

IO Problem Set 4

Michael B. Nattinger

November 16, 2021

At the end of the document I present a series of screenshots of the Stata output. I was going to present nicely formatted tables instead but I don't have the time. Still, all of the information one would expect in the tables should be represented by the output.

1 Question 1: Probit entry

For the Walmart entry probit, I found the following specification to perform the best: Kmart presence, log of county population, log of retail sales, percent of urban population, log distance to Benton, Southern (dummy), number of small stores, and distance weighted num of Kmart stores.

For the Kmart entry probit, the optimal model was similar: Walmart presence, log of county population, log of retail sales, percent of urban population, Midwest (dummy), log distance to Benton, and Southern (dummy).

The model is blind to entry of Walmart into a country that already has a Walmart. This is probably a bad assumption, as Walmart presence will have a massive impact on Walmart entry into a country, as Walmart will try to reduce competition for themselves.

Also, we are proxying for entry with presence in a market. These concepts of course are related, but not the same.

2 Question 2: Probit

As requested, we use the variables we excluded in Question 1 as instruments here. Technically speaking, we should not use distance from Benton county as an instrument as it was not excluded in Question 1. Benton is where Walmart is headquartered so it may directly influence Walmart's decision but not Kmart's, so it may only affect Kmart through its influence on Walmart. That being said, I think the 'Southern' dummy is going to suck up a lot of the variation from this variable as Benton county is in Arkansas. We could potentially take it out of Kmart's regression from Question 1 and use it as an instrument in Kmart's part of problem 2, I suppose.

3 Question 3

Here we do the Bresnahan and Reiss stuff. One problem here is that we are thinking of Big vs Small players but the only big players we are considering are Walmart and Kmart? This is a bad assumption and influences our results dramatically. This is a big-time limitation.

4 Question 4

Here I applied the two-step: Probit with entry of Walmart as dependent variable, then estimate Walmart entry probabilities. These probability estimates are then used in the second stage regression where Kmart entry is the dependent variable and Walmart probabilities are included

as a covariate. We also do the opposite, where we look at Probit regression with Kmart entry as dependent variable in first-stage, then in second stage included probability estimates from first stage as explanatory variables in the second stage regression, where Walmart entry is the dependent variable.

These results indicate that the entry of the other player in the market has a massive influence on the entry decisions of a player. Much more than other sections. When the other firm enters, it reduces my profits from entering, and such reduces my probability of entry.

5 Stata Output

Here is all of the content of the pset. I will include all of my code in with my Canvas submission. Sorry if it is not that well organized, my TA position is absolutely ruining my life right now.

```
. // Walmart entry (specifications that best fit the data)
. probit walmart kmart log_cnty_pop log_cnty_rtl num_stores south perc_pop log_dist_B dist_kmart
```

```
Iteration 0:  log likelihood = -1428.878
Iteration 1:  log likelihood = -683.31965
Iteration 2:  log likelihood = -666.35584
Iteration 3:  log likelihood = -666.21332
Iteration 4:  log likelihood = -666.21331
```

Probit regression	Number of obs	=	2,065
	LR chi2(8)	=	1525.33
	Prob > chi2	=	0.0000
Log likelihood = -666.21331	Pseudo R2	=	0.5338

walmart	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
kmart	- .7385315	.1136444	-6.50	0.000	-.9612704	-.5157927
log_cnty_pop	2.281686	.1170737	19.49	0.000	2.052226	2.511147
log_cnty_rtl	1.599726	.1267704	12.62	0.000	1.35126	1.848191
num_stores	-.0842164	.0200799	-4.19	0.000	-.1235724	-.0448605
south	.8695156	.0910077	9.55	0.000	.6911438	1.047887
perc_pop	.9830968	.2073893	4.74	0.000	.5766213	1.389572
log_dist_B	-1.09051	.0792035	-13.77	0.000	-1.245746	-.9352736
dist_kmart	-3.880088	.6387255	-6.07	0.000	-5.131967	-2.628209
_cons	-13.44497	1.110049	-12.11	0.000	-15.62063	-11.26931

```
. // KMart entry (specifications that best fit the data)
. probit kmart walmart log_cnty_pop log_cnty_rtl midwest south perc_pop log_dist_B
```

```
Iteration 0: log likelihood = -1004.9923
Iteration 1: log likelihood = -606.2221
Iteration 2: log likelihood = -542.76451
Iteration 3: log likelihood = -540.43304
Iteration 4: log likelihood = -540.42834
Iteration 5: log likelihood = -540.42834
```

```
Probit regression                                Number of obs    =      2,065
                                                LR chi2(7)       =      929.13
                                                Prob > chi2      =      0.0000
Log likelihood = -540.42834                    Pseudo R2       =      0.4623
```

