Problem Set 3

ECON 6343: Econometrics III
Prof. Tyler Ransom
University of Oklahoma

Due: September 15, 9:00 AM

Directions: Answer all questions. Each student must turn in their own copy, but you may work in groups. Clearly label all answers. Show all of your code. Turn in jl-file(s), output files and writeup via GitHub. Your writeup may simply consist of comments in jl-file(s). If applicable, put the names of all group members at the top of your writeup or jl-file.

You will need to load the following packages:

`Optim`

`HTTP`

`GLM`

`LinearAlgebra`

`Random`

`Statistics`

`DataFrames`

`CSV`

`FreqTables`

On Github there is a file called `PS2_starter.jl` that has the code blocks below already created.

1. Estimate a multinomial logit (with alternative-specific covariates Z) on the following data set:

```
using DataFrames
using CSV
using HTTP
url = "https://raw.githubusercontent.com/OU-PhD-Econometrics/fall-2020/
master/ProblemSets/PS3-gev/nlsw88w.csv"
```

```
df = CSV.read(HTTP.get(url).body)
X = [df.age df.white df.collgrad]
Z = hcat(df.elnwage1, df.elnwage2, df.elnwage3, df.elnwage4,
         df.elnwage5, df.elnwage6, df.elnwage7, df.elnwage8)
y = df.occupation
```

The choice set represents possible occupations and is structured as follows.

1 Professional/Technical

2 Managers/Administrators

3 Sales

4 Clerical/Unskilled

5 Craftsmen

6 Operatives

7 Transport

8 Other

Hints:

- Index the parameter vector so that the coefficient on Z is the last element and the coefficients on X are the first set of elements.
- You will need to difference the Z's in your likelihood function.
- Normalize $\beta_J = 0$
- The formula for the choice probabilities will thus be

$$P_{ij} = \begin{cases} \frac{\exp\left(X_i\beta_j + \gamma(Z_{ij} - Z_{iJ})\right)}{1+\sum_{k=1}^{J-1}\exp(X_i\beta_k + \gamma(Z_{ik} - Z_{iJ}))}, & j = 1, \ldots, J-1 \\ \frac{1}{1+\sum_{k=1}^{J-1}\exp(X_i\beta_k + \gamma(Z_{ik} - Z_{iJ}))}, & j = J \end{cases}$$

2. Interpret the estimated coefficient $\hat{\gamma}$.

3. Estimate a nested logit with the following nesting structure:

   - White collar occupations (indexed by WC)

     1 Professional/Technical

     2 Managers/Administrators

     3 Sales

   - Blue collar occupations (indexed by BC)

     4 Clerical/Unskilled

     5 Craftsmen

6 Operatives

7 Transport

- Other occupations (indexed by Other)

8 Other

Specify the parameters such that there are only nest-level (rather than alternative-level) coefficients. That is, estimate a model with the following parameters:

- β_{WC}
- β_{BC}
- λ_{WC}
- λ_{BC}
- γ
- β_{Other} is normalized to 0
- The formula for the choice probabilities will thus be

$$P_{ij} = \begin{cases} \frac{\exp\left(\frac{X_i\beta_{WC}+\gamma(Z_{ij}-Z_{iJ})}{\lambda_{WC}}\right)\left[\sum_{\ell\in WC}\exp\left(\frac{X_i\beta_{WC}+\gamma(Z_{i\ell}-Z_{iJ})}{\lambda_{WC}}\right)\right]^{\lambda_{WC}-1}}{1+\left[\sum_{k\in WC}\exp\left(\frac{X_i\beta_{WC}+\gamma(Z_{ik}-Z_{iJ})}{\lambda_{WC}}\right)\right]^{\lambda_{WC}}+\left[\sum_{m\in BC}\exp\left(\frac{X_i\beta_{BC}+\gamma(Z_{im}-Z_{iJ})}{\lambda_{BC}}\right)\right]^{\lambda_{BC}}}, & j\in WC \\ \frac{\exp\left(\frac{X_i\beta_{BC}+\gamma(Z_{ij}-Z_{iJ})}{\lambda_{BC}}\right)\left[\sum_{\ell\in BC}\exp\left(\frac{X_i\beta_{BC}+\gamma(Z_{i\ell}-Z_{iJ})}{\lambda_{BC}}\right)\right]^{\lambda_{BC}-1}}{1+\left[\sum_{k\in WC}\exp\left(\frac{X_i\beta_{WC}+\gamma(Z_{ik}-Z_{iJ})}{\lambda_{WC}}\right)\right]^{\lambda_{WC}}+\left[\sum_{m\in BC}\exp\left(\frac{X_i\beta_{BC}+\gamma(Z_{im}-Z_{iJ})}{\lambda_{BC}}\right)\right]^{\lambda_{BC}}}, & j\in BC \\ \frac{1}{1+\left[\sum_{k\in WC}\exp\left(\frac{X_i\beta_{WC}+\gamma(Z_{ik}-Z_{iJ})}{\lambda_{WC}}\right)\right]^{\lambda_{WC}}+\left[\sum_{m\in BC}\exp\left(\frac{X_i\beta_{BC}+\gamma(Z_{im}-Z_{iJ})}{\lambda_{BC}}\right)\right]^{\lambda_{BC}}}, & j=J \end{cases}$$

4. Wrap all of your code above into a function and then call that function at the very bottom of your script. Make sure you add `println()` statements after obtaining each set of estimates so that you can read them.