored. This range was explored concurrently with the exploration of all dimensions of the state space, as is the nature of the simple method with simulated annealing, gradient-based approaches, and a grid search. The interval points are described as percentages of the initial starting value size. For example, an interval of $(-50\%, +50\%)$ for a starting value of 2 implies that the optimizer explored values between 1 and 3.

Formal likelihood ratio tests suggest that the simple logit model is appropriate. Therefore, we conclude that the nested logit model does not improve the fit of the model. This result may not be as surprising as it appears. We offer two observations that may help to clarify our findings.

First, we should point out that the identification of the correlation parameter in our model with rationing is not as straightforward as in a standard nested logit model. Recall that our model is based on the assumption that households receive, at most, one offer to move into public housing. The offer probabilities that are implied by our estimates are small since there is a fair amount of rationing in the market for public housing. Moreover, households that are on the wait list only obtain, at most, one offer to move into one specific public housing unit. Hence they never have to choose among public housing units, that is, they do not face the blue bus–red bus problem since the red bus and the blue bus are never available at the same time. The main source of identification of the correlation coefficient of the nested logit comes from households that transfer between different public housing units. These are the only households that can (a) move to the private housing market and (b) have choices among two different public housing units. However, the fraction of households that transfer inside of public housing is small relative to number of low-income households. This may help to explain why the model fit does not improve as we relax the error structure and consider a nested logit model.

Second, we have documented in the main paper that our aggregation of public housing units into six broad categories implies that the elements in the choice set are distinctly different. Broadly speaking, we do not distinguish between housing communities that are likely to be similar. Instead, we primarily classify public housing units by size and age of targeted residents. Living in a larger public housing complex with more than 100 units is very different than living in a small community with less than 40 units. Similarly, mixed and senior communities have very different observed characteristics than family communities since they draw a larger fraction of older households.

Submitted March, 2011. Final version accepted April, 2013.