AAEC 6311

LAB #4

Objectives:

1. Learn to estimate system of equations using SUR and SEM procedures

The system of equations we will estimate today is a system of demand equations developed by Deaton and Muellbauer (1980a; 1980b). The system is called the "almost ideal demand system" (AIDS). A simplified version of the model that we will use today is the linear approximate almost ideal demand system (LA/AIDS).

Data for estimation was obtained from the 2006 Ecuadorian consumer expenditure survey.

When estimating demand models with cross sectional data you also need to include in the model socio-demographic characteristics (e.g., number of children in household); however, for purposes of this lab we will only consider prices and total expenditures. We consider a demand system for three commodities: meats, dairy products and pulses (sources of protein).

The LA/AIDS model for the three food commodities can be written as:

where 𝑤𝑤𝑖𝑖is the budget share of good i; 𝑝𝑝𝑖𝑖 is the price of good j; 𝑥𝑥 ∗ are real total expenditures of the three goods in question; 𝜀𝜀𝑖𝑖 is the random disturbance.

To be consistent with economic demand theory, the parameters in the system need to satisfy the following restrictions:

* 1. The homogeneity restriction is satisfied for the AIDS model if, and only if, for all j:
  2. The symmetry restriction is satisfied by:

3

𝑗𝑗=1

∑

𝛾𝛾𝑖𝑖𝑗𝑗=0

* 1. The adding up restriction ∑3

𝑖𝑖=1

𝛾𝛾𝑖𝑖𝑖𝑖 = 𝛾𝛾𝑖𝑖𝑖𝑖

𝑤𝑤𝑖𝑖=1 is satisfied if for all j:

𝑖𝑖=1

𝑗𝑗=1

3

∑

𝑖𝑖=1

𝛼𝛼𝑖𝑖=1, ∑3

𝛽𝛽𝑖𝑖=0 and ∑3

𝛾𝛾𝑗𝑗𝑖𝑖=0.

Part 1. Basic Operations Using SAS

* 1. **Import and manipulate the data**
  2. **Calculate basic summary statistics and report results for all variables included in the model:**

|  |
| --- |
| The SAS System |

The MEANS Procedure

| Variable | Label | N | Mean | Std Dev | Minimum | Maximum |
| --- | --- | --- | --- | --- | --- | --- |
| |  | | --- | | wmeats | | wdairy | | wpulses | | pmeats1 | | pdairy1 | | ppulses1 | | lnpmeats | | lnpdairy | | lnppulses | | yprot | | lnpstar | | realy | | |  | | --- | | wmeats | | wdairy | | wpulses | | pmeats1 | | pdairy1 | | ppulses1 | | lnpmeats | | lnpdairy | | lnppulses | | yprot | | lnpstar | | realy | | |  | | --- | | 2066 | | 2066 | | 2066 | | 2066 | | 2066 | | 2066 | | 2066 | | 2066 | | 2066 | | 2066 | | 2066 | | 2066 | | |  | | --- | | 0.6570403 | | 0.2691918 | | 0.0737679 | | 2.6099958 | | 1.2314839 | | 1.2178835 | | 0.9182512 | | -0.2223833 | | 0.1140849 | | 17.0868642 | | 0.5921247 | | 2.0685185 | | |  | | --- | | 0.1677869 | | 0.1586839 | | 0.0704007 | | 0.8200990 | | 1.5095597 | | 0.4439593 | | 0.2848808 | | 0.9337580 | | 0.4679136 | | 10.6786414 | | 0.3438878 | | 0.6052187 | | |  | | --- | | 0.0010269 | | 0.000101950 | | 0.000030122 | | 0.3651000 | | 0.0107000 | | 0.0096400 | | -1.0075840 | | -4.5375115 | | -4.6418342 | | 0.1988054 | | -1.2368221 | | -2.2930764 | | |  | | --- | | 0.9839341 | | 0.9452825 | | 0.8836403 | | 11.3514000 | | 23.7358000 | | 4.3859600 | | 2.4293411 | | 3.1669845 | | 1.4784085 | | 128.6019441 | | 2.5527479 | | 4.1817308 | |

# **Section 1. Assuming all variables in the system are exogenous**

Part 2. Use proc reg to estimate the three equations of the system separately. Report the estimation results. Interpret all the coefficients in the first equation.

