

**Guideline Solution**  
**Assignment 2**

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  $GDP_t$

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

**Reduce form models:**

$$\begin{aligned} \ln S_t &= \pi_{10} + \pi_{11} \ln P_{X2t} + \pi_{12} \ln P_{X3t} + \pi_{13} \ln P_{X4t} + \pi_{14} \ln GDP_t + w_{1t} \\ \ln D_t &= \pi_{20} + \pi_{21} \ln P_{X2t} + \pi_{22} \ln P_{X3t} + \pi_{23} \ln P_{X4t} + \pi_{24} \ln GDP_t + w_{2t} \\ \ln P_{Dt} &= \pi_{30} + \pi_{31} \ln P_{X2t} + \pi_{32} \ln P_{X3t} + \pi_{33} \ln P_{X4t} + \pi_{34} \ln GDP_t + w_{3t} \end{aligned}$$

**\*Set up STATA analysis as time-series analysis**

```
. tsset obs
      time variable:  obs, 1986 to 2007
                delta:  1 unit
```

**\*Generate Variable in ln-form**

```
. g lnst=ln(st)
. g lndt=ln(dt)
. g pd=pm+t
. g lnpd=ln(pd)
. g lnpx2=ln(px2)
. g lnpx3=ln(px3)
```

```
. g lnpx4=ln(px4)
```

```
. g lngdp=ln(gdp)
```

**\*Estimate Reduced Form model for Supply model**

```
. reg lnst lnpx2 lnpx3 lnpx4 lngdp
```

Source	SS	df	MS	Number of obs =	22
Model	4.64569724	4	1.16142431	F( 4, 17) =	37.32
Residual	.529104674	17	.031123804	Prob > F =	0.0000
				R-squared =	0.8978
				Adj R-squared =	0.8737
Total	5.17480192	21	.246419139	Root MSE =	.17642

lnst	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
lnpx2	-.4503744	.1515961	-2.97	0.009	-.7702142 -.1305347
lnpx3	-.9242052	.2783356	-3.32	0.004	-1.511442 -.3369685
lnpx4	-.3883793	.4222332	-0.92	0.371	-1.279214 .5024549
lngdp	.3438812	.1913463	1.80	0.090	-.0598242 .7475865
_cons	24.65741	5.309757	4.64	0.000	13.4548 35.86002

**\*Predict lnsthat for Supply model**

```
. predict lnsthat, xb
```

**\*Estimate Reduced Form model for Demand model**

```
. reg lndt lnpx2 lnpx3 lnpx4 lngdp
```

Source	SS	df	MS	Number of obs =	22
Model	3.4026552	4	.850663799	F( 4, 17) =	26.43
Residual	.54721789	17	.032189288	Prob > F =	0.0000
				R-squared =	0.8615
				Adj R-squared =	0.8289
Total	3.94987309	21	.188089195	Root MSE =	.17941

lndt	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
lnpx2	-.4887365	.1541691	-3.17	0.006	-.8140049 -.1634682
lnpx3	-.7243134	.2830597	-2.56	0.020	-1.321517 -.1271097
lnpx4	-.577921	.4293997	-1.35	0.196	-1.483875 .3280333
lngdp	.1265855	.194594	0.65	0.524	-.2839719 .5371429
_cons	27.18614	5.399879	5.03	0.000	15.79339 38.57889

**\*Predict lndthat for Supply model**

```
. predict lndthat, xb
```

**\*Estimate Reduced Form model for Domestic Price model**

```
. 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

**\*Predict lnpdhat for Domestic Price model**

```
. predict lnpdhat, xb
```

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

```
. reg lnst lnpdhat lnpx2 lnpx3 lnpx4
```

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
				R-squared	=	0.8978
				Adj R-squared	=	0.8737
Total	5.17480192	21	.246419139	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

```
. reg lndt lnpdhat lngdp
```

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
				R-squared	=	0.8257
				Adj R-squared	=	0.8073
Total	3.94987309	21	.188089195	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

**\*Alternative Command 2sls**

```
. ivregress 2sls lnst lnpx2 lnpx3 lnpx4 (lnpd= lnpx2 lnpx3 lnpx4 lngdp), first
```

First-stage regressions

```
-----
```

Number of obs	=	22
F( 4, 17)	=	6.76
Prob > F	=	0.0019
R-squared	=	0.6142
Adj R-squared	=	0.5234
Root MSE	=	0.0809

	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

```
Instrumental variables (2SLS) regression          Number of obs   =          22
                                                Wald chi2(4)    =          55.22
                                                Prob > chi2     =          0.0000
                                                R-squared      =          0.6424
                                                Root MSE      =          .29004
```

	lnst	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
<b>lnpd</b>		2.10611	1.926677	1.09	0.274	-1.670107	5.882327
lnpx2		-.7279628	.3026471	-2.41	0.016	-1.32114	-.1347853
lnpx3		-1.122146	.464304	-2.42	0.016	-2.032165	-.212127
lnpx4		-1.428722	.7811544	-1.83	0.067	-2.959757	.1023125
_cons		18.59914	14.05113	1.32	0.186	-8.940569	46.13886

```
Instrumented:  lnpd
Instruments:   lnpx2 lnpx3 lnpx4 lngdp
```

```
. ivregress 2sls lndt lngdp (lnpd= lnpx2 lnpx3 lnpx4 lngdp)
```

```
Instrumental variables (2SLS) regression          Number of obs   =          22
                                                Wald chi2(2)    =         178.41
                                                Prob > chi2     =          0.0000
                                                R-squared      =          0.8982
                                                Root MSE      =          .1352
```

	lndt	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
<b>lnpd</b>		-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

```
Instrumented:  lnpd
Instruments:   lngdp lnpx2 lnpx3 lnpx4
```

- 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?

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

```
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]	
lnst						
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
lndt						
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

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

Two-stage least-squares regression

Equation	Obs	Parms	RMSE	"R-sq"	F-Stat	P
lnst	22	4	.329951	0.6424	10.67	0.0000
lndt	22	2	.1454858	0.8982	77.04	0.0000

	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
lnst						
lnpd	2.10611	2.191774	0.96	0.343	-2.339013	6.551233
lnpx2	-.7279628	.3442892	-2.11	0.041	-1.426214	-.0297119
lnpx3	-1.122146	.528189	-2.12	0.041	-2.193363	-.0509293
lnpx4	-1.428722	.8886357	-1.61	0.117	-3.230959	.3735147
_cons	18.59914	15.98447	1.16	0.252	-13.81886	51.01715
lndt						
lnpd	-2.574157	.4354519	-5.91	0.000	-3.457295	-1.69102
lngdp	.5212921	.1027745	5.07	0.000	.3128558	.7297283
_cons	35.93499	5.494663	6.54	0.000	24.7913	47.07868

Endogenous variables: lnst lnpd lndt

Exogenous variables: lnpx2 lnpx3 lnpx4 lngdp

```
. 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
lnst	22	4	.2963642	0.6266	57.47	0.0000
lndt	22	2	.135203	0.8982	178.41	0.0000

	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
lnst						
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
lndt						

```

      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:  lnst lnpd lndt
Exogenous variables:  lnp2 lnp3 lnp4 lngdp
-----

```

```

. reg3 (lnst lnpd lnp2 lnp3 lnp4) (lndt lnpd lngdp), 3sls ireg3 inst(lnp2
lnp3 lnp4 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   Parns      RMSE      "R-sq"      chi2      P
-----
lnst          22     4      .3022006    0.6117      54.83    0.0000
lndt          22     2      .135203    0.8982     178.41    0.0000
-----

```

