

**BOSTON COLLEGE**

Department of Economics

EC 771: Econometrics

Spring 2010

Prof. Baum, Mr. Dmitriev

PROBLEM SET 2: SOLUTIONS

Problem 1 (20 points). Solutions.

We have the following system of equations:

$$\begin{aligned}z_1 &= x_1 - 2x_2, \\z_2 &= x_2 + 4x_3, \\z_3 &= 2x_1 - 3x_2 + 5x_3,\end{aligned}$$

where  $x_i, z_j$  are vectors. Clearly, we can write this system as follows:

$$Z = XA$$

where

$$Z = (z_1 \ z_2 \ z_3), \quad X = (x_1 \ x_2 \ x_3), \quad A = \begin{pmatrix} 1 & 0 & 2 \\ -2 & 1 & -3 \\ 0 & 4 & 5 \end{pmatrix}$$

Next, we need to solve for  $x_i$  in terms of  $z_j$ . Using some linear algebra technique (substitution, Gaussian elimination, Cramer's rule), one obtains the following system:

$$\begin{aligned}x_1 &= 17z_1 + 10z_2 - 8z_3, \\x_2 &= 8z_1 + 5z_2 - 4z_3, \\x_3 &= -2z_1 - z_2 + z_3\end{aligned}$$

Thus, we have that

$$A^{-1} = \begin{pmatrix} 17 & 8 & -2 \\ 10 & 5 & -1 \\ -8 & -4 & 1 \end{pmatrix}$$

Now, recall the solution to question #3 in Problem Set 1. We determined there that the residuals one obtains from a regression will not be changed if the regressors are linearly transformed by an invertible matrix. Thus, the two regressions (on  $x_i$  and on  $z_j$ ) produce the same residuals, and hence the same predicted values.

Finally, since the fitted values are the same for the two regressions,  $X\hat{\beta} = Z\hat{\alpha} = XA\hat{\alpha} \implies \hat{\beta} = A\hat{\alpha}$ . Thus,  $\hat{\beta}_1 = \hat{\alpha}_1 + 2\hat{\alpha}_3$ . It is also the case that  $\hat{\alpha} = A^{-1}\hat{\beta} \implies \hat{\alpha}_1 = 17\hat{\beta}_1 + 8\hat{\beta}_2 - 2\hat{\beta}_3$ .

Problem 2 (10 points). Solutions.

. use <http://fmwww.bc.edu/ec-p/data/greene2008/tbrate>

. regress D.r L.pi LD.y LD.r L2D.r

Source	SS	df	MS	Number of obs =	185
Model	22.1971507	4	5.54928768	F( 4, 180) =	6.99
Residual	142.934504	180	.794080577	Prob > F =	0.0000
				R-squared =	0.1344
				Adj R-squared =	0.1152
Total	165.131655	184	.897454645	Root MSE =	.89111

D.r	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
pi						
L1.	.0160647	.0200335	0.80	0.424	-.023466	.0555955
y						
LD.	18.38055	5.758924	3.19	0.002	7.016859	29.74423
r						
LD.	.2374557	.0740703	3.21	0.002	.0912979	.3836135
L2D.	-.1540175	.0725383	-2.12	0.035	-.2971523	-.0108828
_cons	-.2319403	.1256143	-1.85	0.066	-.4798063	.0159256

. predict rhat  
(option xb assumed; fitted values)  
(3 missing values generated)

. predict uhat, residuals  
(3 missing values generated)

. twoway (connected rhat yq, msize(vsmall)) (line uhat yq)

. reg uhat rhat

Source	SS	df	MS	Number of obs =	185
Model	5.6843e-14	1	5.6843e-14	F( 1, 183) =	0.00
Residual	142.934505	183	.781062871	Prob > F =	1.0000
				R-squared =	0.0000
				Adj R-squared =	-0.0055
Total	142.934505	184	.776817964	Root MSE =	.88378

uhat	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
rhat	7.53e-09	.1875834	0.00	1.000	-.3701043	.3701043
_cons	1.93e-10	.0650252	0.00	1.000	-.1282955	.1282955

Since the residuals are by construction orthogonal to the fitted values, we verify via the above OLS regression that the mean of the residuals is zero and that the fitted values are uncorrelated with the residuals.