Proc Reg Output (OLS)

|  |  |  |  |
| --- | --- | --- | --- |
| Var | W1 | W2 | W3 |
| Int | 0.64955 (0.01787) | 0.19014 (0.01651) | 0.16032 (0.00733) |
| Ln p1 | 0.05704 (0.01269) | -0.02132 (0.01173) | -0.03572 (0.00521) |
| Ln p2 | -0.0327 (0.00391) | 0.04062 (0.00361) | -0.00796 (0.00160) |
| Ln p3 | 0.00244 (0.00778) | -0.02241 (0.00719) | 0.01997 (0.00319) |
| Ln x\* | -0.0254 (0.00599) | 0.05329 (0.00554) | -0.02794 (0.00246) |

If the prices of meats, dairy, and pulses were zero and if the total expenditures on meats, diary, and pulses, were zero, we would expect there to be a budget share of 0.64955 on meats. This, of course, would make no economic sense.

On average, for every percent increase in the price of meats, there will be a 0.05704 increase in the budget share of meats, ceteris paribus. On average, for every percent increase in the price of dairy, there will be a 0.0327 decrease in the budget share of meats, ceteris paribus. On average, for every percent increase in the price of pulses, there will be a 0.00244 increase in the budget share of meats, ceteris paribus. On average, for every percent increase in the real total expenditures of the goods on meats, dairy, and pulses, there will be a 0.0254 decrease in the budget share of meats, ceteris paribus.

Part3. Use proc model to estimate the system of equations

* 1. **Use proc model and ols to estimate the system. Compare the results to those obtained using proc reg for the individual equations in Part 2.**

**Proc Model Output (OLS)**

|  |  |  |  |
| --- | --- | --- | --- |
| **Var** | **W1** | **W2** | **W3** |
| **Int** | 0.649548 (0.0179) | 0.190137 (0.0165) | 0.160315 (0.00733) |
| **Ln p1** | 0.057045 (0.0127) | -0.02132 (0.0117) | -0.03572 (0.00521) |
| **Ln p2** | -0.03266 (0.00391) | 0.040623 (0.00361) | -0.00796 (0.00160) |
| **Ln p3** | 0.002442 (0.00778) | -0.02241 (0.00719) | 0.001997 (0.00319) |
| **Ln x\*** | -0.02535 (0.00599) | 0.053286 (0.00554) | -0.02794 (0.00246) |

As we can see, using proc reg and proc model results in the same values for the coefficients and standard errors of our variables, save for small rounding discrepancies.

* 1. **Use proc model and sur to estimate the system. Check the log window in SAS. Discuss the problem when estimating the model with three equations.**

We receive an error about the covariance across equations or, more specifically, that our system of equations is singular.

* 1. **Drop one equation of the model and estimate a two-equation system using proc model in SUR. Check the log window in SAS. Use HCCME=3 if possible.**

With two equations, our system of equations now have convergence and the criteria has been met to find the coefficients and standard errors of our variables.

|  |  |  |
| --- | --- | --- |
| Var | W1 | W2 |
| Int | 0.649548 (0.0250) | 0.190137 (0.0230) |
| Ln p1 | 0.057045 (0.0183) | -0.02132 (0.0185) |
| Ln p2 | -0.03266 (0.00374) | 0.040623 (0.00342) |
| Ln p3 | 0.002442 (0.00789) | -0.02241 (0.00765) |
| Ln x\* | -0.02535 (0.00783) | 0.053286 (0.00640) |

* 1. **Test the restrictions implied by demand theory using proc model and SUR. Report and interpret the test results. Use HCCME=3 if possible.**

