Stata Textbook Examples

Introductory Econometrics: A Modern Approach by Jeffrey M. Wooldridge (1st & 2nd eds.)

Chapter 15 - Instrumental Variables Estimation and Two Stage Least Squares

Example 15.1: Estimating the Return to Education for Married Women

use http://fmwww.bc.edu/ec-p/data/wooldridge/MROZ

reg lwage educ

Source	SS	df	MS		Number of obs	
Model Residual Total	26.3264237 197.001028 	426 	26.3264237 .462443727 .523015108		R-squared Adj R-squared	= 0.0000 = 0.1179
lwage	Coef.	Std. E	rr. t	P> t	[95% Conf.	Interval]
educ _cons	.1086487 1851969	.01439		0.000	.0803451 5492674	.1369523

ivreg lwage (educ = fatheduc)

educ作为fathedu的 工具变量。

Instrumental variables (2SLS) regression

Source	SS	df	MS		Number of obs F(1, 426)		428 2.84
Model Residual + Total	20.8673618 202.460089 223.327451	426 .	20.8673618 .475258426 		Prob > F R-squared Adj R-squared Root MSE	= = =	0.0929 0.0934 0.0913 .68939
lwage	Coef.	Std. Er	cr. t	P> t	[95% Conf.	 Int	cerval]
educ _cons	.0591735	.035141		0.093 0.323	0098994 4357311		L282463 .317938

Instrumented: educ
Instruments: fatheduc

Example 15.2: Estimating the Return to Education for Men

use http://fmwww.bc.edu/ec-p/data/wooldridge/WAGE2

ivreg lwage (educ = sibs)

Instrumental variables (2SLS) regression

Source	SS	df		MS		Number of obs F(1, 933)	
Model Residual	-1.5197389 167.176033	1 933		197389 181172		Prob > F R-squared Adj R-squared	= 0.0000 = .
Total	165.656294	934	.177	362199		Root MSE	= .4233
lwage	Coef.	Std.	Err.	t	P> t	[95% Conf.	Interval]
educ _cons	.1224327 5.130026	.0263		4.65 14.44	0.000	.0707194 4.432999	.1741459 5.827053
Instrumented: Instruments:	educ sibs						

reg lwage educ

Source	SS	df	MS		Number of obs F(1, 933)	
Model Residual + Total	16.1377074 149.518587 165.656294	933 934	16.1377074 .16025572 .177362199		Prob > F R-squared Adj R-squared	= 0.0000 = 0.0974
lwage	Coef.	Std. E	Err. t	P> t	[95% Conf.	Interval]
educ _cons	.0598392 5.973062	.00596			.0481366	.0715418

Example 15.3: Estimating the Effect of Smoking on Birth Weight

use http://fmwww.bc.edu/ec-p/data/wooldridge/BWGHT

ivreg lbwght (packs = cigprice), first

First-stage regressions

Source | SS df MS Number of obs = 1388

2.988674 8.698884 0.34 0.731 -14.07573 20.05307 packs

4.448137 .9081547 4.90 0.000 2.66663 cons

Instrumented: packs Instruments: cigprice

Example 15.4: Using College Proximity as an IV for Education

use http://fmwww.bc.edu/ec-p/data/wooldridge/CARD

educ = nearc4) exper expersq black smsa south, first ivreg lwage

First-stage regressions educ是工具变量。

SS df Number of obs = 3010Source | MS F(6, 3003) =451.87 10230.4843 6 1705.08072 Prob > F 0.0000 Model | R-squared = Residual 11331.5958 3003 3.77342516 0.4745 Adj R-squared = 0.4734 21562.0801 1.9425 Total | 3009 7.16586243 Root MSE

ed	uc	Coef.	Std. Err.	t	P> t	[95% Conf.	Interval]
exp	er	4100081	.0336939	-12.17	0.000	4760735	3439427
exper	sq	.0007323	.0016499	0.44	0.657	0025029	.0039674
bla	ck	-1.006138	.0896454	-11.22	0.000	-1.181911	8303656
sm	sa	.4038769	.0848872	4.76	0.000	.2374339	.5703199
sou	th	291464	.0792247	-3.68	0.000	4468042	1361238
near	c4	.3373208	.0825004	4.09	0.000	.1755577	.4990839
_co	ns	16.65917	.1763889	94.45	0.000	16.31332	17.00503