. reg rhat uhat

Source	SS	df	MS	Number of obs = 185		
Model	0	1	0	F( 1, 183) =	0.00	
Residual	22.1971507	183	.121295905	Prob > F =	1.0000	
Total	22.1971507	184	.120636688	R-squared =	0.0000	
				Adj R-squared =	-0.0055	
				Root MSE =	.34828	

rhat	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
uhat	1.17e-09	.0291309	0.00	1.000	-.0574757	.0574757
_cons	.0133946	.0256057	0.52	0.602	-.0371258	.063915

As explained earlier, the residuals and the fitted values are uncorrelated, which is borne out by the above OLS regression results. Additionally, we obtain a non-zero constant that is the mean of the fitted values, which is also the mean of the dependent variable in the original regression, since the residuals are constructed to be mean-zero.

### Problem 3 (10 points). Solutions.

. regress D.r LD.y LD.r L2D.r

Source	SS	df	MS	Number of obs = 185		
Model	21.6865324	3	7.22884414	F( 3, 181) =	9.12	
Residual	143.445122	181	.792514488	Prob > F =	0.0000	
Total	165.131655	184	.897454645	R-squared =	0.1313	
				Adj R-squared =	0.1169	
				Root MSE =	.89023	

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

y						
LD.	17.47536	5.641639	3.10	0.002	6.343518	28.6072
r						
LD.	.2437294	.0735833	3.31	0.001	.0985381	.3889208
L2D.	-.1471644	.071962	-2.05	0.042	-.2891568	-.0051721
_cons	-.158009	.0852318	-1.85	0.065	-.3261847	.0101668

. predict ehat, residuals  
(3 missing values generated)

. regress L.pi LD.y LD.r L2D.r

Source	SS	df	MS		Number of obs =	185
Model	127.410931	3	42.4703102		F( 3, 181) =	3.89
Residual	1978.56248	181	10.9312844		Prob > F =	0.0101
					R-squared =	0.0605
					Adj R-squared =	0.0449
Total	2105.97341	184	11.4455077		Root MSE =	3.3062

L.pi	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
y						
LD.	-56.34641	20.95257	-2.69	0.008	-97.68912	-15.00371
r						
LD.	.3905278	.2732821	1.43	0.155	-.1487006	.9297562
L2D.	.4265952	.2672608	1.60	0.112	-.1007524	.9539427
_cons	4.602094	.3165437	14.54	0.000	3.977503	5.226684

. predict vhat, residuals  
(3 missing values generated)

. regress ehat vhat

Source	SS	df	MS		Number of obs =	185
Model	.510618184	1	.510618184		F( 1, 183) =	0.65
Residual	142.934503	183	.781062856		Prob > F =	0.4198
					R-squared =	0.0036
					Adj R-squared =	-0.0019
Total	143.445121	184	.779593048		Root MSE =	.88378

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

vhat		.0160647	.0198686	0.81	0.420	-.0231363	.0552658
_cons		4.07e-11	.0649766	0.00	1.000	-.1281996	.1281996

---

The value of the coefficient on `vhat` is the same as that of  $\pi_{t-1}$  in the original regression in the previous problem, a consequence of the Frisch-Waugh Theorem (also known as the Frisch-Waugh-Lovell Theorem). See Greene 3.3 for details.

Problem 4 (10 points). Solutions.

That the covariance matrix of  $\mathbf{b}$  is positive semidefinite implies that the determinant of this matrix is non-negative. Given that

$$\text{Var}(\mathbf{b}) = \begin{pmatrix} \text{Var}(b_1) & \text{Cov}(b_1, b_2) \\ \text{Cov}(b_1, b_2) & \text{Var}(b_2) \end{pmatrix}$$

the non-negativity of the matrix implies that

$$\text{Var}(b_1)\text{Var}(b_2) - (\text{Cov}(b_1, b_2))^2 \geq 0 \implies (\text{Cov}(b_1, b_2))^2 \leq \text{Var}(b_1)\text{Var}(b_2)$$

Next,  $\rho(b_1, b_2)$ , the correlation between  $b_1$  and  $b_2$  is defined by

$$\rho(b_1, b_2) \equiv \frac{\text{Cov}(b_1, b_2)}{\sqrt{\text{Var}(b_1)\text{Var}(b_2)}}$$

Then, from the above inequality, it must be that case that  $\rho(b_1, b_2)^2 \leq 1 \implies -1 \leq \rho(b_1, b_2) \leq 1$ , as required.

Problem 5 (10 points). Solutions.

