

Gretl Workshop 7

Multicollinearity

Ramanathan data 4-4
Bus Travel Data

Load Data & Take Logs

- Start gretl
- Choose File/open data/sample file and select data4-4 from the Ramanathan tab.
- Read variable definitions.
- Choose Data/Select all.
- Choose Add/logs of selected variables.

Demand Model for Bus Travel

- Specify model based on demand theory (own price, price of substitute, income).
- Why might the other three variables (DENSITY, POP, LANDAREA) be included? What signs do you expect for their coefficients? Should all three be included?
- Do you prefer logs or linear form?

Model 1 – All Variables, Logs

- Choose Model/ordinary least squares.
- Double-click on I_BUSTRAVL (dependent)
- Use shift-click to select and add the 6 remaining variables in their log form. Ok.
- Any evidence of multicollinearity?
- Choose Tests/collinearity to see VIF stats.
What do you conclude?

Exercise

- When all variables are in log form, there should be an exact linear relation among I_POP , $I_DENSITY$, and $I_LANDAREA$. Show this.
- This problem of exact multicollinearity means that Model 1 cannot be estimated. There are rounding errors in the data to break this exact multicollinearity.

Model 2 – Remove LANDAREA

- From main window choose Model/ordinary least squares.
- Remove I_LANDAREA; Ok
- What happens to significance of I_POP, I_DENSITY? Does model have omitted variables problem?
- Choose Tests/collinearity to see VIF stats. What do you conclude?

Model 2 - Correlations

- From main window use ctrl-click to highlight the five explanatory variables in this model (all I_ variables except I_LANDAREA and I_BUSTRVL).
- Choose View/correlation matrix/
- What do you conclude? Do any correlations appear to be large?

Model 2 – Return to equation output window

- What is the interpretation of the coefficient on `I_INCOME`? Is the negative sign surprising?
- The two price elasticities are insignificant. Is this due to multicollinearity? Explain.

Re-specify Model 2?

- Should variables with insignificant effects be eliminated? Explain.
- Remove I_GASPRICE from the model. Do other coefficients change much?
- Remove both prices from the model. Do other coefficients change much?
- What does this suggest about multicollinearity in Model 2?

Interpretation

- Removal of price variables had little effect on remaining parameter estimates, suggesting no omitted variables problem.
- Model 2 with price variables included is more informative.
- Sometimes insignificant results are important (bus travel is not responsive to price changes).

Lessons

- