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

Part Four of Module Two provides a cookbook-type demonstration of the steps required to use SAS to address problems of endogeneity using a two-stage least squares, instrumental variable estimator. The Durbin, Hausman and Wu specification test for endogeneity is also demonstrated. Users of this model need to have completed Module One, Parts One and Four, and Module Two, Part One. That is, from Module One, users are assumed to know how to get data into SAS, recode and create variables within SAS, 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 for demonstration purposes only. Module Two, Parts Two and Three demonstrate in LIMDEP and STATA what is done here in SAS.

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 12^{th} 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 follow 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 ρs 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 estimation of the ρs involves bias 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 (Γs) are functions of the ρs, 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 Xs and a least squares regression of W_i on $\hat{M}_i$, $\bar{W}_i$ and the Xs would yield consistent estimates of the respective ρs, but the standard errors would be incorrect. SAS can automatically do all the required estimations with the two-stage, least squares command:

```
proc syslin data= dataset_name 2sls;
        endogenous  p;
        instruments y u s;
        Equation 1: model q = p y s;
        Equation 2: model q = p u;
 run;
```

This 'proc syslin' package command, however, inappropriate adjustment for degrees of freedom in calculating the standard errors. What follows is the correct two-step procedure for the asymptotic efficient estimators, which involve no adjustment for degrees of freedom.

TWO-STAGE, LEAST SQUARES IN SAS

The Becker and Johnston (1999) data are in the file named "Bill.CSV." The file Bill.CSV can be read into SAS with the following read command (the file may be located anywhere on your hard drive but here it is located on the e drive):

```
data work.bill; infile 'e:\ BILL.CSV' delimiter = ',' missover dsd;

informat student best32.;     informat school best32.;     informat size best32.;
informat other best32.;       informat birthday best32.;   informat sex best32.;
informat eslflag best32.;     informat adultst best32.;    informat mc1 best32.;
informat mc2 best32.;         informat mc3 best32.;        informat mc4 best32.;
informat mc5 best32.;         informat mc6 best32.;        informat mc7 best32.;
informat mc8 best32.;         informat mc9 best32.;        informat mc10 best32.;
informat mc11 best32.;        informat mc12 best32.;       informat mc13 best32.;
informat mc14 best32.;        informat mc15 best32.;       informat mc16 best32.;
informat mc17 best32.;        informat mc18 best32.;       informat mc19 best32.;
informat mc20 best32.;        informat totalmc best32.;    informat avgmc best32.;
informat essay1 best32.;      informat essay2 best32.;     informat essay3 best32.;
informat essay4 best32.;      informat totessay best32.;   informat avgessay best32.;
informat totscore best32.;    informat avgscore best32.;   informat ma081 best32.;
informat ma082 best32.;       informat ec011 best32.;      informat ec012 best32.;
informat ma083 best32.;       informat en093 best32.;

format student best12.;       format school best12.;       format size best12.;
format other best12.;         format birthday best12.;     format sex best12.;
format eslflag best12.;       format adultst best12.;      format mc1 best12.;
format mc2 best12.;           format mc3 best12.;          format mc4 best12.;
format mc5 best12.;           format mc6 best12.;          format mc7 best12.;
format mc8 best12.;           format mc9 best12.;          format mc10 best12.;
format mc11 best12.;          format mc12 best12.;         format mc13 best12.;
format mc14 best12.;          format mc15 best12.;         format mc16 best12.;
format mc17 best12.;          format mc18 best12.;         format mc19 best12.;
format mc20 best12.;          format totalmc best12.;      format avgmc best12.;
format essay1 best12.;        format essay2 best12.;       format essay3 best12.;
format essay4 best12.;        format totessay best12.;     format avgessay best12.;
format totscore best12.;      format avgscore best12.;     format ma081 best12.;
format ma082 best12.;         format ec011 best12.;        format ec012 best12.;
format ma083 best12.;         format en093 best12.;
```

input 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; run;

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

```
data Bill;
        set M2P4.Bill;
                smallest = 0;
                smaller = 0;
                small = 0;
                large = 0;
                larger = 0;
                largest = 0;
                if size > 0  & size < 10 then smallest = 1;
                if size > 9  & size < 20 then smaller = 1;
                if size > 19 & size < 30 then small = 1;
                if size > 29 & size < 40 then large = 1;
                if size > 39 & size < 50 then larger = 1;
                if size > 49 then largest = 1;
run;
```

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

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

TOTESSAY: 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 need to be rejected to treat the mean scores as exogenous and unaffected by any individual student's test performance, which is accomplished with the following command:

```
data Bill; set Bill;
        if smallest = 1 then delete;
        if student = . then delete;
run;
```