The restriction implies that  $\beta_3 = 1 - \beta_2$ . Hence,  $y_t = \beta_1 + \beta_2 x_{t2} + (1 - \beta_2)x_{t3} + u_t \implies y_t - x_{t3} = \beta_1 + \beta_2(x_{t2} - x_{t3}) + u_t$ . Then, define the new variables  $\tilde{y}_t = y_t - x_{t3}$  and  $z_t = x_{t2} - x_{t3}$ . The restricted regression is  $\tilde{y}_t = \beta_1 + \beta_2 z_t + u_t$ . The estimate  $\hat{\beta}_2$  is obtained directly, and, since the estimate of  $\beta_3$  must satisfy the restriction, we see that  $\hat{\beta}_3 = 1 - \hat{\beta}_2$ .

We can add a third term to the right-hand side of restricted regression equation in order to obtain a model that is equivalent to the original one. This yields

$$\hat{y}_t = \beta_1 + \beta_2 z_t + (\beta_2 + \beta_3 - 1)x_{t3} + u_t$$

Thus, if we run the regression

$$\hat{y}_t = \beta_1 + \beta_2 z_t + \gamma x_{t3} + u_t$$

we will obtain estimates for  $\beta_1$  and  $\beta_2$  directly, and we can obtain an estimate of  $\beta_3$  by using the relation  $\hat{\beta}_3 = \hat{\gamma} + 1 - \hat{\beta}_2$ . If the restriction held exactly in the data, the estimate of  $\gamma$  would be zero.

Note that we could have eliminated  $\beta_2$  instead of  $\beta_3$  in the restricted model, and furthermore obtained an appropriately adjusted counterpart to the regression with a zero coefficient when the restriction holds in the data.

Problem 6 (40 points). Solutions.

a)(5 points)

```
. xi i.year  
i.year          _Iyear_66-73      (naturally coded; _Iyear_66 omitted)
```

```
. ivreg2 lw expr s (iq = age kww med)
```

IV (2SLS) estimation

```
-----  
Number of obs =      758  
F( 3, 754) =    105.26  
Prob > F      =     0.0000  
Centered R2   =     0.2886  
Uncentered R2 =     0.9960  
Root MSE     =     .3616  
Total (centered) SS      = 139.2861498  
Total (uncentered) SS   = 24652.24662  
Residual SS            =  99.0915462
```

```
-----  
      lw |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]  
-----+-----  
      iq |   -.0012932   .0047482    -0.27   0.785    - .0105995   .0080132  
     expr |    .0442341   .0065777     6.72   0.000     .0313421   .057126  
        s |    .1107632   .0157675     7.02   0.000     .0798595   .1416668  
   _cons |    4.259495   .3124346    13.63   0.000     3.647134   4.871855  
-----
```

```
Anderson canon. corr. LR statistic (underidentification test):      43.846  
Chi-sq(3) P-val =      0.0000  
-----
```

```
Cragg-Donald F statistic (weak identification test):      14.927  
Stock-Yogo weak ID test critical values: 5% maximal IV relative bias 13.91  
10% maximal IV relative bias 9.08  
20% maximal IV relative bias 6.46  
30% maximal IV relative bias 5.39  
10% maximal IV size 22.30  
15% maximal IV size 12.83  
20% maximal IV size 9.54  
25% maximal IV size 7.80
```

Source: Stock-Yogo (2005). Reproduced by permission.

```
-----  
Sargan statistic (overidentification test of all instruments):      84.806  
Chi-sq(2) P-val =      0.0000  
-----
```

```
Instrumented:      iq  
Included instruments: expr s  
Excluded instruments: age kww med  
-----
```

The Anderson canonical correlation test rejects at the 5% level the null hypothesis of underidentification. However, the rejection of the null of the Sargan test suggests that one or more of the instruments is not uncorrelated with the disturbance process.

b) (5 points)

```
. ivreg2 lw expr s _Iyear_67 _Iyear_68 _Iyear_69 _Iyear_70 _Iyear_71 _Iyear_73 (iq = age kww med)
```