Instrumental variables (2SLS) regression

Source	SS	df	MS		Number of obs F(6, 3003)	= 3010 = 120.83
Model Residual	133.463217 459.178394		2438695 2906558 		Prob > F R-squared Adj R-squared	= 0.0000 = 0.2252
Total	592.641611	3009 .19	6956335		Root MSE	= .39103
lwage	Coef.	 Std. Err.	 t	P> t	 [95% Conf.	 Interval]
educ exper expersq	.1322888 .107498 0022841	.0492332 .0213006 .0003341	2.69 5.05 -6.84	0.007 0.000 0.000	.0357545 .0657327 0029392	.228823 .1492632 0016289
black	130802	.0528723	-2.47	0.013	2344716	0271324
smsa south _cons	.1313237 1049005 3.752783	.0301298 .0230731 .8293408	4.36 -4.55 4.53	0.000 0.000 0.000	.0722465 1501412 2.126649	.1904009 0596599 5.378916

Instrumented: educ

Instruments: exper expersq black smsa south nearc4

reg lwage educ exper expersq black smsa south

Source	SS	df	MS		Number of obs = $F(6, 3003) =$	3010 204.93
Model Residual Total	172.165615 420.475997 592.641611	3003	28.6942691 .140018647 .196956335		Prob > F = R-squared = Adj R-squared =	0.0000 0.2905 0.2891 .37419
lwage	Coef.	Std. E		P> t	 [95% Conf. In	
educ	.074009	.00350	54 21.11	0.000	.0671357 .	0808823

exper	.0835958	.0066478	12.57	0.000	.0705612	.0966305
expersq	0022409	.0003178	-7.05	0.000	0028641	0016177
black	1896316	.0176266	-10.76	0.000	2241929	1550702
smsa	.161423	.0155733	10.37	0.000	.1308876	.1919583
south	1248615	.0151182	-8.26	0.000	1545046	0952184
_cons	4.733664	.0676026	70.02	0.000	4.601112	4.866217

Example 15.5: Return to Education for Working Women

use http://fmwww.bc.edu/ec-p/data/wooldridge/MROZ

reg educ exper expersq motheduc fatheduc

Source	SS	df		MS		Number of obs		753
Model Residual	+ 1025.94324 2884.0966	 4 748		.48581 574412		F(4, 748) Prob > F R-squared Adj R-squared	= = =	66.52 0.0000 0.2624 0.2584
Total	3910.03984	752	5.19	952106		Root MSE	=	1.9636
educ	Coef.	Std.	Err.	t	P> t	[95% Conf.	In	terval]
educ exper	Coef. + .085378	Std. .0255		t 3.34	P> t 0.001	[95% Conf. .0352228		terval] 1355333
	' +		 485				 •	
exper	.085378	.0255	 5485 3276	3.34	0.001	.0352228	 • 	1355333
exper expersq	.085378 .085364	.0255 .0008	485 276 869	3.34	0.001 0.025	.0352228 0034812	 	1355333 0002317

test motheduc fatheduc

- (1) motheduc = 0.0
- (2) fatheduc = 0.0

$$F(2, 748) = 124.76$$

 $Prob > F = 0.0000$

ivreg lwage (educ = motheduc fatheduc) exper expersq

Instrumental variables (2SLS) regression

Source	SS	df	MS	Number of obs =	428
+				F(3, 424) =	8.14
Model	30.3074295	3	10.1024765	Prob > F = 0	.0000
Residual	193.020022	424	.4552359	R-squared = 0.	.1357
+				Adj R-squared = 0.	.1296

Total	223.327451	427 .5230	015108		Root MSE	= .67471
lwage	Coef.	Std. Err.	t	P> t	 [95% Conf.	Interval]
educ exper expersq _cons	.0613966 .0441704 000899 .0481003	.0314367 .0134325 .0004017 .4003281	1.95 3.29 -2.24 0.12	0.051 0.001 0.026 0.904	0003945 .0177679 0016885 7387744	.1231878 .0705729 0001094 .834975
Instrumented: Instruments:	educ exper expers	g motheduc	 fatheduc			

