

MODULE TWO, PART THREE: ENDOGENEITY, INSTRUMENTAL VARIABLES AND TWO-STAGE LEAST SQUARES IN ECONOMIC EDUCATION RESEARCH USING STATA

Part Three of Module Two demonstrates how to address problems of endogeneity using STATA's two-stage least squares instrumental variable estimator, as well as how to perform and interpret the Durbin, Hausman and Wu specification test for endogeneity. Users of this model need to have completed Module One, Parts One and Three, and Module Two, Part One. That is, from Module One, users are assumed to know how to get data into STATA, recode and create variables within STATA, and run and interpret regression results. From Module Two, Part One, they are expected to have an understanding of the problem of and source of endogeneity and the basic idea behind an instrumental variable approach and the two-stage least squares method. The Becker and Johnston (1999) data set is used throughout this module.

THE CASE

As described in Module Two, Part One, Becker and Johnston (1999) called attention to classroom effects that might influence multiple-choice and essay type test taking skills in economics in different ways. For example, if the student is in a classroom that emphasizes skills associated with multiple-choice testing (*e.g.*, risk-taking behavior, question analyzing skills, memorization, and keen sense of judging between close alternatives), then the student can be expected to do better on multiple-choice questions. By the same token, if placed in a classroom that emphasizes the skills of essay test question answering (*e.g.*, organization, good sentence and paragraph construction, obfuscation when uncertain, logical argument, and good penmanship), then the student can be expected to do better on the essay component. Thus, Becker and Johnston attempted to control for the type of class of which the student is a member. Their measure of “teaching to the multiple-choice questions” is the mean score or mark on the multiple-choice questions for the school in which the i^{th} student took the 12th grade economics course. Similarly, the mean school mark or score on the essay questions is their measure of the i^{th} student’s exposure to essay question writing skills.

In equation form, the two equations that summarize the influence of the various covariates on multiple-choice and essay test questions are written as the following structural equations:

$$M_i = \rho_{21} + \rho_{22}W_i + \rho_{23}\bar{M}_i + \sum_{j=4}^J \rho_{2j}X_{ij} + U_i^*$$

$$W_i = \rho_{31} + \rho_{32}M_i + \rho_{33}\bar{W}_i + \sum_{j=4}^J \rho_{3j}X_{ij} + V_i^*$$

M_i and W_i are the i^{th} student's respective scores on the multiple-choice test and essay test. \bar{M}_i and \bar{W}_i are the mean multiple-choice and essay test scores at the school where the i^{th} student took the 12th grade economics course. The X_{ij} variables are the other exogenous variables (such as gender, age, English a second language, etc.) used to explain the i^{th} student's multiple-choice and essay marks, where the ρ s are parameters to be estimated. The inclusion of the mean multiple-choice and mean essay test scores in their respective structural equations, and their exclusion from the other equation, enables both of the structural equations to be identified within the system.

As shown in Module Two, Part One, the least squares estimators of the ρ s are biased because the error term U_i^* is related to W_i , in the first equation, and V_i^* is related to M_i in second equation. Instruments for regressors W_i and M_i are needed. Because the reduced form equations express W_i and M_i solely in terms of exogenous variables, they can be used to generate the respective instruments:

$$M_i = \Gamma_{21} + \Gamma_{22} \bar{W}_i + \Gamma_{23} \bar{M}_i + \sum_{j=4}^J \Gamma_{2j} X_{ij} + U_i^{**}.$$

$$W_i = \Gamma_{31} + \Gamma_{32} \bar{M}_i + \Gamma_{33} \bar{W}_i + \sum_{j=4}^J \Gamma_{3j} X_{ij} + V_i^{**}.$$

The reduced form parameters (Γ s) are functions of the ρ s, and the reduced form error terms U^{**} and V^{**} are functions of U^* and V^* , which are not related to any of the regressors in the reduced form equations.