IV (2SLS) estimation

-----

		Number of obs =	758	
		F( 9, 748) =	47.13	
		Prob > F =	0.0000	
Total (centered) SS	=	139.2861498	Centered R2 =	0.3621
Total (uncentered) SS	=	24652.24662	Uncentered R2 =	0.9964
Residual SS	=	88.85241753	Root MSE =	.3424

lw	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
iq	.007033	.0040735	1.73	0.084	-.0009509	.0150169
expr	.0398175	.0067903	5.86	0.000	.0265086	.0531263
s	.0565379	.0139059	4.07	0.000	.0292829	.0837929
_Iyear_67	-.0725177	.0497367	-1.46	0.145	-.1699999	.0249644
_Iyear_68	.0504323	.0465702	1.08	0.279	-.0408436	.1417082
_Iyear_69	.1605229	.045594	3.52	0.000	.0711604	.2498854
_Iyear_70	.2097466	.053631	3.91	0.000	.1046318	.3148614
_Iyear_71	.183241	.0456348	4.02	0.000	.0937985	.2726836
_Iyear_73	.2792134	.0420477	6.64	0.000	.1968014	.3616254
_cons	4.013944	.2761018	14.54	0.000	3.472795	4.555094

Anderson canon. corr. LR statistic (underidentification test): 54.386  
 Chi-sq(3) P-val = 0.0000

Cragg-Donald F statistic (weak identification test): 18.497  
 Stock-Yogo weak ID test critical values:

5% maximal IV relative bias	13.91
10% maximal IV relative bias	9.08
20% maximal IV relative bias	6.46
30% maximal IV relative bias	5.39
10% maximal IV size	22.30
15% maximal IV size	12.83
20% maximal IV size	9.54
25% maximal IV size	7.80

Source: Stock-Yogo (2005). Reproduced by permission.

Sargan statistic (overidentification test of all instruments): 91.950  
 Chi-sq(2) P-val = 0.0000

-----  
 Instrumented: iq  
 Included instruments: expr s \_Iyear\_67 \_Iyear\_68 \_Iyear\_69 \_Iyear\_70 \_Iyear\_71  
 \_Iyear\_73  
 Excluded instruments: age kww med  
 -----

The year dummies for years after 1968 are all significant and positive, suggesting some unmodeled change in the underlying process determining the wage that isn't captured by the included characteristics of workers. IQ now has a positive coefficient, but one that is still not statistically significantly different from 0 at the 5% level. The Anderson test and the Sargan test produce similar results as in part a).

c) (5 points)

```
. ivreg2 lw expr s _Iyear_67 _Iyear_68 _Iyear_69 _Iyear_70 _Iyear_71 _Iyear_73 (iq = age kww med),
robust
```

IV (2SLS) estimation

-----  
 Statistics robust to heteroskedasticity

		Number of obs =	758
		F( 9, 748) =	42.35
		Prob > F =	0.0000
Total (centered) SS	=	Centered R2	= 0.3621
Total (uncentered) SS	=	Uncentered R2	= 0.9964
Residual SS	=	Root MSE	= .3424

	lw	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
iq		.007033	.004181	1.68	0.093	-.0011616	.0152276
expr		.0398175	.0068121	5.85	0.000	.0264659	.053169
s		.0565379	.0141939	3.98	0.000	.0287185	.0843574
_Iyear_67		-.0725177	.0474303	-1.53	0.126	-.1654794	.0204439
_Iyear_68		.0504323	.046312	1.09	0.276	-.0403376	.1412021
_Iyear_69		.1605229	.0426472	3.76	0.000	.0769361	.2441098
_Iyear_70		.2097466	.0563248	3.72	0.000	.099352	.3201412
_Iyear_71		.183241	.0433592	4.23	0.000	.0982585	.2682235
_Iyear_73		.2792134	.0420768	6.64	0.000	.1967443	.3616824
_cons		4.013944	.285412	14.06	0.000	3.454547	4.573341



Anderson canon. corr. LR statistic (underidentification test): 54.386  
Chi-sq(3) P-val = 0.0000

Test statistic(s) not robust

-----  
Cragg-Donald F statistic (weak identification test): 18.497  
Stock-Yogo weak ID test critical values: 5% maximal IV relative bias 13.91  
10% maximal IV relative bias 9.08  
20% maximal IV relative bias 6.46  
30% maximal IV relative bias 5.39  
10% maximal IV size 22.30  
15% maximal IV size 12.83  
20% maximal IV size 9.54  
25% maximal IV size 7.80

Test statistic(s) not robust

Source: Stock-Yogo (2005). Reproduced by permission.