Example 15.6: Using Two Test Scores as Indicators of Ability

use http://fmwww.bc.edu/ec-p/data/wooldridge/WAGE2

ivreg lwage educ exper tenure married south urban black (IQ =KWW)

Instrumental variables (2SLS) regression

Source	SS	df 	MS 		Number of obs F(8, 926)	= 935 = 36.96
Model Residual + Total	31.4665073 134.189787 165.656294	926 .144 	331341 913377 362199		Prob > F R-squared Adj R-squared Root MSE	= 0.0000 = 0.1900
lwage	Coef.	Std. Err.	 t	P> t	[95% Conf.	Interval]
IQ	.0130473	.0049341	2.64	0.008	.0033641	.0227305
educ	.0250321	.0166068	1.51	0.132	0075591	.0576234
exper	.01442	.0033208	4.34	0.000	.0079029	.0209371
tenure	.0104562	.0026012	4.02	0.000	.0053512	.0155612
married	.2006904	.0406775	4.93	0.000	.1208596	.2805212
south	0515532	.0311279	-1.66	0.098	1126426	.0095362
urban	.1767058	.0282117	6.26	0.000	.1213394	.2320722
black	0225611	.0739597	-0.31	0.760	1677092	.122587
_cons	4.592453	.3257807	14.10	0.000	3.953099	5.231807
Instrumented:	IQ					

Instruments: educ exper tenure married south urban black KWW

Example 15.7: Return to Education for Working Women

use http://fmwww.bc.edu/ec-p/data/wooldridge/MROZ

Model Residual	SS 471.620998 1758.57526	df 4 423	117. 4.157	MS .90525 738833	_	-	= 28.36 = 0.0000 = 0.2115	
'	2230.19626					Adj R-squared Root MSE		
•	Coef.			t		[95% Conf.	Interval]	
exper expersq motheduc fatheduc	.0452254 0010091 .157597 .1895484	.04029 .00120 .03589 .03379	507 033 941 565	1.12 -0.84 4.39 5.62	0.262 0.402 0.000 0.000		.0013562 .2281501 .2558997	
predict double uhat1, res reg lwage educ exper expersq uhat1								

Source	SS	df	MS		Number of obs	
Model Residual	36.2573159 187.070135		9.06432898		F(4, 423) Prob > F R-squared Adj R-squared	= 0.0000 = 0.1624
Total	223.327451	427 .	.523015108		Root MSE	= .66502
lwage	Coef.	Std. Er	rr. t	P> t	[95% Conf.	<pre>Interval]</pre>
educ	.0613966	.030984		0.048	.000493	.1223003
exper	.0441704	.013239	3.34	0.001	.0181471	.0701937
expersq	000899	.000395	59 -2.27	0.024	0016772	0001208
uhat1	.0581666	.034807	1.67	0.095	0102501	.1265834
_cons	.0481003	.394575	0.12	0.903	7274721	.8236727

reg lwage educ exper expersq

Source	SS	df	MS	Number of obs =	428
	+			F(3, 424) =	26.29
Model	35.0223023	3	11.6741008	Prob > F =	0.0000
Residual	188.305149	424	.444115917	R-squared =	0.1568

Total	223.327451	427 .5230	 015108		Adj R-squared Root MSE	= 0.1509 = .66642
lwage	Coef.	Std. Err.	t	P> t	[95% Conf.	Interval]
educ exper expersq _cons	.1074896 .0415665 0008112 5220407	.0141465 .0131752 .0003932 .1986321	7.60 3.15 -2.06 -2.63	0.000 0.002 0.040 0.009	.0796837 .0156697 0015841 9124668	.1352956 .0674633 0000382 1316145

ivreg lwage (educ = motheduc fatheduc) exper expersq

Instrumental variables (2SLS) regression

Source	SS	df		MS		Number of obs	=	428
Model Residual	30.3074295 193.020022	3 424		.024765 1552359		F(3, 424) Prob > F R-squared Adj R-squared	= = =	8.14 0.0000 0.1357 0.1296
Total	223.327451	427	.523	3015108		Root MSE	=	.67471
lwage	Coef.	Std.	Err.	t	P> t	[95% Conf.	In	terval]
educ	.0613966	.0314	367	1.95	0.051	0003945		1231878
exper	.0441704	.0134	325	3.29	0.001	.0177679		0705729
expersq	000899	.0004	017	-2.24	0.026	0016885		0001094
_cons	.0481003	.4003	281	0.12	0.904	7387744		.834975