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

The MEANS Procedure

Variable	N	Mean	Std Dev	Minimum	Maximum
totalmc	3710	12.4355795	3.9619416	0	20.0000000
avgmc	3710	12.4355800	1.9726377	6.4166670	17.0714300
totessay	3710	18.1380054	9.2119137	0	40.0000000
avgessay	3710	18.1380059	4.6680707	5.7000000	29.7857100
adultst	3710	0.0051213	0.0713894	0	1.0000000
sex	3710	0.3905660	0.4879430	0	1.0000000
eslflag	3710	0.0641509	0.2450547	0	1.0000000
ec011	3710	0.6770889	0.4676521	0	1.0000000
en093	3710	0.0622642	0.2416673	0	1.0000000
ma081	3710	0.5913747	0.4916460	0	1.0000000
ma082	3710	0.5487871	0.4976812	0	1.0000000
ma083	3710	0.4202156	0.4936599	0	1.0000000
smaller	3710	0.4622642	0.4986412	0	1.0000000
small	3710	0.2072776	0.4054108	0	1.0000000
large	3710	0.1064690	0.3084785	0	1.0000000
larger	3710	0.0978437	0.2971432	0	1.0000000

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.

```
proc reg data = Bill; model totalmc totessay adultst sex avgmc eslflag
ec011 en093 ma081 ma082 ma083 smaller small large larger; quit;
```

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	14	34384	2456.01457	380.73	<.0001
Error	3695	23836	6.45085		
Corrected Total	3709	58220			

Root MSE	2.53985	R-Square	0.5906
Dependent Mean	12.43558	Adj R-Sq	0.5890
Coeff Var	20.42408		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	2.65496	0.33936	7.82	<.0001
totessay	1	0.27079	0.00547	49.48	<.0001
adultst	1	0.46749	0.59221	0.79	0.4299
sex	1	-0.52595	0.09129	-5.76	<.0001
avgmc	1	0.37938	0.02529	15.00	<.0001
eslflag	1	0.39333	0.85246	0.46	0.6445
ec011	1	0.01723	0.09265	0.19	0.8525
en093	1	-0.31173	0.86494	-0.36	0.7186
ma081	1	-0.12081	0.18084	-0.67	0.5042
ma082	1	0.38271	0.19467	1.97	0.0494
ma083	1	0.37038	0.11848	3.13	0.0018
smaller	1	0.06721	0.14743	0.46	0.6485
small	1	-0.00569	0.15706	-0.04	0.9711
large	1	0.06636	0.17853	0.37	0.7101
larger	1	0.05655	0.18218	0.31	0.7563

proc reg data = Bill; model totessay = totalmc adultst sex avgessay eslflag ec011 en093 ma081 ma082 ma083 smaller small large larger; quit;

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	14	191732	13695	411.37	<.0001
Error	3695	123011	33.29129		
Corrected Total	3709	314743			

Root MSE	5.76986	R-Square	0.6092
Dependent Mean	18.13801	Adj R-Sq	0.6077
Coeff Var	31.81089		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	-8.94870	0.55428	-16.14	<.0001
totalmc	1	1.40890	0.02822	49.92	<.0001
adultst	1	-0.82915	1.34546	-0.62	0.5378
sex	1	1.23996	0.20801	5.96	<.0001
avgessay	1	0.40002	0.02371	16.87	<.0001
eslflag	1	0.45114	1.93694	0.23	0.8158
ec011	1	0.29854	0.21045	1.42	0.1561
en093	1	-2.02088	1.96470	-1.03	0.3037
ma081	1	0.84951	0.41061	2.07	0.0386
ma082	1	0.15909	0.44250	0.36	0.7192
ma083	1	1.80954	0.26794	6.75	<.0001
smaller	1	0.61707	0.33054	1.87	0.0620
small	1	0.26934	0.35477	0.76	0.4478
large	1	0.26464	0.40526	0.65	0.5138
larger	1	0.06150	0.41437	0.15	0.8820

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 ($\hat{W}_i$) 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. SAS will do these regressions and the subsequent regressions for *M* and *W* employing these instruments via the following commands, which yield the subsequent outputs:

```
proc model data = bill;
    instruments sex adultst avgmc avgessay eslflag smaller
        small large larger ec011 en093 ma081 ma082 ma083;
    totalmc = constant_1 + totessay_1*totessay + adultst_1*adultst +
        sex_1*sex + avgmc_1*avgmc + eslflag_1*eslflag +
        eco11_1*ec011 + en093_1*en093 + ma081_1*ma081 +
        ma082_1*ma082 + ma083_1*ma083 + smaller_1*smaller +
        small_1*small + large_1*large + larger_1*larger;

    totessay = constant_2 + totalm_2*totalmc + adultst_2*adultst +
        sex_2*sex + avgessay_2*avgessay + eslflag_2*eslflag +
        ec011_2*ec011 + en093_2*en093 + ma081_2*ma081 +
        a082_2*ma082+ ma083_2*ma083 + smaller_2*smaller +
        small_2*small + large_2*large + larger_2*larger;

     fit totalmc totessay / 2sls outv=vdata vardef=N;
quit;
```