```

-----
          |      Coef.   Std. Err.      z    P>|z|      [95% Conf. Interval]
-----+-----
lnst     |
      lnpd |  2.212666   2.005956     1.10  0.270   -1.718936   6.144268
      lnp2 | -1.8435967  .3049354    -2.77  0.006   -1.441259  -.2459342
      lnp3 | -1.460044   .4623671    -3.16  0.002   -2.366267  -.5538216
      lnp4 | -1.009892   .7998393    -1.26  0.207   -2.577548   .557764
      _cons |  17.37893   14.61488     1.19  0.234   -11.26571  46.02357
-----+-----
lndt     |
      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:  lnst lnpd lndt
Exogenous variables:  lnp2 lnp3 lnp4 lngdp
-----

```

Concerning on the asymptotic property, 3SLS or I3SLS is the most appropriated model. Because if there exists correlation of the error terms across equations, 3SLS or I3SLS will lead to a more asymptotically efficient estimators.

However, if there exists specification error problem in any model(s), the specification error will spread throughout the whole system.

$\beta_{21}$  represents price elasticity of quantity demand.

$\beta_{22}$  represents income elasticity of quantity demand.

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  $t$ .

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

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

- 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.

**. \* Generate lnqt**

```
. g lnqt=lnst if lnst<lnst
(22 missing values generated)
```

```
. replace lnqt=lnst if lnst<lnst
(22 real changes made)
```

**. \* Estimate model (3) and (4) using OLS**

```
. reg3 (lnqt lnpd lnpx2 lnpx3 lnpx4) (lnqt lnpd lngdp), ols
```

Multivariate regression

Equation	Obs	Parms	RMSE	"R-sq"	F-Stat	P
lnqt	22	4	.136235	0.9201	48.95	0.0000
2lnqt	22	2	.1391259	0.9069	92.53	0.0000

	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
lnqt						
lnpd	-1.353506	.3722912	-3.64	0.001	-2.108548	-.5984645
lnpx2	-.3864994	.1180437	-3.27	0.002	-.6259031	-.1470957
lnpx3	-.6782817	.2131651	-3.18	0.003	-1.110601	-.2459629
lnpx4	-.3606189	.2837767	-1.27	0.212	-.9361448	.2149069
_cons	40.10218	3.019386	13.28	0.000	33.97858	46.22578
2lnqt						
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

**. \* Estimate model (3) and (4) using 2SLS**

```
. reg3 (lnqt lnpd lnpx2 lnpx3 lnpx4) (lnqt lnpd lngdp), 2sls inst(lnpx2 lnpx3 lnpx4 lngdp)
```

Two-stage least-squares regression

Equation	Obs	Parms	RMSE	"R-sq"	F-Stat	P
lnqt	22	4	.2329302	0.7665	15.68	0.0000
2lnqt	22	2	.1454858	0.8982	77.04	0.0000

	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
--	-------	-----------	---	------	----------------------	--

```

-----+-----
lnqt |
  lnpd | .7752765 1.547291 0.50 0.619 -2.362776 3.913329
  lnpx2 | -.5909191 .2430523 -2.43 0.020 -1.083852 -.0979861
  lnpx3 | -.7971771 .3728771 -2.14 0.039 -1.553407 -.0409473
  lnpx4 | -.9608797 .6273359 -1.53 0.134 -2.233176 .3114165
  _cons | 24.95604 11.2843 2.21 0.033 2.07042 47.84166
-----+-----
2lnqt |
  lnpd | -2.574157 .4354519 -5.91 0.000 -3.457295 -1.69102
  lngdp | .5212921 .1027745 5.07 0.000 .3128558 .7297283
  _cons | 35.93499 5.494663 6.54 0.000 24.7913 47.07868
-----+-----
Endogenous variables: lnqt lnpd
Exogenous variables: lnpx2 lnpx3 lnpx4 lngdp
-----+-----

```

**. \* Estimate model (3) and (4) using 3SLS**

```

. 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
lnqt	22	4	.2056595	0.7644	81.74	0.0000
2lnqt	22	2	.135203	0.8982	178.41	0.0000

```

-----+-----
          |      Coef.   Std. Err.   z   P>|z|   [95% Conf. Interval]
-----+-----
lnqt |
  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
-----+-----
2lnqt |
  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
-----+-----

```

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

In term of economic and econometric concept, the problem in **d** is identification problem. Unfortunately, the models are identified system in this case. Additionally, the estimated results of coefficient of price in demand and supply models both have correct signs. However, the equilibrium condition assumption is the most important assumption that make the models become simultaneous equation model. If the equilibrium condition assumption doesn't hold, the models will be incorrect.