Instrumented: educ

Instruments: exper expersq motheduc fatheduc

Example 15.8: Return to Education for Working Women

use http://fmwww.bc.edu/ec-p/data/wooldridge/MROZ

ivreg lwage (educ = motheduc fatheduc) exper expersq

Instrumental variables (2SLS) regression

Source	SS	df	MS	Number of obs =	428
+				F(3, 424) =	8.14
Model	30.3074295	3	10.1024765	Prob > F = 0	.0000
Residual	193.020022	424	.4552359	R-squared = 0	.1357
+	- 			Adj R-squared = 0	.1296

Total 223.327451 427 .523015108 Root MSE = .67471

lwage | Coef. Std. Err. t P>|t| [95% Conf. Interval] educ | .0613966 .0314367 1.95 0.051 -.0003945 .1231878 .0705729 3.29 0.001 .0177679 exper | .0441704 .0134325 expersq | -.000899 .0004017 -2.24 0.026 -.0016885 -.0001094 _cons | .0481003 .4003281 0.12 0.904 -.7387744 .834975

Instrumented: educ

Instruments: exper expersq motheduc fatheduc

ssc install overid, replace

overid

Test of overidentifying restrictions: .378071 Chi-sq(1) P-value = .5386

ivreg lwage (educ = motheduc fatheduc huseduc) exper expersq

Instrumental variables (2SLS) regression

Source	SS	df	MS	Number of obs =	428
+				F(3, 424) =	11.52
Model	33.3927427	3	11.1309142	Prob > F =	0.0000
Residual	189.934709	424	.447959218	R-squared =	0.1495
+	+			Adj R-squared =	0.1435
Total	223.327451	427	.523015108	Root MSE =	.6693

lwage	Coef.	Std. Err.	t	P> t	[95% Conf.	Interval]
educ	.0803918	.021774	3.69	0.000	.0375934	.1231901
exper	.0430973	.0132649	3.25	0.001	.0170242	.0691704
expersq	0008628	.0003962	-2.18	0.030	0016415	0000841
_cons	1868574	.2853959	-0.65	0.513	7478243	.3741096

Instrumented: educ

Instruments: exper expersq motheduc fatheduc huseduc

overid

Test of overidentifying restrictions: 1.115043 Chi-sq(2) P-value = .5726

Example 15.9: Return of Education to Fertility

use http://fmwww.bc.edu/ec-p/data/wooldridge/FERTIL1

ivreg kids (educ = meduc feduc) age agesq black east northcen west farm othrural town smcity y74 y76 y78 y80 y82 y84

Instrumental variables (2SLS) regression

Source	SS	df	MS		Number of obs F(17, 1111)	
Model Residual	395.36632 2690.14298		568424 137082		Prob > F R-squared Adj R-squared	= 0.0000 = 0.1281
Total	3085.5093	1128 2.73	538059		Root MSE	= 1.5561
kids	Coef.	Std. Err.	t 	P> t	[95% Conf.	Interval]
educ	1527395	.0392232	-3.89	0.000	2296993	0757796
age	.5235536	.1390348	3.77	0.000	.2507532	.796354
agesq	005716	.0015705	-3.64	0.000	0087976	0026345
black	1.072952	.1737155	6.18	0.000	.732105	1.4138
east	.2285554	.1338537	1.71	0.088	0340792	.49119
northcen	.3744188	.122061	3.07	0.002	.1349228	.6139148
west	.2076398	.1676568	1.24	0.216	1213199	.5365995
farm	0770015	.1513718	-0.51	0.611	3740083	.2200052
othrural	1952451	.181551	-1.08	0.282	5514666	.1609764
town	.08181	.1246821	0.66	0.512	162829	.3264489
smcity	.2124996	.160425	1.32	0.186	1022706	.5272698
y74	.2721292	.172944	1.57	0.116	0672045	.6114629
y76	0945483	.1792324	-0.53	0.598	4462205	.2571239
y78	0572543	.1825536	-0.31	0.754	415443	.3009343
у80	053248	.1847175	-0.29	0.773	4156825	.3091865
y82	4962149	.1765888	-2.81	0.005	8427	1497298
y84	5213604	.1779205	-2.93	0.003	8704586	1722623
_cons	-7.241244 	3.136642	-2.31	0.021	-13.39565 	-1.086834

Instrumented: educ

Instruments: age agesq black east northcen west farm othrural town smcity y74

y76 y78 y80 y82 y84 meduc feduc

reg kids educ age agesq black east northcen west farm othrural town smcity y74 y76 y78 y80 y82 y84

Source	SS	df	MS	Number of obs =	1129
 +				F(17, 1111) =	9.72
Model	399.610888	17	23.5065228	Prob > F =	0.0000
Residual	2685.89841	1111	2.41755033	R-squared =	0.1295
 +				Adj R-squared =	0.1162
Total	3085.5093	1128	2.73538059	Root MSE =	1.5548

kids	Coef.	Std. Err.	t	P> t	[95% Conf.	Interval]
educ	1284268	.0183486	-7.00	0.000	1644286	092425
age	.5321346	.1383863	3.85	0.000	.2606065	.8036626
agesq	005804	.0015643	-3.71	0.000	0088733	0027347
black	1.075658	.1735356	6.20	0.000	.7351631	1.416152
east	.217324	.1327878	1.64	0.102	0432192	.4778672
northcen	.363114	.1208969	3.00	0.003	.125902	.6003261
west	.1976032	.1669134	1.18	0.237	1298978	.5251041
farm	0525575	.14719	-0.36	0.721	3413592	.2362443
othrural	1628537	.175442	-0.93	0.353	5070887	.1813814
town	.0843532	.124531	0.68	0.498	1599893	.3286957
smcity	.2118791	.160296	1.32	0.187	1026379	.5263961
y74	.2681825	.172716	1.55	0.121	0707039	.6070689
y76	0973795	.1790456	-0.54	0.587	448685	.2539261
y78	0686665	.1816837	-0.38	0.706	4251483	.2878154
у80	0713053	.1827707	-0.39	0.697	42992	.2873093
y82	5224842	.1724361	-3.03	0.003	8608214	184147
y84	5451661	.1745162	-3.12	0.002	8875846	2027477
_cons	-7.742457	3.051767	-2.54	0.011	-13.73033	-1.754579

reg educ meduc feduc age agesq black east northcen west farm othrural town smcity y74 y76 y78 y80 y82 y84

Source	SS	df	MS		Number of obs	
Model Residual	2256.26171 5606.85432		347873 122011		F(18, 1110) Prob > F R-squared Adj R-squared	= 0.0000 = 0.2869
Total	7863.11603	1128 6.97	084755		Root MSE	= 2.2475
educ	Coef.	Std. Err.	t t	P> t	[95% Conf.	Interval]
meduc	.1723015	.0221964	7.76	0.000	.1287499	.2158531
feduc	.2074188	.0254604	8.15	0.000	.1574629	.2573747
age	2243687	.2000013	-1.12	0.262	616792	.1680546
agesq	.0025664	.0022605	1.14	0.256	001869	.0070018
black	.3667819	.2522869	1.45	0.146	1282311	.861795
east	.2488042	.1920135	1.30	0.195	1279462	.6255546
northcen	.0913945	.1757744	0.52	0.603	2534931	.4362821
west	.1010676	.2422408	0.42	0.677	3742339	.5763691
farm	3792615	.2143864	-1.77	0.077	7999099	.0413869
othrural	560814	.2551196	-2.20	0.028	-1.061385	060243
town	.0616337	.1807832	0.34	0.733	2930816	.416349
smcity	.0806634	.2317387	0.35	0.728	3740319	.5353587
y74	.0060993	.249827	0.02	0.981	4840872	.4962858
y76	.1239104	.2587922	0.48	0.632	3838667	.6316874