Nonlinear 2SLS Summary of Residual Errors

Equation	DF Model	DF Error	SSE	MSE	Root MSE	R-Square	Adj R-Sq
totalmc	15	3695	46345.2	12.4920	3.5344	0.2040	0.2010
totessay	15	3695	202719	54.6411	7.3920	0.3559	0.3535

- _

Number of Observations		Statistics for System	
Used	3710	Objective	4.76E-23
Missing	0	Objective*N	1.766E-19

Nonlinear 2SLS Parameter Estimates

Parameter	Estimate	Approx Std Err	t Value	Approx Pr > |t|
constant_1	-0.30383	0.5738	-0.53	0.5965
totessay_1	-0.05248	0.0364	-1.44	0.1496
adultst_1	0.253349	0.8244	0.31	0.7586
sex_1	-0.08972	0.1358	-0.66	0.5089
avgmc_1	0.974884	0.0744	13.10	<.0001
eslflag_1	0.674447	1.1867	0.57	0.5698
eco11_1	0.292543	0.1324	2.21	0.0273
en093_1	-1.58872	1.2118	-1.31	0.1899
ma081_1	0.299566	0.2559	1.17	0.2418
ma082_1	0.815971	0.2751	2.97	0.0030
ma083_1	1.635256	0.2158	7.58	<.0001
smaller_1	0.271592	0.2064	1.32	0.1883
small_1	0.04373	0.2186	0.20	0.8415
large_1	0.198118	0.2489	0.80	0.4260
larger_1	-0.08677	0.2540	-0.34	0.7327
constant_2	-1.17974	1.1193	-1.05	0.2920
totalm_2	0.027888	0.1580	0.18	0.8599
adultst_2	-0.16908	1.7253	-0.10	0.9219
sex_2	0.685463	0.2736	2.51	0.0123
avgessay_2	0.841762	0.0578	14.56	<.0001
eslflag_2	1.723699	2.4855	0.69	0.4880
ec011_2	0.71287	0.2735	2.61	0.0092
en093_2	-3.98325	2.5265	-1.58	0.1150
ma081_2	1.069629	0.5266	2.03	0.0423
ma082_2	1.217027	0.5790	2.10	0.0356
ma083_2	3.892551	0.4143	9.40	<.0001
smaller_2	0.334822	0.4246	0.79	0.4305
small_2	-0.13648	0.4567	-0.30	0.7651
large_2	0.341892	0.5193	0.66	0.5103
larger_2	-0.08251	0.5311	-0.16	0.8765

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, which is a two-step procedure in SAS.

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 **X*,** on the exogenous variables (and column of ones for the constant term) in matrix **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 **Z** are exogenous

H_A : $\boldsymbol{\gamma} \neq 0$, at least one of the variables in **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 SAS, the predicted essay score is obtained by the following command, where the specification "output out=essaypredict p=Essayhat;" tells SAS to predict the essay scores and keep them as a variable called "Essayhat":

```
proc reg data = bill; model totessay = adultst sex avgessay avgmc eslflag
        ec011 en093 ma081 ma082 ma083 smaller small large larger;
output out=essaypredict p=Essayhat; quit;
```

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

```
proc reg data = essaypredict; model totalmc = totessay adultst sex avgmc eslflag
        ec011 en093 ma081 ma082 ma083 smaller small large larger Essayhat;
quit;
```

The test statistic for the Essayhat coefficient is then used in the test of endogeneity. In the below SAS output, we see that the calculated standard normal test statistic z value is −12.916, 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.

proc reg data = bill; model totessay = adultst sex avgessay avgmc eslflag ec011 en093 ma081 ma082 ma083 smaller small large larger; output out=essaypredict p=Essayhat; quit;

-

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	14	108775	7769.62501	139.38	<.0001
Error	3695	205969	55.74251		
Corrected Total	3709	314743			