The MODEL Procedure

| **Nonlinear SUR Summary of Residual Errors** | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Equation** | **DF Model** | **DF Error** | **SSE** | **MSE** | **Root MSE** | **R-Square** | **Adj R-Sq** | **Label** |
| **wmeats** | 5 | 2061 | 55.2617 | 0.0268 | 0.1637 | 0.0494 | 0.0476 | wmeats |
| **wdairy** | 5 | 2061 | 47.1754 | 0.0229 | 0.1513 | 0.0927 | 0.0910 | wdairy |

| **Nonlinear SUR Parameter Estimates** | | | | |
| --- | --- | --- | --- | --- |
| **Parameter** | **Estimate** | **Approx Std Err** | **t Value** | **Approx Pr > |t|** |
| **a1** | 0.649548 | 0.0179 | 36.35 | <.0001 |
| **a2** | 0.190137 | 0.0165 | 11.52 | <.0001 |
| **b1** | -0.02535 | 0.00599 | -4.23 | <.0001 |
| **b2** | 0.053286 | 0.00554 | 9.62 | <.0001 |
| **g11** | 0.057045 | 0.0127 | 4.49 | <.0001 |
| **g12** | -0.03266 | 0.00391 | -8.35 | <.0001 |
| **g13** | 0.002442 | 0.00778 | 0.31 | 0.7537 |
| **g21** | -0.02132 | 0.0117 | -1.82 | 0.0692 |
| **g22** | 0.040623 | 0.00361 | 11.24 | <.0001 |
| **g23** | -0.02241 | 0.00719 | -3.12 | 0.0019 |

| **Test Results** | | | | |
| --- | --- | --- | --- | --- |
| **Test** | **Type** | **Statistic** | **Pr > ChiSq** | **Label** |
| **Test0** | Wald | 15.15 | 0.0005 | g11+g12+g13=0, g21+g22+g23=0 |
| **Test1** | Wald | 0.83 | 0.3629 | g12=g21 |

| **Number of Observations** | | **Statistics for System** | |
| --- | --- | --- | --- |
| **Used** | 2066 | **Objective** | 1.9952 |
| **Missing** | 0 | **Objective\*N** | 4122 |

Testing the homogeneity restriction:

H0: the system of equations is homogeneous

HA: the system of equations is not homogeneous

Since the Wald statistic is 15.15 and is significant, we reject the null hypothesis and conclude there is evidence to suggest our system of equations is not homogeneous

Testing the symmetry restriction:

H0: the system of equations is symmetrical

HA: the system of equations is not symmetrical

Since the Wald statistic is 0.83 and is not significant, we fail to reject our null hypothesis and conclude there is evidence to suggest our system of equations is symmetrical

* 1. **Estimate the system of equations imposing the demand restrictions using proc model and SUR. Use HCCME=3 if possible.**

|  |  |  |
| --- | --- | --- |
| **Var** | **W1** | **W2** |
| **Int** | 0.658629 (0.0185) | 0.198717 (0.0137) |
| **Ln p1** | 0.049592 (0.00796) | -0.03133 (0.00284) |
| **Ln p2** | -0.03133 (0.00293) | 0.038775 (0.00311) |
| **Ln p3** | -0.01826 (0.00879) | -0.00744 (0.00793) |
| **Ln x\*** | -0.02514 (0.00783) | 0.052559 (0.00638) |

