

ProblemSet5

Xinya (Lucia) Yu

Data

There are two markets in this problem set: 2007 and 2008. There are 45 buyers and 45 targets in market 2007, and there are 54 buyers and 54 targets in market 2008. In each market, I need to compare 'observed' matches and 'counterfactual' matches. Therefore, in the new dataset `df_match` and `df_match2`, there are 2,421 unique comparisons.

Model 1

The first payoff function has the form

$$f_m(b, t) = x_{1bm}y_{1tm} + \alpha x_{2bm}y_{1tm} + \beta distance_{btm} + \varepsilon_{btm} \quad (1)$$

Therefore, the maximum score objective function is

$$Q(\alpha, \beta) = \sum_{m \in M} \sum_{i \in U_m} \sum_{j \in U_m \setminus i} 1[(f_{\alpha, \beta}(i, i) + f_{\alpha, \beta}(j, j)) > (f_{\alpha, \beta}(i, j) + f_{\alpha, \beta}(j, i))] \quad (2)$$

As ε_{btm} is independent across matches, so it can be eliminated in the inequation

For model 2, the first payoff function is

$$f_m(b, t) = \delta x_{1bm}y_{1tm} + \alpha x_{2bm}y_{1tm} + \gamma HHI_{tm} + \beta distance_{btm} + \varepsilon_{btm} \quad (3)$$

Therefore, the maximum score objective function is

$$Q(\delta, \alpha, \gamma, \beta) = \sum_{m \in M} \sum_{i \in U_m} \sum_{j \in U_m \setminus i} 1[(f_{\delta, \alpha, \gamma, \beta}(i, i) + f_{\delta, \alpha, \gamma, \beta}(j, j)) > (f_{\delta, \alpha, \gamma, \beta}(i, j) + f_{\delta, \alpha, \gamma, \beta}(j, i))] \quad (4)$$

Result

The result for model 1 is

```
For the model1:  
The estimate of alpha is 0.5374999999999981  
The estimate of beta is -0.7699999999999984  
The maximum socre estimator is -2286
```

This means that from all the 2,421 comparisons, my model estimates 2286 comparisons correctly. The estimated model is

$$f_m(b, t) = x_{1bm}y_{1tm} + 0.54x_{2bm}y_{1tm} - 0.77distance_{btm} + \varepsilon_{btm} \quad (5)$$

The result for model 1 is

```
For the model2:  
The estimate of delta is 0.10187499999999997  
The estimate of alpha is 0.09671874999999999  
The estimate of gamma is 0.09796874999999997  
The estimate of beta is -0.10421875  
The maximum socre estimator is -2286
```

This means that from all the 2,421 comparisons, my model estimates 2286 comparisons correctly. The estimated model is

$$f_m(b, t) = 0.1x_{1bm}y_{1tm} + 0.1x_{2bm}y_{1tm} + 0.1HHI_t m - 0.1distance_{btm} + \varepsilon_{btm} \quad (6)$$