Basic Exercise on Conditional Logit
Graduate Urban Economics, SUFE

This in-class exercise is intended to give a very basic introduction to simulating a location choice model where individuals have type 1 extreme value errors ("logit errors"). We then try to back out the chosen parameters using the Stata *cmclogit* command. The basic setup we will use is:

1. There are J locations index by j and N individuals indexed by i.
2. Each location has two continuous characteristics, $x1_j$ and $x2_j$.
3. In one part of the exercise we will allow for simple heterogeneity by having two types of individuals, where each type has a different utility function. A share t_1 of individuals are type 1 while $(1 - t_1)$ are type 2.

Step 0: Create the dataset Create a dataset in Stata with J locations and N individuals, where $0 < t_1 < 1$ are type 1. In order to estimate logit models in Stata, there needs to be an observation for each individual-choice alternative. This means your dataset should have $N \times J$ observations: J possible choices (locations) for each individual. Create a variable *i* which indexes individuals and a variable *j* which indexes locations. Assign each location j values of $x1_j$ and $x2_j$, where both variables are simply draws from a standard normal distribution. You should code J, N, and t_1 as Stata *global* variables so that you can easily change these. The Stata command *rnormal()* returns normally distributed values. Lastly, create the error term $\epsilon_i j$ for each individual-choice alternative using the code[1]:

```
gen e_ij=-ln(-ln(uniform()))
```

Step 1: Simplest conditional logit model Estimate a basic model with no heterogeneity and a single continuous variable using the utility function: $V_{ij} = b_1 * x1_j + \epsilon_{ij}$. To do this, calculate V_{ij} for each individual-alternative pair. Individuals then choose the single alternative with the largest utility V_{ij}. Define a variable *choice* as a binary equal to one if the individual chose the alternative and zero otherwise; your dataset should only have N observations where *choice* equals one. You can then estimate this conditional logit model using the following two commands:

```
cmset i j
cmclogit choice x1j, noconst
```

Step 2: Varying the strength of idiosyncratic preferences A common way to vary the strength of idiosyncratic preferences is to multiply ϵ_{ij} by a constant, $\sigma > 0$. This is equivalent to changing the scale parameter of the ϵ_{ij} distribution (see footnote on Gumbel). Specifically, let the utility function be: $V_{ij} = b_1 * x1_j + \sigma\epsilon_{ij}$. When utility takes this form, then the resulting logit probabilities are:

$$P_j = \frac{exp(\frac{b_1}{\sigma} * x1_j)}{\sum_{k=1}^{J} exp(\frac{b_1}{\sigma} * x1_k)} \tag{1}$$

[1]The extreme value type 1 distribution, or Gumbel, has CDF: $Pr(X < x) = exp(-exp(\frac{x-\mu}{\sigma}))$. Since most computer programs can easily generate draws from a uniform distribution, a common trick is to invert the CDF and apply to uniform draws—the $Pr(X < x)$—to generate draws from a given distribution. Inverting this CDF gives: $ln(-ln(Pr(X < x))) = \frac{x-\mu}{\sigma}$. The μ parameter has no effect on anything—adding a constant to the utility of all choices has no effect on the maximum—and thus we normalize to 0. The σ parameter is known as the scale parameter and determines the weight of the idiosyncratic part of utility, ϵ_{ij}, versus the part affected by the covariates. Here we assume $\sigma = 1$.

Code σ as a *global* variable and then try simulating and estimating the basic model from step 1 using different values of σ. How does σ affect the choice shares? What will be choice shares as $\sigma \rightarrow \infty$?

Step 3: Tricks: Berry (RAND, 1994) shows that when there are only alternative-specific variables (only j variables), then we can simply take logs and estimate the model with OLS. Try the following, where *uniq_j* is an indicator for a unique observation of alternative j:

```
gen ln_cshare=ln(choice_share) //converges to cmclogit coefs as N->infinity
reg ln_cshare x1j if uniq_j
```

Another trick is from Guimaraes, Figureres, and Wood (ReStat, 2004), who show that when there are only alternative-specific variables and we have the counts of agents making choices, then we can estimate the model directly with a poisson model. This can be *much* faster. Try:

```
poisson choice_count x1j if uniq_j
```

Step 4: Heterogeneity Now try simulating and estimating a model with heterogeneity. Specifically, let the utility function be: $V_{ij}^t = b_1^t * x1_j + b_2^t * x2_j + \epsilon_{ij}$, where $t \in 1, 2$. This model can be estimating using the same strategy as above, but with interaction variables for one of the types. This alone is not that interesting, but with a bit more work we could use this setup to simulate a simple sorting model with two types.

Step 5: Simple equilibrium sorting model In the DO file "logit_sim_cmap.DO" I show how to simulate a sorting model with two types and two characteristics (same heterogeneity as in step 4). The key difference is that we now solve for equilibrium prices and thus can try a hedonic regression to estimate MWTP. To do so, we first need a supply of housing for every location, S_j, which we will assume is exogenous (does not depend on prices—completely inelastic supply). We can then define the equilibrium prices, p_j, as the set of prices such that for every location $j \in J$, $P_j(p_j) = S_j$, where P_j is the probability (share) of all consumers choosing location j. This involves solving J non-linear equations—a difficult problem—but luckily Bayer et al. gives us a simple contraction mapping that can do this for us. We can iterate through successive guesses for prices using:

$$p_j^t = p_j^{t-1} + ln\left(\frac{Pr_j(p_j^{t-1})}{S_j}\right)$$

Please go through my code and see if you can understand how it works. Some questions to think about:

1. Sometimes the equilibrium price of a given location will be negative, $p_j < 0$. Can negative prices be an equilibrium in this model?

2. Does the hedonic regression yield the average MWTP? Does it matter whether $x1_j$ and $x2_j$ are continuous versus discrete?

3. Do we need an instrument for prices in the logit regression?