3.6 Recover the parameters of the third equation using the demand restrictions.

Running the other two pairs of equations, we can recover the coefficients for our third equation as such:

|  |  |  |
| --- | --- | --- |
| Var | W1 | W3 |
| Int | 0.658616 (0.0179) | 0.142665 (0.0103) |
| Ln p1 | 0.049608 (0.00751) | -0.01827 (0.00400) |
| Ln p2 | -0.03134 (0.00389) | -0.00743 (0.00206) |
| Ln p3 | -0.01827 (0.00391) | 0.025703 (0.00541) |
| Ln x\* | -0.02514 (0.00779) | -0.02741 (0.00455) |

|  |  |  |
| --- | --- | --- |
| Var | W2 | W3 |
| Int | 0.198712 (0.0154) | 0.142681 (0.0106) |
| Ln p1 | -0.03133 (0.00489) | -0.01829 (0.00554) |
| Ln p2 | 0.038766 (0.00342) | -0.00744 (0.00207) |
| Ln p3 | -0.00744 (0.00206) | 0.025723 (0.00256) |
| Ln x\* | 0.052558 (0.00640) | -0.02741 (0.00455) |

As we can see, using either combinations of the remaining 2 equations will give us the coefficients for the last equation.

# **Section 2. Considering endogeneity of total expenditures in each demand equation**

**It is well documented that the variable x\* is endogenous. Thus, consider an additional equation model for real income, which is assumed to be exogenous:**

**When considering this additional equation, the equation system including the share demand equations and the income equation becomes a simultaneous equation system.**

Meat share equation:



Dairy products equation:



Pulses:



Real expenditures equation:

4.1. For the equation system comprised by share equations and the real expenditures equation, are the order conditions satisfied? Explain.

For w1

* Endogenous variables: x\*
* Exogenous variables: income
* Since exogenous variables (1) >= endogenous variable (1), the order condition is satisfied

For w2

* Endogenous variables: x\*
* Exogenous variables: income
* Since exogenous variables (1) >= endogenous variable (1), the order condition is satisfied

For w3

* Endogenous variables: x\*
* Exogenous variables: income
* Since exogenous variables (1) >= endogenous variable (1), the order condition is satisfied

For ln(x\*)

* Endogenous variables: none
* Exogenous variables: none
* Since exogenous variables (0) >= endogenous variable (0), the order condition is satisfied

Since the order condition is satisfied for every equation in our simultaneous equation model, we can conclude our SEM is indeed identified.

Overall identification of the system:

* 1. **Estimate the three share equations individually using 2SLS procedures to account for the endogeneity of total expenditures. For each equation report the estimation results using both OLS and 2SLS procedures.**
     1. **Use the necessary procedures to evaluate the weakness of the instrument (lnincome).**

To test the weakness of instruments, since we have one instrument for one endogenous variable, we would need to find the first stage t-value of the instrument, which is 14.18. The F-value is simply t squared, which would be 14.18^2=201. Since this is greater than 10, we have evidence to conclude ln income is not a weak instrument.

* + 1. **Test for heteroscedasticity and use a HCCME matrix if necessary.**

With White’s test for the 2SLS procedure on meats, dairy, and pulses, the values are 160.0, 258.3, and 171.2, respectively. Therefore, we have evidence to suggest we should reject the null hypothesis of homoskedasticity and conclude there is evidence of heteroskedasticity.

* + 1. **Report OLS/2SLS estimation results and discuss the differences in terms of coefficients and standard errors.**

|  |  |  |  |
| --- | --- | --- | --- |
| **OLS** | | | |
| **Var** | **W1** | **W2** | **W3** |
| **Int** | 0.649548  (0.0179) | 0.190137  (0.0165) | 0.160315 (0.00733) |
| **Ln p1** | 0.057045  (0.0127) | -0.02132  (0.0117) | -0.03572 (0.00521) |
| **Ln p2** | 0.057045 (0.00391) | 0.040623 (0.00361) | -0.00796 (0.00160) |
| **Ln p3** | -0.03266 (0.00778) | -0.02241 (0.00719) | 0.019969 (0.00319) |
| **Ln x\*** | -0.02535 (0.00599) | 0.053286 (0.00719) | -0.02794 (0.00246) |

|  |  |  |  |
| --- | --- | --- | --- |
| 2SLS | | | |
| Var | W1 | W2 | W3 |
| Int | 0.518422 (0.0485) | 0.20802 (0.0432) | 0.273559 (0.0234) |
| Ln p1 | 0.06556 (0.0180) | -0.02248 (0.0182) | -0.04308 (0.00657) |
| Ln p2 | -0.02907 (0.00408) | 0.040132 (0.00368) | -0.01107 (0.00202) |
| Ln p3 | 0.000061 (0.00774) | -0.02209 (0.00757) | 0.022025 (0.00319) |
| Ln x\* | 0.034782 (0.0215) | 0.045086 (0.0191) | -0.07987 (0.00983) |