-----  
Hansen J statistic (overidentification test of all instruments): 72.328  
Chi-sq(2) P-val = 0.0000

-----  
Instrumented: iq  
Included instruments: expr s \_Iyear\_67 \_Iyear\_68 \_Iyear\_69 \_Iyear\_70 \_Iyear\_71  
\_Iyear\_73  
Excluded instruments: age kww med  
-----

Using robust standard errors does not seem to affect the standard errors very much, suggesting that heteroskedasticity is not an issue.

d)(10 points)

```
. ivreg2 lw expr s _Iyear_67 _Iyear_68 _Iyear_69 _Iyear_70 _Iyear_71 _Iyear_73 (iq = age kww med),  
gmm
```

2-Step GMM estimation

-----  
Statistics robust to heteroskedasticity

		Number of obs =	758
		F( 9, 748) =	41.49
		Prob > F =	0.0000
Total (centered) SS	=	Centered R2 =	0.3562
Total (uncentered) SS	=	Uncentered R2 =	0.9964
Residual SS	=	Root MSE =	.344

-----

lw	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
iq	.0077785	.004178	1.86	0.063	-.0004103	.0159673
expr	.0457042	.0067758	6.75	0.000	.032424	.0589845
s	.0550396	.0141922	3.88	0.000	.0272233	.0828558
_Iyear_67	-.0602284	.0474059	-1.27	0.204	-.1531422	.0326854
_Iyear_68	.0569231	.0463035	1.23	0.219	-.0338301	.1476763
_Iyear_69	.1601399	.0426317	3.76	0.000	.0765833	.2436965
_Iyear_70	.1794522	.0561917	3.19	0.001	.0693184	.289586
_Iyear_71	.1548847	.04323	3.58	0.000	.0701555	.2396139
_Iyear_73	.2763517	.0420029	6.58	0.000	.1940274	.358676
_cons	3.940908	.285036	13.83	0.000	3.382248	4.499568

Anderson canon. corr. LR statistic (underidentification test): 54.386  
Chi-sq(3) P-val = 0.0000

Test statistic(s) not robust

Cragg-Donald F statistic (weak identification test): 18.497  
Stock-Yogo weak ID test critical values: 5% maximal IV relative bias 13.91  
10% maximal IV relative bias 9.08  
20% maximal IV relative bias 6.46  
30% maximal IV relative bias 5.39  
10% maximal IV size 22.30  
15% maximal IV size 12.83  
20% maximal IV size 9.54  
25% maximal IV size 7.80

Test statistic(s) not robust

Source: Stock-Yogo (2005). Reproduced by permission.

Hansen J statistic (overidentification test of all instruments): 72.328  
Chi-sq(2) P-val = 0.0000

Instrumented: iq  
Included instruments: expr s \_Iyear\_67 \_Iyear\_68 \_Iyear\_69 \_Iyear\_70 \_Iyear\_71  
\_Iyear\_73  
Excluded instruments: age kww med

None of the results are markedly different from that obtained in part b). In the GMM model we estimate here, we do not maintain the assumption of conditional homoskedasticity, but rather allow arbitrary heteroskedasticity. The GMM model also delivers efficient estimates. The Hansen *J* statistic allows a test of overidentification similar to that provided by the Sargan statistic in the 2SLS model; the Hansen *J* is consistent in the presence of heteroskedasticity. The rejection of the null in this test suggests that one or more of the instruments is not uncorrelated with the disturbance process. The Anderson test as before indicates that the model is not underidentified.

e) (5 points)

```
. ivreg2 lw expr _Iyear_67 _Iyear_68 _Iyear_69 _Iyear_70 _Iyear_71 _Iyear_73 (s iq = age kww med),  
gmm endog(s)
```

2-Step GMM estimation

-----  
Statistics robust to heteroskedasticity

	Number of obs =	758
	F( 9, 748) =	37.83
	Prob > F =	0.0000
Total (centered) SS =	Centered R2 =	0.0906
Total (uncentered) SS =	Uncentered R2 =	0.9949
Residual SS =	Root MSE =	.4088

```
-----  
      |           Robust  
      |           Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]  
-----+-----  
      s |   .1993476   .0254187    7.84  0.000    .1495279   .2491674  
      iq |  -.0089693   .0054021   -1.66  0.097   -.0195573   .0016187  
      expr |   .0630694   .0081395    7.75  0.000    .0471162   .0790225  
      _Iyear_67 |  -.0753593   .0560256   -1.35  0.179   -.1851675   .0344488  
      _Iyear_68 |   .012483    .0531677    0.23  0.814   -.0917237   .1166897  
      _Iyear_69 |   .0967016   .050023    1.93  0.053   -.0013417   .1947449  
      _Iyear_70 |   .1450002   .0670161    2.16  0.030    .013651    .2763494  
      _Iyear_71 |   .0198738   .0584071    0.34  0.734   -.094602    .1343495  
      _Iyear_73 |  -.0100273   .0670913   -0.15  0.881   -.1415238   .1214693  
      _cons |   3.81719    .3332255   11.46  0.000    3.16408    4.4703  
-----
```