y78	.2077861	.2627738	0.79	0.429	3078033	.7233755
y80	.3828911	.2642433	1.45	0.148	1355816	.9013638
y82	.5820401	.2492372	2.34	0.020	.0930108	1.071069
y84	.4250429	.2529006	1.68	0.093	0711741	.92126
_cons	13.63334	4.396773	3.10	0.002	5.006421	22.26027

predict v, res

reg kids educ age agesq black east northcen west farm othrural town smcity y74 y76 y78 y80 y82 y84 v

Source	ss	df	MS		Number of obs F(18, 1110)	
Model Residual Total	400.802376 2684.70692 +		.2667987 41865489 		Prob > F R-squared Adj R-squared Root MSE	= 0.0000 = 0.1299
kids	Coef.	Std. Err	. t	P> t	[95% Conf.	Interval]
educ	1527395	.0392012	-3.90	0.000	2296562	0758227
age	.5235536	.1389568	3.77	0.000	.250906	.7962012
agesq	005716	.0015697	-3.64	0.000	0087959	0026362
black	1.072952	.173618	6.18	0.000	.7322958	1.413609
east	.2285554	.1337787	1.71	0.088	0339321	.491043
northcen	.3744188	.1219925	3.07	0.002	.1350569	.6137807
west	.2076398	.1675628	1.24	0.216	1211357	.5364153
farm	0770015	.1512869	-0.51	0.611	373842	.2198389
othrural	1952451	.1814491	-1.08	0.282	5512671	.1607769
town	.08181	.1246122	0.66	0.512	162692	.3263119
smcity	.2124996	.160335	1.33	0.185	1020943	.5270935
y74	.2721292	.172847	1.57	0.116	0670144	.6112729
y76	0945483	.1791319	-0.53	0.598	4460236	.2569269
y78	0572543	.1824512	-0.31	0.754	4152424	.3007337
y80	053248	.1846139	-0.29	0.773	4154795	.3089836
y82	4962149	.1764897	-2.81	0.005	842506	1499238
y84	5213604	.1778207	-2.93	0.003	8702631	1724578
V	.0311374	.0443634	0.70	0.483	0559081	.1181829
_cons	-7.241244 	3.134883	-2.31	0.021	-13.39221	-1.09028

Example 15.10: Job Training and Worker Productivity

use http://fmwww.bc.edu/ec-p/data/wooldridge/JTRAIN

tsset fcode year

sort fcode year

drop if year==1989

ivreg D.lscrap (D.hrsemp = D.grant)

Instrumental variables (2SLS) regression

Source	SS	df		MS		Number of obs = 45 F(1, 43) = 3.20
Model Residual	.274952567 17.0148863	1 43		952567 569503		F(1, 43) = 3.20 Prob > F = 0.0808 R-squared = 0.0159 Adj R-squared = -0.0070
Total	17.2898389	44	.392	950883		Root MSE = .62904
D.lscrap	Coef.	Std.	Err.	t	P> t	[95% Conf. Interval]
hrsemp D1 _cons	0141532 0326684	.0079		-1.79 -0.26	0.081	0301148 .0018084 2886898 .223353
<pre>Instrumented: Instruments:</pre>	D.hrsemp D.grant					

reg D.lscrap D.hrsemp

Source	SS	df	MS		Number of obs	
Model Residual + Total	1.07071319 16.2191257 17.2898389	43 .377	071319 188969 		F(1, 43) Prob > F R-squared Adj R-squared Root MSE	= 0.0993 = 0.0619
D.lscrap	Coef.	Std. Err.	t	P> t	[95% Conf.	Interval]
hrsemp D1 _cons	0076007 1035161	.0045112 .103736	-1.68 -1.00	0.099	0166984 3127197	.0014971

This page prepared by Oleksandr Talavera (revised 8 Dec 2002)

Send your questions/comments/suggestions to Kit Baum at **baum@bc.edu**These pages are maintained by the Faculty Micro Resource Center's **GSA Program**,

a unit of Boston College **Academic Technology Services**