* 1. **Estimate the simultaneous equation model involving the share and the expenditure equation using 3sls.**

4.2.a) Without imposing any restrictions. Compare the results with those obtained using SUR in 3.3. Present a Table and discuss the differences in terms of coefficients and standard errors. Use HCCME=3 if possible.

3SLS (Unrestricted)

|  |  |  |  |
| --- | --- | --- | --- |
| Var | W1 | W2 | Ln(x\*) |
| Int | 0.518422 (0.0485) | 0.20802 (0.0432) | 0.855154 (0.1192) |
| Ln p1 | 0.06556 (0.0180) | -0.02248 (0.0182) | -0.25884 (0.0533) |
| Ln p2 | -0.02907 (0.00408) | 0.040132 (0.00368) | -0.05128 (0.0163) |
| Ln p3 | 0.000061 (0.00774) | -0.02209 (0.00757) | -0.05128 (0.0163) |
| Ln x\* | 0.034782 (0.0215) | 0.045086 (0.0191) | 0.169977 (0.0141) |

SUR (Unrestricted)

|  |  |  |
| --- | --- | --- |
| Var | W1 | W2 |
| Int | 0.649548 (0.0250) | 0.190137 (0.0230) |
| Ln p1 | 0.057045 (0.0183) | -0.02132 (0.0185) |
| Ln p2 | -0.03266 (0.00374) | 0.040623 (0.00342) |
| Ln p3 | 0.002442 (0.00789) | -0.02241 (0.00765) |
| Ln x\* | -0.02535 (0.00783) | 0.053286 (0.00640) |

For the coefficients between 3SLS and SUR, they remain similar to each other and with matching signs. Using SUR, the standard errors tend to be smaller, although for some variables, 3SLS has smaller standard errors.

* 1. b) Imposing the demand restrictions. Compare the results with those obtained using SUR in 3.5. Present a Table and discuss the differences in terms of coefficients and standard errors. Use HCCME=3 if possible.

3SLS (Restricted)

|  |  |  |  |
| --- | --- | --- | --- |
| Var | W1 | W2 | Ln(x\*) |
| Int | 0.541609 (0.0446) | 0.223586 (0.0425) | 0.758587 (0.1201) |
| Ln p1 | 0.048512 (0.00722) | -0.02891 (0.00302) | -0.19362 (0.0533) |
| Ln p2 | -0.02891 (0.00302) | 0.038538 (0.00313) | -0.04832 (0.0163) |
| Ln p3 | -0.0196 (0.00848) | -0.00962 (0.00581) | 0.049879 (0.0313) |
| Ln x\* | 0.03224 (0.0212) | 0.039557 (0.0202) | 0.174059 (0.0142) |

SUR (Restricted)

|  |  |  |
| --- | --- | --- |
| **Var** | **W1** | **W2** |
| **Int** | 0.658629 (0.0185) | 0.198717 (0.0137) |
| **Ln p1** | 0.049592 (0.00796) | -0.03133 (0.00284) |
| **Ln p2** | -0.03133 (0.00293) | 0.038775 (0.00311) |
| **Ln p3** | -0.01826 (0.00879) | -0.00744 (0.00793) |
| **Ln x\*** | -0.02514 (0.00783) | 0.052559 (0.00638) |

As for the coefficients and standard errors between 3SLS and SUR when adding restrictions, they also remain similar to each other and with matching signs. Additionally, 3SLS is consistent and asymptotically efficient, so the 3SLS procedure in this example produces different estimates (and hopefully, more efficient).