

MIS7420

Seminar in Management Information Systems:

Paper Replication with R

Prof. Zhiqiang Zheng, Prof. Vijay Mookerjee

Author: Yihong Liu,
NETID: YXL180111,

Contents

List of Figures	3
List of Tables	4
1 Data Cleaning Process	5
1.1 BOBO1	5
2 Paper Replication	5
2.1 Table 1	5
2.2 Table 2	5
2.3 Table 3	9
2.4 Table 4	9
2.5 Table 5	9
2.6 Table 6	9
2.7 Table 7	9
2.8 Table 8	9
2.9 Table 9	9
2.10 Table 10	9
2.11 Table 11	9
2.12 Table 12	9
2.13 Table 13	9
2.14 Table 14	9
2.15 Table C1	9
2.16 Table D1-D4	10
2.17 Table E1-E2	10
2.18 Table G1-G3	10
3 Advanced Method	10
3.1 Synthetic Control Method	10
3.2 Generalized Synthetic Control Method	10
3.3 Heckit Approach	10
3.4 PSM and LA-PSM	10
3.5 Measurement Error Bias Correction	10
4 References	11

List of Figures

List of Tables

1	Summary Statistics of Top Five Vendors by Sales Volume	5
2	Summary Statistics of Referring Domain Categories	5
3	Average Difference-in-Difference (DID) of the Outcome Variables	6
4	Results of the Sales Effect (All Product Categories)	7
5	Results of the Search Effect (Pages Per Dollar, All Product Categories)	7
6	Results of the Search Effect (Minutes Per Dollar, All Product Categories)	8
7	Change in Demographics after Circuit City Store Closure	9

1 Data Cleaning Process

1.1 BOBO1

This is a reference test. Hlavac (2015) is used to generate many summary tables in this report.

2 Paper Replication

In this section, we provide our replication for this paper.

2.1 Table 1

Table 1 shows the summary statistics of top five vendors by sales volume.

Table 1: Summary Statistics of Top Five Vendors by Sales Volume

DomainName	TotalTransaction	TotalSales	TotalPagesViewed	PagesPerDollar	TotalDuration	MinsPerDollar
dell.com	1,620	483,703.300	66,953	0.138	57,225.660	0.118
amazon.com	10,904	354,573.300	464,383	1.310	369,227.900	1.041
staples.com	5,927	236,982.300	247,163	1.043	166,189.900	0.701
walmart.com	1,977	156,606.100	80,397	0.513	68,434.890	0.437
bestbuy.com	1,230	149,950.400	50,627	0.338	36,735.900	0.245

2.2 Table 2

Table 2: Summary Statistics of Referring Domain Categories

Domain Name	Total Transaction	Referred by SearchEngine	Direct to Website	Referred by Others
amazon.com	10,904	2,955(27.1%)	7,018(64.4%)	931(8.6%)
bestbuy.com	1,230	258(21.0%)	901(73.3%)	71(5.8%)
All Others	36,794	6,999(19.0%)	25,483(69.3%)	4,312(11.7%)
All Transactions	48,928	10,212(20.9%)	33,402(68.3%)	5,314(10.9%)

Table 3: Average Difference-in-Difference (DID) of the Outcome Variables

Outcome Variable	Groups	After Store Closure	Before Store Closure	First Difference (se)	DID
Amazon Sales	Control	3.418	3.303	0.115 (0.031)	-0.167
	Treatment	3.351	3.403	-0.052 (0.212)	
Amazon PagesPerDollar	Control	1.188	1.147	0.041 (0.025)	0.257
	Treatment	1.363	1.065	0.298 (0.153)	
Amazon MinsPerDollar	Control	1.016	0.975	0.041 (0.025)	0.263
	Treatment	1.187	0.882	0.304 (0.137)	
bestbuy.com Sales	Control	3.418	3.303	0.354 (0.031)	0.623
	Treatment	3.351	3.403	0.976 (0.212)	
bestbuy.com PagesPerDollar	Control	1.188	1.147	-0.109 (0.025)	0.074
	Treatment	1.363	1.065	-0.035 (0.153)	
bestbuy.com MinsPerDollar	Control	1.016	0.975	-0.084 (0.025)	-0.012
	Treatment	1.187	0.882	-0.096 (0.137)	

Table 4: Results of the Sales Effect (All Product Categories)

<i>Dependent variable:</i>				
	log(TotalMonthlySales + 1)			
	(1)	(2)	(3)	(4)
β_1	0.140 (0.525)	-0.206 (0.198)	0.087 (1.183)	-1.020 (0.906)
β_2	0.100 (0.669)	0.094 (0.245)		1.120 (1.235)
Observations	5,089	5,740	725	831
R ²	0.0002	0.001	0.0001	0.012
Adjusted R ²	-1.293	-1.209	-6.869	-6.593
F Statistic	0.196 (df = 2; 2219)	0.716 (df = 2; 2596)	0.005 (df = 1; 92)	0.653 (df = 2; 108)

Note:

*p<0.1; **p<0.05; ***p<0.01

Table 5: Results of the Search Effect (Pages Per Dollar, All Product Categories)

<i>Dependent variable:</i>				
	log(PagesPerDollar + 1)			
	(1)	(2)	(3)	(4)
β_1	-0.075 (0.413)	0.075 (0.157)	0.467 (0.677)	0.397 (0.511)
β_2	-0.262 (0.526)	-0.056 (0.194)		-0.045 (0.696)
Observations	5,089	5,740	725	831
R ²	0.0005	0.0001	0.005	0.009
Adjusted R ²	-1.292	-1.211	-6.829	-6.615
F Statistic	0.531 (df = 2; 2219)	0.120 (df = 2; 2596)	0.477 (df = 1; 92)	0.499 (df = 2; 108)

Note:

*p<0.1; **p<0.05; ***p<0.01

Table 6: Results of the Search Effect (Minutes Per Dollar, All Product Categories)

<i>Dependent variable:</i>				
	log(MinsPerDollar)			
	(1)	(2)	(3)	(4)
β_1	0.248 (0.700)	0.049 (0.275)	0.165 (1.570)	1.821 (1.211)
β_2	-0.744 (0.928)	-0.089 (0.336)		-0.459 (1.650)
Observations	4,941	5,565	677	775
R ²	0.0004	0.00003	0.0001	0.031
Adjusted R ²	-1.321	-1.237	-7.243	-6.811
F Statistic	0.389 (df = 2; 2128)	0.035 (df = 2; 2487)	0.011 (df = 1; 82)	1.548 (df = 2; 96)

Note:

*p<0.1; **p<0.05; ***p<0.01

2.3 Table 3

2.4 Table 4

2.5 Table 5

2.6 Table 6

2.7 Table 7

2.8 Table 8

2.9 Table 9

2.10 Table 10

2.11 Table 11

2.12 Table 12

2.13 Table 13

2.14 Table 14

2.15 Table C1

Table 7: Change in Demographics after Circuit City Store Closure

Group	Before Store Closure			After Store Closure			First Difference of Mean (p-value)		
	Mean Age	Mean Income	Mean Education	Mean Age	Mean Income	Mean Education	Mean Age	Mean Income	Mean Education
Control	7.048	4.479	97.957	6.937	4.498	97.999	-0.111 (<0.0001)	0.019 (0.300)	0.042 (0.639)
Treated	7.68	4.971	98.632	6.645	4.739	96.843	-1.035 (<0.0001)	-0.232 (0.029)	-1.789 (0.004)

2.16 Table D1-D4

2.17 Table E1-E2

2.18 Table G1-G3

3 Advanced Method

3.1 Synthetic Control Method

3.2 Generalized Synthetic Control Method

3.3 Heckit Approach

3.4 PSM and LA-PSM

3.5 Measurement Error Bias Correction

4 References

- [1] Marek Hlavac. Stargazer: Well-formatted regression and summary statistics tables. *R package version*, 5(1), 2015.