kmart	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
walmart	-.556266	.1141986	-4.87	0.000	-.7800911	-.3324408
log_cnty_pop	1.752464	.1122638	15.61	0.000	1.532431	1.972497
log_cnty_rtl	1.682724	.1423129	11.82	0.000	1.403796	1.961652
midwest	1.184473	.1730199	6.85	0.000	.8453606	1.523586
south	.9215157	.1715677	5.37	0.000	.5852491	1.257782
perc_pop	1.315929	.230556	5.71	0.000	.8640476	1.76781
log_dist_B	.6816653	.0996091	6.84	0.000	.486435	.8768956
_cons	-26.21273	1.554173	-16.87	0.000	-29.25885	-23.1666

Probit model with endogenous regressors Number of obs = **2,065**
Wald chi2(8) = **780.78**
Log likelihood = **-244.0615** Prob > chi2 = **0.0000**

	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
kmart	-2.957882	.1479295	-20.00	0.000	-3.247819	-2.667945
log_cnty_pop	1.760684	.0985244	17.87	0.000	1.56758	1.953789
log_cnty_rtl	1.227663	.103535	11.86	0.000	1.024738	1.430587
num_stores	-.0510097	.0149323	-3.42	0.001	-.0802766	-.0217429
south	.3769998	.0771848	4.88	0.000	.2257204	.5282792
perc_pop	.8654146	.1632083	5.30	0.000	.5455323	1.185297
log_dist_B	-.3326666	.0711245	-4.68	0.000	-.472068	-.1932652
dist_kmart	-2.256799	.4819749	-4.68	0.000	-3.201453	-1.312146
_cons	-13.0684	.9525141	-13.72	0.000	-14.93529	-11.20151
corr(e.kmart,e.walmart)	.9654912	.0105585			.937384	.9811046
sd(e.kmart)	.2511887	.0039202			.2436217	.2589908

Instrumented: kmart

Instruments: log_cnty_pop log_cnty_rtl num_stores south perc_pop log_dist_B
dist_kmart cnty midwest dist_walmart opt1 opt2 opt3

Wald test of exogeneity (corr = 0): chi2(1) = **168.57** Prob > chi2 = **0.0000**

Probit model with endogenous regressors Number of obs = **2,065**
Wald chi2(7) = **562.83**
Log likelihood = **-220.56831** Prob > chi2 = **0.0000**

	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
walmart	-2.796373	.1840061	-15.20	0.000	-3.157018	-2.435728
log_cnty_pop	1.49954	.0859387	17.45	0.000	1.331103	1.667976
log_cnty_rtl	1.178443	.1077138	10.94	0.000	.9673279	1.389558
midwest	.6913683	.13597	5.08	0.000	.4248721	.9578645
south	.7758942	.131138	5.92	0.000	.5188684	1.03292
perc_pop	1.174436	.1646542	7.13	0.000	.85172	1.497153
log_dist_B	-.0517235	.0835062	-0.62	0.536	-.2153927	.1119457
_cons	-14.29317	1.340861	-10.66	0.000	-16.92121	-11.66513
corr(e.walmart,e.kmart)	.9433428	.0183989			.8937088	.9701647
sd(e.walmart)	.2461716	.0038554			.23873	.2538452

Instrumented: walmart

Instruments: log_cnty_pop log_cnty_rtl midwest south perc_pop log_dist_B cnty
num_stores dist_kmart dist_walmart opt1 opt2 opt3

Wald test of exogeneity (corr = 0): chi2(1) = **111.89** Prob > chi2 = **0.0000**

```

. //////////3. Bresnahan and Reiss analysis of industry//////////
> gen large_stores = kmart+walmart

. gen total_stores = num_stores+large_stores

.
. // dependent variable = number of large players
. oprobit large_stores log_cnty_pop log_cnty_rtl num_stores south perc_pop log_dist_B dist_kmart

Iteration 0:   log likelihood = -2056.9421
Iteration 1:   log likelihood = -1138.5256
Iteration 2:   log likelihood = -1089.0606
Iteration 3:   log likelihood = -1088.3555
Iteration 4:   log likelihood = -1088.3555

Ordered probit regression               Number of obs       =      2,065
                                         LR chi2(7)           =     1937.17
                                         Prob > chi2          =      0.0000
Log likelihood = -1088.3555             Pseudo R2            =      0.4709