Anderson canon. corr. LR statistic (underidentification test): 45.115  
Chi-sq(2) P-val = 0.0000

Test statistic(s) not robust

```
-----  
Cragg-Donald F statistic (weak identification test): 15.270  
Stock-Yogo weak ID test critical values: 10% maximal IV size 13.43  
15% maximal IV size 8.18  
20% maximal IV size 6.40  
25% maximal IV size 5.45
```

Test statistic(s) not robust

Source: Stock-Yogo (2005). Reproduced by permission.

```
-----  
Hansen J statistic (overidentification test of all instruments): 0.482  
Chi-sq(1) P-val = 0.4873
```

```

-endog- option:
Endogeneity test of endogenous regressors:          71.528
                                                    Chi-sq(1) P-val = 0.0000

Regressors tested:    s

```

```

-----
Instrumented:          s iq
Included instruments:  expr _Iyear_67 _Iyear_68 _Iyear_69 _Iyear_70 _Iyear_71
                     _Iyear_73
Excluded instruments: age kww med
-----

```

The endogeneity test rejects the null hypothesis of exogeneity of the variable *s*, years of schooling.

f)(5 points)

```

. ivreg2 lw expr _Iyear_67 _Iyear_68 _Iyear_69 _Iyear_70 _Iyear_71 _Iyear_73 (s iq = age kww med),
  gmm

```

2-Step GMM estimation

Statistics robust to heteroskedasticity

```

-----
Total (centered) SS      = 139.2861498
Total (uncentered) SS   = 24652.24662
Residual SS              = 126.6665339

Number of obs = 758
F( 9, 748) = 37.83
Prob > F = 0.0000
Centered R2 = 0.0906
Uncentered R2 = 0.9949
Root MSE = .4088

```

```

-----
          |           Robust
          |           Coef.   Std. Err.   z   P>|z|   [95% Conf. Interval]
-----+-----
          s |   .1993476   .0254187   7.84   0.000   .1495279   .2491674
          iq |  -.0089693   .0054021  -1.66   0.097  -.0195573   .0016187
          expr |   .0630694   .0081395   7.75   0.000   .0471162   .0790225
    _Iyear_67 |  -.0753593   .0560256  -1.35   0.179  -.1851675   .0344488
    _Iyear_68 |   .012483   .0531677   0.23   0.814  -.0917237   .1166897
    _Iyear_69 |   .0967016   .050023   1.93   0.053  -.0013417   .1947449
    _Iyear_70 |   .1450002   .0670161   2.16   0.030   .013651   .2763494
    _Iyear_71 |   .0198738   .0584071   0.34   0.734  -.094602   .1343495
    _Iyear_73 |  -.0100273   .0670913  -0.15   0.881  -.1415238   .1214693
          _cons |   3.81719   .3332255  11.46   0.000   3.16408   4.4703

```

```

-----
Anderson canon. corr. LR statistic (underidentification test):          45.115

```

Chi-sq(2) P-val = 0.0000

Test statistic(s) not robust

---

Cragg-Donald F statistic (weak identification test): 15.270  
Stock-Yogo weak ID test critical values: 10% maximal IV size 13.43  
15% maximal IV size 8.18  
20% maximal IV size 6.40  
25% maximal IV size 5.45

Test statistic(s) not robust

Source: Stock-Yogo (2005). Reproduced by permission.

---

Hansen J statistic (overidentification test of all instruments): 0.482  
Chi-sq(1) P-val = 0.4873

---

Instrumented: s iq  
Included instruments: expr \_Iyear\_67 \_Iyear\_68 \_Iyear\_69 \_Iyear\_70 \_Iyear\_71  
\_Iyear\_73  
Excluded instruments: age kww med

---

Unlike for the previous regression models, the Hansen *J* test fails to reject the null hypothesis of instruments uncorrelated with the disturbance process, suggesting, together with the successful rejection of underidentification via the Anderson test, that the instrument set and endogenous variables set used are valid.