We could estimate the reduced form equations and get \hat{M}_i and \hat{W}_i . We could then substitute \hat{M}_i and \hat{W}_i into the structural equations as proxy regressors (instruments) for M_i and W_i . The least squares regression of M_i on \hat{W}_i , \bar{M}_i and the X s and a least squares regression of W_i on \hat{M}_i , \bar{W}_i and the X s would yield consistent estimates of the respective ρ s, but the standard errors would be incorrect. STATA automatically performs the required estimations with the instrumental variables command:¹

```
ivreg dependent_variable independent_variables
      (exogenous_var_name=instruments)
```

Here, *independent_variables* should be all of your included, exogenous variables, and in the parentheses, we must specify the endogenous variable as a function of its instruments.

TWO-STAGE, LEAST SQUARES IN STATA

The Becker and Johnston (1999) data are in the file named “Bill.CSV.” Since this is a large dataset, users may need to increase the size of STATA following the procedures described in Module One, Part Three. For the version used in this Module (Intercooled STATA), the default memory is sufficient. Now, the file Bill.CSV can be read into STATA with the following insheet command. Note that, in this case, the directory has been changed beforehand so that we need only specify the file BILL.csv. For instance, say the file is located in the folder, “C:\Documents and Settings\My Documents\BILL.csv”. Then users can change the directory with the command, *cd* “C:\Documents and Settings\My Documents”, in which case the file may be accessed simply by specifying the actual file name, BILL.csv, as in the following:

```
insheet student school size other birthday sex eslflag ///
adultst mc1 mc2 mc3 mc4 mc5 mc6 mc7 mc8 mc9 mc10 mc11 ///
mc12 mc13 mc14 mc15 mc16 mc17 mc18 mc19 mc20 totalmc ///
avgmc essay1 essay2 essay3 essay4 totessay avgessay ///
totscore avgscore ma081 ma082 ec011 ec012 ma083 en093 ///
using "BILL.csv", comma
```

Using these recode and generate commands yields the following relevant variable definitions:

```
recode size (0/9=1) (10/19=2) (20/29=3) (30/39=4) ///
(40/49=5) (50/100=6)
gen smallest=(size==1)
gen smaller=(size==2)
gen small=(size==3)
gen large=(size==4)
gen larger=(size==5)
gen largest=(size==6)
```

TOTALMC: Student’s score on 12th grade economics multiple-choice exam (M_i).

AVGMC: Mean multiple-choice score for students at school (\bar{M}_i).

TOTESAY: Student’s score on 12th grade economics essay exam (W_i).

AVGESSAY: Mean essay score for students at school (\bar{W}_i).

ADULTST = 1, if a returning adult student, and 0 otherwise.

SEX = GENDER = 1 if student is female and 0 is male.

ESLFLAG = 1 if English is not student’s first language and 0 if it is.

EC011 = 1 if student enrolled in first semester 11 grade economics course, 0 if not.

EN093 = 1 if student was enrolled in ESL English course, 0 if not

MA081 = 1 if student enrolled in the first semester 11 grade math course, 0 if not.

MA082 = 1 if student was enrolled in the second semester 11 grade math course, 0 if not.

MA083 = 1 if student was enrolled in the first semester 12 grade math course, 0 if not.

SMALLER = 1 if student from a school with 10 to 19 test takers, 0 if not.

SMALL = 1 if student from a school with 20 to 29 test takers, 0 if not.

LARGE = 1 if student from a school with 30 to 39 test takers, 0 if not.

LARGER = 1 if student from a school with 40 to 49 test takers, 0 if not.

In all of the regressions, the effect of being at a school with more than 49 test takers is captured in the constant term, against which the other dummy variables are compared. The smallest schools should not be included so that we can treat the mean scores as exogenous and unaffected by any individual student's test performance, which is accomplished by adding the following command to the end of the summary statistics and regression commands:

```
if smallest!=1
```

This command is added to the end of our regression and summary statistics commands as an option, and it says to only perform the desired command if smallest is not equal to 1. We could also completely remove these observations with the command:

```
drop if smallest==1
```

The problem with this approach, however, is that we cannot retrieve observations once they've been dropped (at least not easily), so it's generally sound practice to follow the first approach.

The descriptive statistics on the relevant variables are then obtained with the following command, yielding the STATA output shown:ⁱⁱ

```
sum totalmc avgmc totessay avgessay adultst sex eslflag ///  
ec011 en093 ma081 ma082 ma083 smaller small large ///  
larger if smallest!=1
```

Variable	Obs	Mean	Std. Dev.	Min	Max
totalmc	3710	12.43558	3.961942	0	20
avgmc	3710	12.43558	1.972638	6.416667	17.07143
totessay	3710	18.13801	9.211914	0	40
avgessay	3710	18.13801	4.668071	5.7	29.78571
adultst	3710	.0051213	.0713894	0	1
sex	3710	.390566	.487943	0	1
eslflag	3710	.0641509	.2450547	0	1
ec011	3710	.6770889	.4676521	0	1
en093	3710	.0622642	.2416673	0	1
ma081	3710	.5913747	.491646	0	1
ma082	3710	.5487871	.4976812	0	1
ma083	3710	.4202156	.4936599	0	1
smaller	3711	.4621396	.4986317	0	1
small	3711	.2072218	.4053704	0	1
large	3711	.1064403	.3084419	0	1
larger	3711	.0978173	.2971075	0	1

For comparison with the two-stage least squares results, we start with the least squares regressions shown after this paragraph. The least squares estimations are typical of those found in multiple-choice and essay score correlation studies, with correlation coefficients of 0.77 and 0.78. The essay mark or score, *W*, is the most significant variable in the multiple-choice score regression (first of the two tables) and the multiple-choice mark, *M*, is the most significant variable in the essay regression (second of the two tables). Results like these have led researchers to conclude that the essay and multiple-choice marks are good predictors of each other. Notice also that both the mean multiple-choice and mean essay marks are significant in their respective equations, suggesting that something in the classroom environment or group experience influences individual test scores. Finally, being female has a significant negative effect on the multiple choice-test score, but a significant positive effect on the essay score, as expected from the least squares results reported by others. We will see how these results hold up in the two-stage least squares regressions.

```
regress totalmc totessay adultst sex avgmc eslflag ec011 ///
en093 ma081 ma082 ma083 smaller small large larger if ///
smallest!=1
```

Source	SS	df	MS	Number of obs =	3710
Model	34384.2039	14	2456.01457	F(14, 3695) =	380.73
Residual	23835.8996	3695	6.45085239	Prob > F =	0.0000
Total	58220.1035	3709	15.6969813	R-squared =	0.5906
				Adj R-squared =	0.5890
				Root MSE =	2.5399

totalmc	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
totessay	.2707917	.0054726	49.48	0.000	.2600621 .2815213
adultst	.4674948	.592213	0.79	0.430	-.6936016 1.628591
sex	-.5259549	.0912871	-5.76	0.000	-.7049329 -.3469768
avgmc	.3793819	.0252904	15.00	0.000	.3297974 .4289663
eslflag	.3933259	.8524557	0.46	0.645	-1.278004 2.064656
ec011	.0172264	.0926488	0.19	0.853	-.1644214 .1988743
en093	-.3117338	.8649386	-0.36	0.719	-2.007538 1.38407
ma081	-.120807	.1808402	-0.67	0.504	-.4753635 .2337494
ma082	.3827058	.1946737	1.97	0.049	.0010273 .7643843
ma083	.3703758	.1184767	3.13	0.002	.1380896 .602662
smaller	.0672105	.147435	0.46	0.649	-.2218514 .3562725
small	-.0056878	.1570632	-0.04	0.971	-.3136269 .3022513
large	.0663582	.1785263	0.37	0.710	-.2836616 .4163781
larger	.0565486	.1821756	0.31	0.756	-.300626 .4137232
_cons	2.654957	.3393615	7.82	0.000	1.989603 3.320311

```
regress totessay totalmc adultst sex avgessay eslflag ec011 ///
en093 ma081 ma082 ma083 smaller small large larger if ///
smallest!=1
```

Source	SS	df	MS	Number of obs =	3710
Model	191732.026	14	13695.1447	F(14, 3695) =	411.37
Residual	123011.315	3695	33.2912896	Prob > F =	0.0000
Total	314743.341	3709	84.8593533	R-squared =	0.6092
				Adj R-squared =	0.6077
				Root MSE =	5.7699

totessay	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
totalmc	1.408896	.0282236	49.92	0.000	1.353561 1.464231
adultst	-.8291495	1.345456	-0.62	0.538	-3.467058 1.808759
sex	1.239957	.2080107	5.96	0.000	.8321298 1.647784
avgessay	.4000235	.0237117	16.87	0.000	.3535343 .4465128
eslflag	.4511402	1.936935	0.23	0.816	-3.346427 4.248707
ec011	.2985372	.2104486	1.42	0.156	-.1140697 .7111441
en093	-2.020882	1.9647	-1.03	0.304	-5.872885 1.831122
ma081	.8495121	.4106127	2.07	0.039	.0444623 1.654562
ma082	.1590916	.4424986	0.36	0.719	-.7084739 1.026657
ma083	1.809542	.2679394	6.75	0.000	1.284218 2.334865
smaller	.6170663	.3305425	1.87	0.062	-.0309973 1.26513
small	.2693409	.3547691	0.76	0.448	-.4262217 .9649034
large	.2646448	.4052628	0.65	0.514	-.5299159 1.059206
larger	.0615029	.414367	0.15	0.882	-.7509076 .8739135
_cons	-8.948704	.5542766	-16.14	0.000	-10.03542 -7.861986

Theoretical considerations discussed in Module Two, Part One, suggest that these least squares estimates involve a simultaneous equation bias that is brought about by an apparent reverse causality between the two forms of testing. Consistent estimation of the parameters in this simultaneous equation system is possible with two-stage least squares, where our instrument \hat{M}_i for M_i is obtained by a least squares regression of M_i on SEX, ADULTST, AVGMC, AVGESSAY, ESLFLAG, SMALLER, SMALL, LARGE, LARGER, EC011, EN093, MA081, MA082, and MA083. Our instrument for \hat{W}_i for W_i is obtained by a least squares regression of W_i on SEX, ADULTST, AVGMC, AVGESSAY, ESLFLAG, SMALLER, SMALL, LARGE, LARGER, EC011, EN093, MA081, MA082, and MA083. STATA will do these regressions and the subsequent regressions for M and W employing these instruments via the following commands, which yield the subsequent output. Note that we should only specify as instruments variables that we are not including as independent variables in the full regression. As seen in the output tables, STATA correctly includes all of the exogenous variables as instruments in the two-stage least squares estimation:

```
ivreg totalmc adultst sex avgmc eslflag ec011 en093 ///
ma081 ma082 ma083 smaller small large larger ///
(totessay=avgessay) if smallest!=1
```

Instrumental variables (2SLS) regression

Source	SS	df	MS	Number of obs =	3710
Model	11874.9352	14	848.209657	F(14, 3695) =	106.01
Residual	46345.1683	3695	12.5426707	Prob > F =	0.0000
Total	58220.1035	3709	15.6969813	R-squared =	0.2040
				Adj R-squared =	0.2010
				Root MSE =	3.5416

totalmc	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
totessay	-.0524779	.036481	-1.44	0.150	-.1240029 .019047
adultst	.2533495	.8261181	0.31	0.759	-1.366343 1.873042
sex	-.0897195	.1360894	-0.66	0.510	-.3565373 .1770983
avgmc	.9748841	.0745801	13.07	0.000	.8286619 1.121106
eslflag	.6744471	1.189066	0.57	0.571	-1.656844 3.005738
ec011	.2925431	.1327137	2.20	0.028	.0323437 .5527425
en093	-1.588716	1.214273	-1.31	0.191	-3.969426 .7919948
ma081	.2995655	.2563946	1.17	0.243	-.2031234 .8022545

ma082	.8159711	.275631	2.96	0.003	.2755672	1.356375
ma083	1.635256	.2162776	7.56	0.000	1.211221	2.059291
smaller	.2715921	.2068164	1.31	0.189	-.1338934	.6770776
small	.04373	.2190764	0.20	0.842	-.3857925	.4732525
large	.1981185	.2493609	0.79	0.427	-.29078	.687017
larger	-.086771	.254517	-0.34	0.733	-.5857787	.4122366
_cons	-.3038295	.5749204	-0.53	0.597	-1.431022	.8233631

```
Instrumented: totessay
Instruments: adultst sex avgmc eslflag ec011 en093 ma081 ma082 ma083
              smaller small large larger avgessay
```

```
ivreg totessay adultst sex avgessay eslflag ec011 ///
en093 ma081 ma082 ma083 smaller small large larger ///
(totalmc=avgmc) if smallest!=1
```

Instrumental variables (2SLS) regression

Source	SS	df	MS	Number of obs =	3710
Model	112024.731	14	8001.76652	F(14, 3695) =	141.62
Residual	202718.61	3695	54.8629526	Prob > F =	0.0000
Total	314743.341	3709	84.8593533	R-squared =	0.3559
				Adj R-squared =	0.3535
				Root MSE =	7.407

totessay	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
totalmc	.0278877	.1583175	0.18	0.860	-.2825105 .338286
adultst	-.1690792	1.728774	-0.10	0.922	-3.558524 3.220366
sex	.6854633	.274106	2.50	0.012	.1480494 1.222877
avgessay	.8417622	.0579371	14.53	0.000	.7281703 .9553541
eslflag	1.723698	2.490557	0.69	0.489	-3.159304 6.6067
ec011	.7128703	.2740879	2.60	0.009	.1754919 1.250249
en093	-3.983249	2.531637	-1.57	0.116	-8.946793 .980295
ma081	1.069629	.5276886	2.03	0.043	.0350393 2.104218
ma082	1.217027	.5801887	2.10	0.036	.0795055 2.354548
ma083	3.892551	.4151461	9.38	0.000	3.078613 4.706489
smaller	.3348224	.4254953	0.79	0.431	-.4994062 1.169051
small	-.1364833	.4576746	-0.30	0.766	-1.033803 .7608364
large	.3418925	.5203201	0.66	0.511	-.6782504 1.362035
larger	-.0825121	.5321788	-0.16	0.877	-1.125905 .960881
_cons	-1.179741	1.12159	-1.05	0.293	-3.378737 1.019256

```
Instrumented: totalmc
Instruments: adultst sex avgessay eslflag ec011 en093 ma081 ma082 ma083
              smaller small large larger avgmc
```

The 2SLS results differ from the least squares results in many ways. The essay mark or score, W , is no longer a significant variable in the multiple-choice regression and the multiple-choice mark, M , is likewise insignificant in the essay regression. Each score appears to be measuring something different when the regressor and error-term-induced bias is eliminated by our instrumental variable estimators.

Both the mean multiple-choice and mean essay scores continue to be significant in their respective equations. But now being female is insignificant in explaining the multiple-choice test score. Being female continues to have a significant positive effect on the essay score.

DURBIN, HAUSMAN AND WU TEST FOR ENDOGENEITY

The theoretical argument is strong for treating multiple-choice and essay scores as endogenous when employed as regressors in the explanation of the other. Nevertheless, this endogeneity can be tested with the Durbin, Hausman and Wu specification test. There are at least two ways to perform this test in STATA: One is with the auxiliary regression as done with LIMDEP in Module Two, Part Two, and the other is with the general Hausman command. For consistency across Parts Two and Three, we use the auxiliary regression method here; however, for those interested in the more general Hausman command, type *help hausman* in the command window for a brief description.

Either a Wald statistic, in a Chi-square (χ^2) test with K^* degrees of freedom, or an F statistic with K^* and $n - (K + K^*)$ degrees of freedom, is used to test the joint significance of the contribution of the predicted values ($\hat{\mathbf{X}}^*$) of a regression of the K^* endogenous regressors, in matrix \mathbf{X}^* , on the exogenous variables (and column of ones for the constant term) in matrix \mathbf{Z} :

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \hat{\mathbf{X}}^* \boldsymbol{\gamma} + \boldsymbol{\varepsilon}^*,$$

where $\mathbf{X}^* = \mathbf{Z}\boldsymbol{\lambda} + \mathbf{u}$, $\hat{\mathbf{X}}^* = \mathbf{Z}\hat{\boldsymbol{\lambda}}$, and $\hat{\boldsymbol{\lambda}}$ is a least squares estimator of $\boldsymbol{\lambda}$.

H_o : $\boldsymbol{\gamma} = 0$, the variables in \mathbf{Z} are exogenous

H_A : $\boldsymbol{\gamma} \neq 0$, at least one of the variables in \mathbf{Z} is endogenous

In our case, $K^* = 1$ when the essay score is to be tested as an endogenous regressor in the multiple-choice equation and when the multiple-choice regressor is to be tested as endogenous in the essay equation. $\hat{\mathbf{X}}^*$ is an $n \times 1$ vector of predicted essay scores from a regression of essay scores on all the exogenous variables (for subsequent use in the multiple-choice equation) or an $n \times 1$ vector of predicted multiple-choice scores from a regression of multiple-choice scores on all the exogenous variables (for subsequent use in the essay equation). Because $K^* = 1$, the relevant test statistic is either the t , with $n - (K + K^*)$ degrees of freedom for small n or the standard normal, for large n .

In STATA, the predicted essay score is obtained by the following command, where the specification “predict totesshat, xb” tells STATA to predict the essay scores and keep them as a variable called “totesshat”:

```
regress totessay adultst sex avgmc avgessay eslflag ///
ec011 en093 ma081 ma082 ma083 smaller small large ///
larger if smallest!=1

predict totesshat, xb
```

The predicted essay scores are then added as a regressor in the original multiple-choice regression:

```
regress totalmc totessay adultst sex avgmc eslflag ///
```



```
ec011 en093 ma081 ma082 ma083 smaller small large ///
larger totesshat if smallest!=1
```

The test statistic for the totesshat coefficient is then used in the test of endogeneity. In the below STATA output, we see that the calculated standard normal test statistic z value is -12.92 , which far exceeds the absolute value of the 0.05 percent Type I error critical 1.96 standard normal value. Thus, the null hypothesis of an exogenous essay score as an explanatory variable for the multiple-choice score is rejected. As theorized, the essay score is endogenous in an explanation of the multiple-choice score.

```
. regress totessay adultst sex avgmc avgessay eslflag ec011 en093 ma081 ma082 ///
ma083 smaller small large larger if smallest!=1
```

Source	SS	df	MS	Number of obs =	3710
Model	108774.75	14	7769.62501	F(14, 3695) =	139.38
Residual	205968.591	3695	55.7425145	Prob > F =	0.0000
				R-squared =	0.3456
				Adj R-squared =	0.3431
Total	314743.341	3709	84.8593533	Root MSE =	7.4661

totessay	Coef.	Std. Err.	t	P> t	[95% Conf. Intervals]
adultst	-.1617771	1.741248	-0.09	0.926	-3.575679 3.252125
sex	.6819632	.2723475	2.50	0.012	.1479971 1.215929
avgmc	.0271476	.1553461	0.17	0.861	-.277425 .3317202
avgessay	.8405321	.0646428	13.00	0.000	.7137931 .9672711
eslflag	1.739961	2.506713	0.69	0.488	-3.174717 6.654638
ec011	.719975	.2721919	2.65	0.008	.1863138 1.253636
en093	-4.021669	2.541765	-1.58	0.114	-9.005069 .9617305
ma081	1.076408	.531461	2.03	0.043	.0344219 2.118393
ma082	1.237971	.571906	2.16	0.030	.1166884 2.359253
ma083	3.9324	.3425393	11.48	0.000	3.260815 4.603984
smaller	.3418961	.433852	0.79	0.431	-.5087167 1.192509
small	-.1350661	.4622235	-0.29	0.770	-1.041304 .7711722
large	.3469099	.5247716	0.66	0.509	-.6819605 1.37578
larger	-.0848079	.5361811	-0.16	0.874	-1.136048 .9664321
_cons	-1.186477	1.161393	-1.02	0.307	-3.463511 1.090556

```
. predict totesshat, xb
(1 missing value generated)
```

```
. regress totalmc totessay adultst sex avgmc eslflag ec011 en093 ma081 ma082 ma083 ///
> smaller small large larger totesshat if smallest!=1
```

Source	SS	df	MS	Number of obs =	3710
Model	35414.1533	15	2360.94355	F(15, 3694) =	382.41
Residual	22805.9502	3694	6.17378186	Prob > F =	0.0000
				R-squared =	0.6083
				Adj R-squared =	0.6067
Total	58220.1035	3709	15.6969813	Root MSE =	2.4847

totalmc	Coef.	Std. Err.	t	P> t	[95% Conf. Intervals]
totessay	.2855834	.0054749	52.16	0.000	.2748493 .2963175
adultst	.2533495	.5795925	0.44	0.662	-.8830033 1.389702
sex	-.0897195	.0954784	-0.94	0.347	-.2769151 .097476
avgmc	.974884	.0523243	18.63	0.000	.8722967 1.077471
eslflag	.6744471	.8342319	0.81	0.419	-.9611532 2.310047
ec011	.2925431	.09311	3.14	0.002	.1099909 .4750952
en093	-1.588716	.8519162	-1.86	0.062	-3.258988 .0815565
ma081	.2995655	.1798828	1.67	0.096	-.0531138 .6522448

ma082	.8159711	.1933787	4.22	0.000	.4368315	1.195111
ma083	1.635256	.1517372	10.78	0.000	1.337759	1.932753
smaller	.2715921	.1450994	1.87	0.061	-.0128907	.5560749
small	.04373	.1537008	0.28	0.776	-.2576168	.3450769
large	.1981185	.174948	1.13	0.258	-.1448856	.5411227
larger	-.086771	.1785655	-0.49	0.627	-.4368675	.2633256
totesshat	-.3380613	.0261736	-12.92	0.000	-.3893774	-.2867452
_cons	-.3038295	.4033559	-0.75	0.451	-1.094652	.4869926

The similar estimation routine to test for the endogeneity of the multiple-choice test score in the essay equation yields a calculated z test statistic of -11.71 , which far exceeds the absolute value of its 1.96 critical value. Thus, the null hypothesis of an exogenous multiple-choice score as an explanatory variable for the essay score is rejected. As theorized, the multiple-choice score is endogenous in an explanation of the essay score.

```
. regress totalmc avgessay adultst sex avgmc eslflag ec011 en093 ma081 ma082 ma083 ///
> smaller small large larger if smallest!=1
```

Source	SS	df	MS	Number of obs =	3710
Model	18615.7882	14	1329.69916	F(14, 3695) =	124.06
Residual	39604.3153	3695	10.7183533	Prob > F =	0.0000
Total	58220.1035	3709	15.6969813	R-squared =	0.3197
				Adj R-squared =	0.3172
				Root MSE =	3.2739

totalmc	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
avgessay	-.0441094	.0283459	-1.56	0.120	-.0996846	.0114658
adultst	.2618392	.7635394	0.34	0.732	-1.235161	1.758839
sex	-.1255075	.1194247	-1.05	0.293	-.3596523	.1086372
avgmc	.9734594	.0681195	14.29	0.000	.839904	1.107015
eslflag	.5831376	1.099197	0.53	0.596	-1.571954	2.738229
ec011	.2547603	.1193565	2.13	0.033	.0207492	.4887713
en093	-1.377667	1.114567	-1.24	0.217	-3.562894	.8075597
ma081	.2430779	.2330463	1.04	0.297	-.2138341	.6999899
ma082	.7510049	.2507815	2.99	0.003	.2593213	1.242689
ma083	1.428892	.1502039	9.51	0.000	1.134401	1.723382
smaller	.2536501	.1902446	1.33	0.183	-.1193446	.6266448
small	.050818	.2026856	0.25	0.802	-.3465686	.4482046
large	.1799134	.2301129	0.78	0.434	-.2712475	.6310742
larger	-.0823205	.235116	-0.35	0.726	-.5432904	.3786495
_cons	-.2415657	.509272	-0.47	0.635	-1.240048	.7569162

```
. predict totmchat, xb
(1 missing value generated)
```

```
. regress totessay totalmc adultst sex avgessay eslflag ec011 en093 ma081 ma082 ///
ma083 smaller small large larger totmchat if smallest!=1
```

Source	SS	df	MS	Number of obs =	3710
Model	196137.341	15	13075.8227	F(15, 3694) =	407.25
Residual	118606	3694	32.1077423	Prob > F =	0.0000
Total	314743.341	3709	84.8593533	R-squared =	0.6232
				Adj R-squared =	0.6216
				Root MSE =	5.6664

totessay	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
totalmc	1.485222	.028473	52.16	0.000	1.429398	1.541047
adultst	-.1690792	1.322524	-0.13	0.898	-2.762028	2.42387
sex	.6854633	.2096929	3.27	0.001	.274338	1.096589
avgessay	.8417622	.0443223	18.99	0.000	.7548637	.9286608
eslflag	1.723699	1.905293	0.90	0.366	-2.011832	5.459229

ec011	.7128703	.2096791	3.40	0.001	.3017721	1.123969
en093	-3.983249	1.93672	-2.06	0.040	-7.780395	-.186104
ma081	1.069629	.4036853	2.65	0.008	.2781608	1.861097
ma082	1.217027	.4438483	2.74	0.006	.3468152	2.087239
ma083	3.892551	.3175896	12.26	0.000	3.269883	4.515219
smaller	.3348226	.3255068	1.03	0.304	-.303368	.9730132
small	-.1364831	.3501242	-0.39	0.697	-.8229389	.5499726
large	.3418926	.3980484	0.86	0.390	-.4385237	1.122309
larger	-.0825121	.4071204	-0.20	0.839	-.880715	.7156909
totmchat	-1.457335	.1244158	-11.71	0.000	-1.701265	-1.213404
_cons	-1.179741	.8580241	-1.37	0.169	-2.861988	.5025071

CONCLUDING COMMENTS

This cookbook-type introduction to the use of instrumental variables and two-stage least squares regression and testing for endogeneity has just scratched the surface of this controversial problem in statistical estimation and inference. It was intended to enable researchers to begin using instrumental variables in their work and to enable readers of that work to have an idea of what is being done. To learn more about these methods there is no substitute for a graduate level textbook treatment such as that found in William Greene's *Econometric Analysis*.

REFERENCES

Becker, William E. and Carol Johnston (1999). "The Relationship Between Multiple Choice and Essay Response Questions in Assessing Economics Understanding," *Economic Record* (Economic Society of Australia), Vol. 75 (December): 348-357.

Greene, William (2003). *Econometric Analysis*. 5th Edition, New Jersey: Prentice Hall.

ENDNOTES

ⁱ As stated in Module Two, Part One, the 2SLS coefficients are consistent but not necessarily unbiased. Consistency is an asymptotic property for which there are no adjustments for degrees of freedom. Nevertheless, the default in the standard STATA routine for 2SLS, "ivreg," adjusts standard errors for the degrees of freedom. As an alternative, if your institution permits downloads from STATA's user-written routines, then the "ivreg2" command rather than "ivreg" can be employed. The "ivreg2" command makes no adjustment for degrees of freedom.

To use ivreg2, type *findit ivreg2* into the STATA command window, where a list of information and links to download this routine appears. Click on one of the download links and STATA automatically downloads and installs the routine for use. Users can then access the documentation for this routine by typing *help ivreg2*.

If users do not have access to `ivreg2` or are not permitted to download user-written routines on the machine in use, the following code provides the unadjusted standard errors after running a model using `ivreg`:

```
matrix large_sample_se=e(b)
matrix large_sample_var=e(V)*e(df_r)/e(N)
    local ncol=colsof(large_sample_se)
    forvalues i=1/`ncol' {
        matrix large_sample_se[1,`i']=sqrt(large_sample_var[`i',`i'])
    }
matrix list large_sample_se
```

ⁱⁱ Notice that the size variables (smaller to larger) show 3711 observations but the others show the correct 3710. This was an artifact of the way the size variables were created in STATA. The extra blank space has no relevance and is ignored in the calculations that are all based on the original 3710 observations.