```

large_stores	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
log_cnty_pop	2.018999	.0862905	23.40	0.000	1.849873	2.188126
log_cnty_rtl	1.611048	.1015303	15.87	0.000	1.412052	1.810044
num_stores	-.0613252	.015618	-3.93	0.000	-.0919358	-.0307145
south	.4549787	.0707108	6.43	0.000	.316388	.5935693
perc_pop	1.090231	.1699265	6.42	0.000	.7571814	1.423281
log_dist_B	-.3904347	.051262	-7.62	0.000	-.4909063	-.289963
dist_kmart	-2.560454	.5113532	-5.01	0.000	-3.562687	-1.55822
/cut1	16.97882	.9052189			15.20462	18.75301
/cut2	19.3068	.9342387			17.47573	21.13788

Ordered probit regression

Number of obs = 2,065

LR chi2(7) = 7139.58

Prob > chi2 = 0.0000

Pseudo R2 = 0.7413

Log likelihood = -1245.8258

total_stores	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
log_cnty_pop	1.865377	.0787969	23.67	0.000	1.710938	2.019816
log_cnty_rtl	1.276735	.085546	14.92	0.000	1.109068	1.444402
num_stores	2.787096	.0707761	39.38	0.000	2.648378	2.925815
south	.4012114	.0653627	6.14	0.000	.2731029	.5293199
perc_pop	1.064469	.1541317	6.91	0.000	.7623767	1.366562
log_dist_B	-.3453631	.0480654	-7.19	0.000	-.4395695	-.2511567
dist_kmart	-2.430058	.4853371	-5.01	0.000	-3.381301	-1.478814
/cut1	13.22565	.7651624			11.72596	14.72534
/cut2	16.57341	.8033898			14.99879	18.14802
/cut3	19.33081	.835399			17.69346	20.96816
/cut4	22.31705	.8837957			20.58484	24.04925
/cut5	25.10649	.9278527			23.28793	26.92504
/cut6	27.982	.9788047			26.06358	29.90042
/cut7	30.7608	1.029307			28.7434	32.77821
/cut8	33.4581	1.079023			31.34326	35.57295
/cut9	35.85331	1.127148			33.64414	38.06248
/cut10	39.48027	1.204218			37.12004	41.84049
/cut11	41.8972	1.258822			39.42995	44.36444
/cut12	44.81581	1.325428			42.21802	47.4136
/cut13	47.3233	1.372976			44.63232	50.01428

```
. //////////4. two-step method from Bajari et al. (2012)//////////
> // Walmart first stage, Kmart second stage
. // first stage: find estimates using probit regression from question 1
. probit walmart kmart log_cnty_pop log_cnty_rtl num_stores south perc_pop log_dist_B dist_kmart
```

```
Iteration 0: log likelihood = -1428.878
Iteration 1: log likelihood = -683.31965
Iteration 2: log likelihood = -666.35584
Iteration 3: log likelihood = -666.21332
Iteration 4: log likelihood = -666.21331
```

```
Probit regression                                Number of obs      =      2,065
                                                LR chi2(8)         =     1525.33
                                                Prob > chi2        =      0.0000
Log likelihood = -666.21331                    Pseudo R2         =      0.5338
```

walmart	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
kmart	-.7385315	.1136444	-6.50	0.000	-.9612704	-.5157927
log_cnty_pop	2.281686	.1170737	19.49	0.000	2.052226	2.511147
log_cnty_rtl	1.599726	.1267704	12.62	0.000	1.35126	1.848191
num_stores	-.0842164	.0200799	-4.19	0.000	-.1235724	-.0448605
south	.8695156	.0910077	9.55	0.000	.6911438	1.047887
perc_pop	.9830968	.2073893	4.74	0.000	.5766213	1.389572
log_dist_B	-1.09051	.0792035	-13.77	0.000	-1.245746	-.9352736
dist_kmart	-3.880088	.6387255	-6.07	0.000	-5.131967	-2.628209
_cons	-13.44497	1.110049	-12.11	0.000	-15.62063	-11.26931

```
. // second stage: use predicted value of Walmart entry in regression of Kmart entry
. probit kmart walmart_hat log_cnty_pop log_cnty_rtl midwest south perc_pop log_dist_B
```

```
Iteration 0:  log likelihood = -1004.9923
Iteration 1:  log likelihood = -539.47244
Iteration 2:  log likelihood = -434.25997
Iteration 3:  log likelihood = -428.29195
Iteration 4:  log likelihood = -428.28097
Iteration 5:  log likelihood = -428.28097
```

```
Probit regression                                Number of obs      =       2,065
                                                LR chi2(7)         =      1153.42
                                                Prob > chi2        =       0.0000
Log likelihood = -428.28097                    Pseudo R2         =       0.5738
```

kmart	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
walmart_hat	-5.797998	.4194399	-13.82	0.000	-6.620085	-4.975911
log_cnty_pop	4.000033	.2212028	18.08	0.000	3.566484	4.433583
log_cnty_rtl	3.507878	.2170918	16.16	0.000	3.082386	3.93337
midwest	1.620243	.1962461	8.26	0.000	1.235608	2.004879
south	2.096183	.2132785	9.83	0.000	1.678164	2.514201
perc_pop	3.020569	.2915511	10.36	0.000	2.449139	3.591999
log_dist_B	-.0643957	.1186302	-0.54	0.587	-.2969066	.1681151
_cons	-42.53766	2.271352	-18.73	0.000	-46.98943	-38.08589


```
. // Kmart first stage, Walmart second stage
. // first stage: find estimates using probit regression from question 1
. probit kmart walmart log_cnty_pop log_cnty_rtl midwest south perc_pop log_dist_B
```

```
Iteration 0: log likelihood = -1004.9923
Iteration 1: log likelihood = -606.2221
Iteration 2: log likelihood = -542.76451
Iteration 3: log likelihood = -540.43304
Iteration 4: log likelihood = -540.42834
Iteration 5: log likelihood = -540.42834
```

```
Probit regression                                Number of obs    =      2,065
                                                LR chi2(7)       =      929.13
                                                Prob > chi2      =      0.0000
Log likelihood = -540.42834                    Pseudo R2       =      0.4623
```

kmart	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
walmart	-.556266	.1141986	-4.87	0.000	-.7800911	-.3324408
log_cnty_pop	1.752464	.1122638	15.61	0.000	1.532431	1.972497
log_cnty_rtl	1.682724	.1423129	11.82	0.000	1.403796	1.961652
midwest	1.184473	.1730199	6.85	0.000	.8453606	1.523586
south	.9215157	.1715677	5.37	0.000	.5852491	1.257782
perc_pop	1.315929	.230556	5.71	0.000	.8640476	1.76781
log_dist_B	.6816653	.0996091	6.84	0.000	.486435	.8768956
_cons	-26.21273	1.554173	-16.87	0.000	-29.25885	-23.1666

```
. // second stage: use predicted value of Walmart entry in regression of Kmart entry
. probit walmart kmart_hat log_cnty_pop log_cnty_rtl num_stores south perc_pop log_dist_B dist_kr
```

```
Iteration 0:   log likelihood =  -1428.878
Iteration 1:   log likelihood =  -621.97459
Iteration 2:   log likelihood =  -548.88016
Iteration 3:   log likelihood =  -544.63233
Iteration 4:   log likelihood =  -544.62519
Iteration 5:   log likelihood =  -544.62519
```

```
Probit regression                               Number of obs       =      2,065
                                                LR chi2(8)           =      1768.51
                                                Prob > chi2          =      0.0000
Log likelihood = -544.62519                    Pseudo R2            =      0.6188
```

walmart	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
kmart_hat	-5.755623	.3929097	-14.65	0.000	-6.525712	-4.985534
log_cnty_pop	4.083843	.2048522	19.94	0.000	3.682341	4.485346
log_cnty_rtl	3.268508	.2047872	15.96	0.000	2.867133	3.669884
num_stores	-.1204672	.021901	-5.50	0.000	-.1633925	-.0775419
south	.9377875	.1028542	9.12	0.000	.736197	1.139378
perc_pop	2.225374	.2517886	8.84	0.000	1.731877	2.71887
log_dist_B	-1.162508	.0919088	-12.65	0.000	-1.342646	-.9823701
dist_kmart	-4.410239	.6880629	-6.41	0.000	-5.758818	-3.061661
_cons	-31.68619	1.994733	-15.88	0.000	-35.5958	-27.77659