

# Gretl Workshop 10

## Model Specification

Ramanathan data 7-17  
Cross Country Inequality Data

# Load Data

- Choose File/open data/sample file and select data7-17 from the Ramanathan tab.
- Data set to model cross country determinants of income inequality
- Read variable definitions.
  - Dependent variable is gini coefficient
  - Look for redundancies (similarities) among remaining variables

# Correlations

- From main window use shift-click to highlight all variables except Y.
- Choose View/correlation matrix
- Do any correlations appear to be large, especially between redundant pairs?

# Model 1 – All Variables

- General to specific approach – begin with complete model
- Choose Model/Ordinary least squares.
- Double-click on Y (dependent variable)
- Use shift-click to highlight remaining variables and add them as explanatory variables.

# Eliminate Redundancies

- edu and lit are similar; prefer lit on theoretical grounds. Why?
- In model 1 output window choose Tests/omit variables, and select edu to be omitted. Then Ok
- Scroll to bottom to see if model selection criteria improved.
- Any other model improvements?

# Eliminate Redundancies

- urb and agr may be similar.
- From Model 2 output window choose Tests/omit variables and select urb to be omitted. Do selection statistics improve?
- Return to Model 2 output window; choose Tests/omit variables and select agr to be omitted. Do selection statistics improve?
- Which variable should be omitted?

# New Base Model

- Model selection criteria suggest we can omit urb (but not agr) and edu.
- From main gretl window choose Model/ordinary least squares.
- Regress  $y$  on all variables except urb, edu.

# Diagnostic Tests

- From output window choose Tests/Ramsey's RESET and notice form of the regression for this test.
- What is null hypothesis of RESET test?
- What do you conclude from the test (scroll to bottom)?
- Close this window.

# Diagnostic Tests

- From base model output window choose Tests/heteroscedasticity to get White test.
- Notice form of this test regression and test statistic.
- What is null hypothesis and what do you conclude?
- Close this window

# Diagnostic Tests

- From equation output window choose Tests/normality of residual to get Jarque-Bera statistic.
- What is null hypothesis, and what is conclusion?
- From the histogram does there appear to be outliers problem? Skewness?

# Diagnostic Tests

- From equation output window choose Graphs/fitted, actual plot/against gnp (gnp=gnp per capita).
- Notice only 6 observations are for countries with  $\text{gnp} > 1000$ .
- Nonhomogenous sample: possible source of non-normality.

# Restrict sample to lower income countries

- In main gretl window choose Sample/restrict based on criterion, and type in `gnp<1000`
- Repeat base model regression (previous model with edu and urb excluded).
- Check normality of residuals.
- Apply RESET and heteroscedasticity tests again.

# Interpretation of results

- Model now satisfies diagnostic tests.
- What is major change in model estimates for the restricted sample?
- $Y$  is gini coefficient: increase in  $Y \Rightarrow$  increase in inequality.
  - Which effects on inequality are significant?
  - Give economic argument for the signs of significant coefficients.

# Econometric Lessons

- Model selection criteria, together with theoretical reasoning, can be useful in eliminating redundant variables.
- Use diagnostic tests as model selection criteria.
- Sample should be homogenous; small number of unusual observations may cause statistical problems (non-normality).