Root MSE	7.46609	R-Square	0.3456
Dependent Mean	18.13801	Adj R-Sq	0.3431
Coeff Var	41.16269		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	-1.18648	1.16139	-1.02	0.3070
adultst	1	-0.16178	1.74125	-0.09	0.9260
sex	1	0.68196	0.27235	2.50	0.0123
avgessay	1	0.84053	0.06464	13.00	<.0001
avgmc	1	0.02715	0.15535	0.17	0.8613
eslflag	1	1.73996	2.50671	0.69	0.4877
ec011	1	0.71997	0.27219	2.65	0.0082
en093	1	-4.02167	2.54176	-1.58	0.1137
ma081	1	1.07641	0.53146	2.03	0.0429
ma082	1	1.23797	0.57191	2.16	0.0305
ma083	1	3.93240	0.34254	11.48	<.0001
smaller	1	0.34190	0.43385	0.79	0.4307
small	1	-0.13507	0.46222	-0.29	0.7701
large	1	0.34691	0.52477	0.66	0.5086
larger	1	-0.08481	0.53618	-0.16	0.8743

proc reg data = essaypredict; model totalmc = totessay adultst sex avgmc eslflag ec011 en093 ma081 ma082 ma083 smaller small large larger Essayhat; quit;

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	Intercept	1	-0.30383	0.40336	-0.75	0.4513
totessay		1	0.28558	0.00547	52.16	<.0001
adultst		1	0.25335	0.57959	0.44	0.6621
sex		1	-0.08972	0.09548	-0.94	0.3474
avgmc		1	0.97488	0.05232	18.63	<.0001
eslflag		1	0.67445	0.83423	0.81	0.4189
ec011		1	0.29254	0.09311	3.14	0.0017
en093		1	-1.58872	0.85192	-1.86	0.0623
ma081		1	0.29957	0.17988	1.67	0.0959
ma082		1	0.81597	0.19338	4.22	<.0001
ma083		1	1.63526	0.15174	10.78	<.0001
smaller		1	0.27159	0.14510	1.87	0.0613
small		1	0.04373	0.15370	0.28	0.7760
large		1	0.19812	0.17495	1.13	0.2575
larger		1	-0.08677	0.17857	-0.49	0.6270
essayhat	Predicted Value of totessay	1	-0.33806	0.02617	-12.92	<.0001

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 −11.713, 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.

```
proc reg data = bill; model totalmc = adultst sex avgmc avgessay eslflag ec011 en093 ma081 ma082 ma083 smaller small large larger; output out=mcpredict p=MChat; quit;
```

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	14	18616	1329.69916	124.06	<.0001
Error	3695	39604	10.71835		
Corrected Total	3709	58220			

Root MSE	3.27389	R-Square	0.3197
Dependent Mean	12.43558	Adj R-Sq	0.3172
Coeff Var	26.32680		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	-0.24157	0.50927	-0.47	0.6353
adultst	1	0.26184	0.76354	0.34	0.7317
sex	1	-0.12551	0.11942	-1.05	0.2934
avgmc	1	0.97346	0.06812	14.29	<.0001
avgessay	1	-0.04411	0.02835	-1.56	0.1198
eslflag	1	0.58314	1.09920	0.53	0.5958
ec011	1	0.25476	0.11936	2.13	0.0329
en093	1	-1.37767	1.11457	-1.24	0.2165
ma081	1	0.24308	0.23305	1.04	0.2970
ma082	1	0.75100	0.25078	2.99	0.0028
ma083	1	1.42889	0.15020	9.51	<.0001
smaller	1	0.25365	0.19024	1.33	0.1825
small	1	0.05082	0.20269	0.25	0.8020
large	1	0.17991	0.23011	0.78	0.4344
larger	1	-0.08232	0.23512	-0.35	0.7263

```
proc reg data = mcpredict; model totessay = totalmc adultst sex avgessay eslflag ec011 en093 ma081 ma082 ma083 smaller small large larger MChat; quit;
```

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	Intercept	1	-1.17974	0.85802	-1.37	0.1692
totalmc		1	1.48522	0.02847	52.16	<.0001
adultst		1	-0.16908	1.32252	-0.13	0.8983
sex		1	0.68546	0.20969	3.27	0.0011
avgessay		1	0.84176	0.04432	18.99	<.0001
eslflag		1	1.72370	1.90529	0.90	0.3657
ec011		1	0.71287	0.20968	3.40	0.0007
en093		1	-3.98325	1.93672	-2.06	0.0398
ma081		1	1.06963	0.40369	2.65	0.0081
ma082		1	1.21703	0.44385	2.74	0.0061
ma083		1	3.89255	0.31759	12.26	<.0001
smaller		1	0.33482	0.32551	1.03	0.3037
small		1	-0.13648	0.35012	-0.39	0.6967
large		1	0.34189	0.39805	0.86	0.3904
larger		1	-0.08251	0.40712	-0.20	0.8394
MChat	Predicted Value of totalmc	1	-1.45733	0.12442	-11.71	<.0001

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.