

**Assignment 2**  
**Simultaneous Equation Models**

From the data set `assign2.dta`:

**Demand and Supply Equations**

$$\ln S_t = \beta_{10} + \beta_{11} \ln P_{Dt} + \beta_{12} \ln P_{X2t} + \beta_{13} \ln P_{X3t} + \beta_{14} \ln P_{X4t} + \varepsilon_{1t} \quad (1)$$

$$\ln D_t = \beta_{20} + \beta_{21} \ln P_{Dt} + \beta_{22} \ln GDP_t + \varepsilon_{2t} \quad (2)$$

Equilibrium condition can be achieved by  $D_t = S_t$  through the price  $P_{Dt}$  mechanism.

where:  $S_t$  = Domestic Supply at time  $t$

$D_t$  = Domestic Demand at time  $t$

$P_{Dt}$  = Domestic Price at time  $t = P_{Mt} + T_t$

$T_t$  = Tariff at time  $t$

$P_{X2t}$  = Price of Input 2 at time  $t$

$P_{X3t}$  = Price of Input 3 at time  $t$

$P_{X4t}$  = Price of Input 4 at time  $t$

$GDP_t$  = Gross Domestic Product (Representing Income) at time  $t$

Endogenous variables in this system include  $S_t$ ,  $D_t$ , and  $P_{Dt}$

Exogenous variables in this system include  $P_{X2t}$ ,  $P_{X3t}$ ,  $P_{X4t}$ , and  $GDP_t$

- State reduce form models of this system. Estimate reduce form models using OLS and prediction of the endogenous variables.
- Estimate structural form using predicted endogenous variables as independent variables in the structural form models.
- Estimate this system equations model using OLS, 2SLS, 3SLS, and I3SLS. Determine whether there exists endogeneity bias in the estimated results. Concerning on the asymptotic property, which model is the most appropriated model? Why? What do  $\beta_{21}$  and  $\beta_{22}$  mean?

**Additional Issue:**

If equilibrium doesn't hold  $D_t \neq S_t$ , when  $D_t > S_t$ ; then  $Q_t = S_t$  but when  $D_t < S_t$ ; then  $Q_t = D_t$ , where  $Q_t$  is transaction quantity at time  $t$ .

$$\ln Q_t = \beta_{10} + \beta_{11} \ln P_{Dt} + \beta_{12} \ln P_{X2t} + \beta_{13} \ln P_{X3t} + \beta_{14} \ln P_{X4t} + \varepsilon_{1t} \quad (3)$$

$$\ln Q_t = \beta_{20} + \beta_{21} \ln P_{Dt} + \beta_{22} \ln GDP_t + \varepsilon_{2t} \quad (4)$$

- Generate  $\ln Q_t$  and estimate the above system equations (model (3) and model (4)) using OLS, 2SLS, and 3SLS using  $Q_t$ , and  $P_{Dt}$  as endogenous variables and  $P_{X2t}$ ,  $P_{X3t}$ ,  $P_{X4t}$ , and  $GDP_t$  as exogenous variables.
- What are the problems, in term of economic concept and econometric technique, of the estimated results in **d**?

- a. State reduced form models of this system. Estimate reduced form models using OLS and prediction of the endogenous variables.

6204641218

$$\ln S_t = \ln D_t$$

$$\beta_{10} + \beta_{11} \ln P_{Dt} + \beta_{12} \ln P_{X2t} + \beta_{13} \ln P_{X3t} + \beta_{14} \ln P_{X4t} + \epsilon_{1t} = \beta_{20} + \beta_{21} \ln P_{Dt} + \beta_{22} \ln GDP_t + \epsilon_{2t}$$

$$\beta_{11} \ln P_{Dt} - \beta_{21} \ln P_{Dt} = \beta_{20} + \beta_{22} \ln GDP_t + \epsilon_{2t} - \beta_{10} - \beta_{12} \ln P_{X2t} - \beta_{13} \ln P_{X3t} - \beta_{14} \ln P_{X4t} - \epsilon_{1t}$$

$$\ln P_{Dt} (\beta_{11} - \beta_{21}) = \beta_{20} + \beta_{22} \ln GDP_t + \epsilon_{2t} - \beta_{10} - \beta_{12} \ln P_{X2t} - \beta_{13} \ln P_{X3t} - \beta_{14} \ln P_{X4t} - \epsilon_{1t}$$

$$\ln P_{Dt} = \frac{\beta_{20} - \beta_{10}}{(\beta_{11} - \beta_{21})} + \frac{\beta_{22}}{(\beta_{11} - \beta_{21})} \ln GDP_t - \frac{\beta_{12}}{(\beta_{11} - \beta_{21})} \ln P_{X2t} - \frac{\beta_{13}}{(\beta_{11} - \beta_{21})} \ln P_{X3t} - \frac{\beta_{14}}{(\beta_{11} - \beta_{21})} \ln P_{X4t} - \frac{\epsilon_{1t} + \epsilon_{2t}}{(\beta_{11} - \beta_{21})}$$

$$\begin{aligned} \ln P_{Dt} = & 2.87652 + .1318015 \ln P_{X2t} + .0939842 \ln P_{X3t} + .4939641 \ln P_{X4t} \\ & (2.434717) \quad (.0695123) \quad (.127627) \quad (.1936093) \\ & + .1632779 \ln GDP \\ & (.0877392) \end{aligned}$$

- b. Estimate structural form using predicted endogenous variables as independent variables in the structural form models.

$$\ln S_t = 18.59912 + 2.106112 \ln P_{Dt} - .727963 \ln P_{X2t} - 1.122146 \ln P_{X3t} - 1.428722 \ln P_{X4t}$$

(8.546622) (1.171903) (1.840856) (2.824139) (4.751381)

$$\ln D_t = 35.93498 - 2.574157 \ln P_{Dt} + .5212927 \ln GDP$$

(7.189835) (1.5697943) (1.1344816)

- c. Estimate this system equations model using OLS, 2SLS, 3SLS, and I3SLS.

Determine whether there exists endogeneity bias in the estimated results.

Concerning on the asymptotic property, which model is the most appropriated model? Why? What do  $\beta_{21}$  and  $\beta_{22}$  mean?

Hausman Test:  $\text{prob} > \chi^2 = 0.5659 \therefore$  fail to reject  $H_0$ , there is no endogeneity bias. As there is no endogeneity bias, OLS method still be the most appropriated one. Moreover, s.d. of OLS are the least as compared with 2SLS, 3SLS, I3SLS.

$\beta_{12}$  = percentage change of domestic demand when domestic price changes for one unit at time  $t$ .

$\beta_{22}$  = percentage change of domestic demand when GDP changes for one unit at time  $t$ .

- d. Generate  $\ln Q_t$  and estimate the above system equations (model (3) and model (4)) using OLS, 2SLS, and 3SLS using  $Q_t$  and  $P_{Dt}$  as endogenous variables and  $P_{X2t}$ ,  $P_{X3t}$ ,  $P_{X4t}$ , and  $GDP_t$  as exogenous variables.

OLS

$$\ln Q_t = 40.10218 - 1.353506 \ln P_{Dt} - .3864994 \ln P_{X2t} - .6782817 \ln P_{X3t} - .3606189 \ln P_{X4t}$$

(3.019386) (1.3722912) (1.1180437) (2.131651) (2.2837767)

$$\ln Q_t = 31.03578 - 2.181329 \ln P_{Dt} + .5776586 \ln GDP$$

(3.761201) (1.2946999) (1.0887536)

2SLS

$$\ln q_t = 24.95604 + .7752765 \ln P_{bt} - .5909191 \ln P_{x_{2t}} - .7971771 \ln P_{x_{3t}} - .9608797 \ln P_{x_{4t}}$$

(1.919453) (1.360145) (2.2136549) (3.277773) (5.51459)

$$\ln q_t = 35.93499 - 2.574157 \ln P_{bt} + .5212921 \ln yelp$$

(5.106302) (1.4046743) (1.0955104)

3SLS

$$\ln q_t = 24.81121 + .7879239 \ln P_{bt} - .6046439 \ln P_{x_{2t}} - .8372831 \ln P_{x_{3t}} - .9111679 \ln P_{x_{4t}}$$

(1.918922) (1.360114) (2.213472) (3.273419) (5.511692)

$$\ln q_t = 35.93499 - 2.574157 \ln P_{bt} + .5212921 \ln yelp$$

(5.106302) (1.4046743) (1.0955104)

I 3SLS

$$\ln q_t = 24.72031 + .795862 \ln P_{bt} - .6132584 \ln P_{x_{2t}} - .8624556 \ln P_{x_{3t}} - .8799661 \ln P_{x_{4t}}$$

(1.9194251) (1.370509) (2.214736) (3.291482) (5.549316)

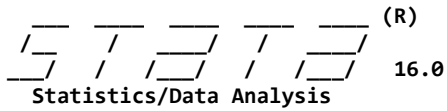
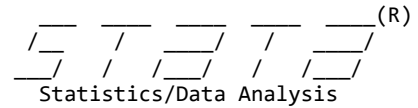
$$\ln q_t = 35.93499 - 2.574157 \ln P_{bt} + .5212921 \ln yelp$$

(5.106302) (1.4046743) (1.0955104)

- e. What are the problems, in term of economic concept and econometric technique, of the estimated results in d?

Once estimated, it is hardly recognized that whether it is demand or supply curve.

However, the result will always on the left side of equilibrium.



*MP - Parallel Edition*

Copyright 1985-2019 StataCorp LLC  
 StataCorp  
 4905 Lakeway Drive  
 College Station, Texas 77845 USA  
 800-STATA-PC <http://www.stata.com>  
 979-696-4600 [stata@stata.com](mailto:stata@stata.com)  
 979-696-4601 (fax)

20-student 2-core Stata lab perpetual license:

Serial number: 501606222284  
 Licensed to: Faculty of economics  
 Thammasat University

Notes:

1. Unicode is supported; see [help unicode advice](#).
2. More than 2 billion observations are allowed; see [help obs advice](#).
3. Maximum number of variables is set to 5000; see [help set maxvar](#).

```
1 . use "C:\Users\A\Downloads\assign2.dta"
2 . log using "C:\Users\A\Desktop\426.a2.smcl"
```

---

```
name: <unnamed>
log: C:\Users\A\Desktop\426.a2.smcl
log type: smcl
opened on: 3 Feb 2021, 17:43:54
```

```
3 . gen lnst=ln(st)
4 . gen lndt=ln(dt)
5 . gen lngdp=ln(gdp)
6 . gen pd=pm+t
7 . gen lnpd=ln(pd)
8 . gen lnpx2=ln(px2)
9 . gen lnpx3=ln(px3)
10 . gen lnpx4=ln(px4)
11 . tsset obs
    time variable: obs, 1986 to 2007
    delta: 1 unit
```

```
12 . reg lnpd lnpx2 lnpx3 lnpx4 lngdp
```

Source	SS	df	MS	Number of obs	=	22
Model	.17707359	4	.044268398	F(4, 17)	=	6.76
Residual	.111247189	17	.006543952	Prob > F	=	0.0019
				R-squared	=	0.6142
				Adj R-squared	=	0.5234
Total	.288320779	21	.013729561	Root MSE	=	.08089

lnpd	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
lnpx2	.1318015	.0695123	1.90	0.075	-.0148567	.2784596
lnpx3	.0939842	.127627	0.74	0.472	-.1752851	.3632535
lnpx4	.4939641	.1936093	2.55	0.021	.0854842	.9024439
lngdp	.1632779	.0877392	1.86	0.080	-.0218357	.3483914
_cons	2.87652	2.434717	1.18	0.254	-2.260283	8.013322

13 . predict lnpdhat (option xb assumed; fitted values) (a)

14 . reg lnst lnpdhat lnpx2 lnpx3 lnpx4 (b).

Source	SS	df	MS	Number of obs	=	22
Model	4.64569773	4	1.16142443	F(4, 17)	=	37.32
Residual	.529104183	17	.031123775	Prob > F	=	0.0000
Total	5.17480192	21	.246419139	R-squared	=	0.8978
				Adj R-squared	=	0.8737
				Root MSE	=	.17642

lnst	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
lnpdhat	2.106112	1.171903	1.80	0.090	-.3663879	4.578612
lnpx2	-.727963	.1840856	-3.95	0.001	-1.11635	-.3395762
lnpx3	-1.122146	.2824139	-3.97	0.001	-1.717988	-.5263052
lnpx4	-1.428722	.4751381	-3.01	0.008	-2.431176	-.4262679
_cons	18.59912	8.546622	2.18	0.044	.5673274	36.63092

15 . reg lndt lnpdhat lngdp (b).

Source	SS	df	MS	Number of obs	=	22
Model	3.26129847	2	1.63064924	F(2, 19)	=	44.99
Residual	.688574614	19	.036240769	Prob > F	=	0.0000
Total	3.94987309	21	.188089195	R-squared	=	0.8257
				Adj R-squared	=	0.8073
				Root MSE	=	.19037

lndt	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
lnpdhat	-2.574157	.5697943	-4.52	0.000	-3.76675	-1.381563
lngdp	.5212927	.1344816	3.88	0.001	.2398194	.802766
_cons	35.93498	7.189835	5.00	0.000	20.88648	50.98347

16 . reg3 (lnst lnpd lnpx2 lnpx3 lnpx4) (lndt lnpd lngdp), ols

Multivariate regression

Equation	Obs	Parms	RMSE	"R-sq"	F-Stat	P
lnst	22	4	.1652258	0.9103	43.14	0.0000
lndt	22	2	.1391259	0.9069	92.53	0.0000

	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
<b>lnst</b>						
lnpd	-1.111835	.4515147	-2.46	0.019	-2.027549	-.1961207
lnpx2	-.4189546	.1431634	-2.93	0.006	-.7093034	-.1286059
lnpx3	-.9424196	.2585266	-3.65	0.001	-1.466736	-.4181034
lnpx4	-.521346	.3441643	-1.51	0.139	-1.219344	.1766516
_cons	41.4946	3.661911	11.33	0.000	34.0679	48.9213
<b>lndt</b>						
lnpd	-2.181329	.2946999	-7.40	0.000	-2.779008	-1.58365
lngdp	.5776586	.0887536	6.51	0.000	.397658	.7576593
_cons	31.03578	3.761201	8.25	0.000	23.40771	38.66385

17 . est store ols

18 . reg3 (lnst lnpd lnpx2 lnpx3 lnpx4) (lndt lnpd lngdp), 2sls nodfk inst(lnpx2 lnpx3 lnpx4 lngdp)

Two-stage least-squares regression

Equation	Obs	Parms	RMSE	"R-sq"	F-Stat	P
<b>lnst</b>	22	4	.329951	0.6424	13.81	0.0000
<b>lndt</b>	22	2	.1454858	0.8982	89.20	0.0000

	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
<b>lnst</b>						
lnpd	2.10611	1.926677	1.09	0.282	-1.801371	6.013591
lnpx2	-.7279628	.3026471	-2.41	0.021	-1.34176	-.114166
lnpx3	-1.122146	.464304	-2.42	0.021	-2.063798	-.180494
lnpx4	-1.428722	.7811544	-1.83	0.076	-3.012977	.1555325
_cons	18.59914	14.05113	1.32	0.194	-9.897873	47.09616
<b>lndt</b>						
lnpd	-2.574157	.4046743	-6.36	0.000	-3.394875	-1.75344
lngdp	.5212921	.0955104	5.46	0.000	.327588	.7149961
_cons	35.93499	5.106302	7.04	0.000	25.57893	46.29105

Endogenous variables: lnst lnpd lndt  
 Exogenous variables: lnpx2 lnpx3 lnpx4 lngdp

19 . est store twostage

20 . reg3 (lnst lnpd lnpx2 lnpx3 lnpx4) (lndt lnpd lngdp), 3sls inst(lnpx2 lnpx3 lnpx4 lngdp)

Three-stage least-squares regression

Equation	Obs	Parms	RMSE	"R-sq"	chi2	P
<b>lnst</b>	22	4	.2963642	0.6266	57.47	0.0000
<b>lndt</b>	22	2	.135203	0.8982	178.41	0.0000

	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
<b>Inst</b>						
lnpd	2.171576	1.926095	1.13	0.260	-1.603501	5.946652
lnpx2	-.7990055	.2985983	-2.68	0.007	-1.384247	-.2137635
lnpx3	-1.329743	.4560002	-2.92	0.004	-2.223487	-.4359989
lnpx4	-1.171403	.775654	-1.51	0.131	-2.691657	.348851
_cons	17.84948	14.04122	1.27	0.204	-9.670808	45.36976
<b>Indt</b>						
lnpd	-2.574157	.4046743	-6.36	0.000	-3.367304	-1.78101
lngdp	.5212921	.0955104	5.46	0.000	.3340951	.708489
_cons	35.93499	5.106302	7.04	0.000	25.92682	45.94316

Endogenous variables: Inst lnpd Indt  
 Exogenous variables: lnpx2 lnpx3 lnpx4 lngdp

21 . reg3 (Inst lnpd lnpx2 lnpx3 lnpx4) (Indt lnpd lngdp), 3sls ireg3 inst(lnpx2 lnpx3 lnpx4 lngdp)

Iteration 1: tolerance = .1059484  
 Iteration 2: tolerance = .04569793  
 Iteration 3: tolerance = .01846611  
 Iteration 4: tolerance = .00725496  
 Iteration 5: tolerance = .00281814  
 Iteration 6: tolerance = .00108981  
 Iteration 7: tolerance = .00042072  
 Iteration 8: tolerance = .00016231  
 Iteration 9: tolerance = .0000626  
 Iteration 10: tolerance = .00002414  
 Iteration 11: tolerance = 9.310e-06  
 Iteration 12: tolerance = 3.590e-06  
 Iteration 13: tolerance = 1.384e-06  
 Iteration 14: tolerance = 5.339e-07

Three-stage least-squares regression, iterated

Equation	Obs	Parms	RMSE	"R-sq"	chi2	P
<b>Inst</b>	22	4	.3022006	0.6117	54.83	0.0000
<b>Indt</b>	22	2	.135203	0.8982	178.41	0.0000

	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
<b>Inst</b>						
lnpd	2.212666	2.005956	1.10	0.270	-1.718936	6.144268
lnpx2	-.8435967	.3049354	-2.77	0.006	-1.441259	-.2459342
lnpx3	-1.460044	.4623671	-3.16	0.002	-2.366267	-.5538216
lnpx4	-1.009892	.7998393	-1.26	0.207	-2.577548	.557764
_cons	17.37893	14.61488	1.19	0.234	-11.26571	46.02357
<b>Indt</b>						
lnpd	-2.574157	.4046743	-6.36	0.000	-3.367304	-1.78101
lngdp	.5212921	.0955104	5.46	0.000	.3340951	.708489
_cons	35.93499	5.106302	7.04	0.000	25.92682	45.94316

Endogenous variables: Inst lnpd Indt  
 Exogenous variables: lnpx2 lnpx3 lnpx4 lngdp



	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
<b>lnqt</b>						
lnpd	.7752765	1.360145	0.57	0.572	-1.983225	3.533778
lnpx2	-.5909191	.2136549	-2.77	0.009	-1.024231	-.1576068
lnpx3	-.7971771	.3277773	-2.43	0.020	-1.46194	-.132414
lnpx4	-.9608797	.551459	-1.74	0.090	-2.079291	.1575311
_cons	24.95604	9.919453	2.52	0.016	4.838457	45.07362
<b>2lnqt</b>						
lnpd	-2.574157	.4046743	-6.36	0.000	-3.394875	-1.75344
lngdp	.5212921	.0955104	5.46	0.000	.327588	.7149961
_cons	35.93499	5.106302	7.04	0.000	25.57893	46.29105

Endogenous variables: lnqt lnpd  
 Exogenous variables: lnpx2 lnpx3 lnpx4 lngdp

27 . reg3 (lnqt lnpd lnpx2 lnpx3 lnpx4) (lnqt lnpd lngdp), 3sls inst(lnpx2 lnpx3 lnpx4 lngdp)

Three-stage least-squares regression

Equation	Obs	Parms	RMSE	"R-sq"	chi2	P
<b>lnqt</b>	22	4	.2056595	0.7644	81.74	0.0000
<b>2lnqt</b>	22	2	.135203	0.8982	178.41	0.0000

	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
<b>lnqt</b>						
lnpd	.7879239	1.360114	0.58	0.562	-1.877851	3.453699
lnpx2	-.6046439	.2134422	-2.83	0.005	-1.022983	-.1863049
lnpx3	-.8372831	.3273419	-2.56	0.011	-1.478861	-.1957047
lnpx4	-.9111679	.5511692	-1.65	0.098	-1.99144	.1691039
_cons	24.81121	9.918929	2.50	0.012	5.370467	44.25195
<b>2lnqt</b>						
lnpd	-2.574157	.4046743	-6.36	0.000	-3.367304	-1.78101
lngdp	.5212921	.0955104	5.46	0.000	.3340951	.708489
_cons	35.93499	5.106302	7.04	0.000	25.92682	45.94316

Endogenous variables: lnqt lnpd  
 Exogenous variables: lnpx2 lnpx3 lnpx4 lngdp

28 . reg3 (lnqt lnpd lnpx2 lnpx3 lnpx4) (lnqt lnpd lngdp), 3sls ireg3 inst(lnpx2 lnpx3 lnpx4 lngdp)

Iteration 1: tolerance = .02535182  
 Iteration 2: tolerance = .01003058  
 Iteration 3: tolerance = .00390723  
 Iteration 4: tolerance = .00151264  
 Iteration 5: tolerance = .00058419  
 Iteration 6: tolerance = .00022541  
 Iteration 7: tolerance = .00008694  
 Iteration 8: tolerance = .00003353  
 Iteration 9: tolerance = .00001293  
 Iteration 10: tolerance = 4.986e-06  
 Iteration 11: tolerance = 1.923e-06  
 Iteration 12: tolerance = 7.415e-07

Three-stage least-squares regression, iterated

Equation	Obs	Parms	RMSE	"R-sq"	chi2	P
<b>lnqt</b>	22	4	.2063295	0.7629	80.89	0.0000
<b>2lnqt</b>	22	2	.135203	0.8982	178.41	0.0000

	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
<b>lnqt</b>						
lnpd	.795862	1.370509	0.58	0.561	-1.890287	3.482011
lnpx2	-.6132584	.214736	-2.86	0.004	-1.034133	-.1923836
lnpx3	-.8624556	.3291486	-2.62	0.009	-1.507575	-.2173363
lnpx4	-.8799661	.5549316	-1.59	0.113	-1.967612	.2076798
_cons	24.72031	9.994251	2.47	0.013	5.131936	44.30868
<b>2lnqt</b>						
lnpd	-2.574157	.4046743	-6.36	0.000	-3.367304	-1.78101
lngdp	.5212921	.0955104	5.46	0.000	.3340951	.708489
_cons	35.93499	5.106302	7.04	0.000	25.92682	45.94316

Endogenous variables: lnqt lnpd  
 Exogenous variables: lnpx2 lnpx3 lnpx4 lngdp

```
29 . log close
    name: <unnamed>
    log: C:\Users\A\Desktop\426.a2.smcl
    log type: smcl
    closed on: 3 Feb 2021, 18:22:59
```

30 .