Spring 2020 | BUAN6337.002 | HW 4

Predictive Analytics using SAS

Group2

Adam Butcher

Dong In Kim

Maria Phetteplace

Mahmoud Ali

Q1- 2SLS

BOLD significant in the 95% confidence level

	MEAN	ms	qual	plb	price	dc
Intercept		44.33492	**-275.791**	**109.3299**	**100.4467**	**1.057601**
qual	22.7404	**0.516232**			**0.150865**	**0.035122**
plb	102.3	**-1.00879**				
price	103.8	0.825898	**2.695719**			
pion	0.5190	**7.186911**	-0.67893	1.907349	1.632240	-0.07919
ef	0.3256	**5.792290**	-2.20296	-0.10913	0.107405	0.146221
phpf	0.1570	0.562618				
plpf	0.0225	0.123434				
psc	0.00155	**-31.0497**				
papc	0.0754	-1.54248				
ncomp	2.2859	**-7.54525**				
mktexp	7.1239	-0.28075	**-0.51456**		**0.223327**	
dc	2.0909		**10.52300**	**-9.66750**	-0.61591	
tyrp	0.4903		0.380049	-0.25890	-1.37677	0.230687
pnp	5.9587		**0.209849**	**0.055289**	-0.02351	
custtyp	2.0629			**4.416412**		

ncust	5.6954			0.268310		
custsize	2.0653			0.687859		
ms	24.8689				-0.01783	0.005151
penew	51.1960					-0.00180
cap	76.6761					-0.00002
rbvi	1.9122					-0.04844
emprody	39.6422					**0.002486**
union	49.1445					0.001527

Q2- OLS

a. OLS regression
MS = 47.11 + 0.17 qual – 0.49 plb + 0.33 price + 9.85 pion + 4.50 ef + 1.50 phpf + 1.15 plpf – 20.90 psc – 1.11 papc – 7.53 ncomp – 0.11 mktexp

b. Effect of pioneering on market share
pion(9.85044) + phpf (1.49996) + plpf (1.15149) + psc (-20.89740) + papc (-1.11110) = -9.50661

Q3- 2SLS Analysis

a. Endogenous - MS; QUAL; PLB; PRICE; DC
b. Exogenous - PION; EF; PHPF; PLPF; PSC; PAPC; NCOMP; TYRP; PNP; CUSTTYPE; CUSTSIZE: PENEW; CAP; RBVI; EMPRODY; UNION; NCUST; MKTEXP
c. Effect of price on MS - For each dollar increase, relative market share has an increase of 0.83
d. Effect of MS on price - For each unit increased in relative market share, price decreases by $0.01783
e. Effect of PION on market share using all equations

Direct Effect: pion(7.186911) + phpf (0.562618) + plpf (0.123434) + psc (-31.0497) + papc (-1.54248) = -24.719217

Indirect Effect: qual(-0.67893*0.516232) + plb(1.907349*-1.00879) + price(1.632240*0.825898) =-0.35049 -1.924 +1.3481 = -0.92639

Total Effect: -25.645607

f. Effects of Endogeneity

Bias - Since independent variables and dependent variables simultaneously cause each other, the OLS model will give biased estimates (simultaneity bias).
Efficiency - OLS estimation is more precise than using IV estimation is (efficient), but we are outweighing biases against efficiency.
Consistency - Since the OLS model does not correct endogeneity problems, the OLS estimators will be inconsistent. This means that they will never converge to the true parameter values even in very large samples. Because of this, our hypothesis testing will not be valid.

Q4- **3SLS**

a. Effect of price on MS - For each dollar increase, relative market share has a unit increase of 1.32.

b. Effect of MS on price - For each unit increased in relative market share, price decreases by $ 0.01935.

c. Effect of PION on market share using all equations

Direct Effect: pion(7.122928) + phpf(-0.46357) + plpf(2.310126) + psc(-29.4711) + papc(0.166600) = -20.335

Indirect Effect: qual(-5.77661*0.444076) + plb(2.281837*-0.90687) + price(1.7999*1.325643) + dc(0.025330*) = -2.565 - 2.069 + 2.386 = -2.248
Total Effect: -22.583

d. Cross equation correlations

Cross Model Correlation

	MARKETSHA	QUALITY	PRODUCTLI	PRICE	DIRECTCOS
MARKETSHA	1.00000	-0.44161	0.14837	0.13052	0.50612
QUALITY	-0.44161	1.00000	-0.29560	-0.80001	-0.74613
PRODUCTLI	0.14837	-0.29560	1.00000	0.32422	0.31916
PRICE	0.13052	-0.80001	0.32422	1.00000	0.25720
DIRECTCOS	0.50612	-0.74613	0.31916	0.25720	1.00000

Quality is highly correlated with price and direct cost. In this circumstance, 3SLS will be different than 2SLS. Variables like price and market share have low correlation with each other so the difference between 2SLS and 3SLS will not be different. Overall, we do not think there will be a dramatic difference between the two models.

e. 3SLS provides more efficient estimates for linear regression models where the predictor variables are correlated with the error term(endogeneity) although it is difficult to use 3sls in the real world.

Looking at the R^2 value, 2SLS is better. (2SLS R^2 = 0.28195 > 3SLS R^2 = 0.1976)

3SLS - BOLD significant in the 95% confidence level

	MEAN	ms	qual	plb	price	dc
Intercept		-14.2380	**-329.446**	**111.4939**	**102.0895**	**1.052429**
qual	22.7404	**0.444076**			**0.189188**	**0.028613**
plb	102.3	**-0.90687**				
price	103.8	**1.325643**	**3.074788**			
pion	0.5190	**7.122928**	-5.77661	2.281837	1.799900	0.025330
ef	0.3256	**5.808653**	-3.12720	0.058599	0.331576	-.142785
phpf	0.1570	-0.46357				
plpf	0.0225	2.310126				
psc	0.00155	**-29.4711**				
papc	0.0754	0.166600				
ncomp	2.2859	**-8.26671**				
mktexp	7.1239	**-0.34159**	**-0.63916**		**0.205466**	
dc	2.0909		**19.35266**	**-10.0689**	**-1.86366**	
tyrp	0.4903		1.349149	-0.25492	-1.19794	0.164342
pnp	5.9587		0.076226	**0.075093**	-0.02153	
custtyp	2.0629			**3.717468**		

ncust	5.6954			0.198660		
custsize	2.0653			0.758051		
ms	24.8689				-0.01935	**0.005231**
penew	51.1960					-0.00011
cap	76.6761					0.001537(almost significant)
rbvi	1.9122					0.010385
emprody	39.6422					-0.00023
union	49.1445					-0.00011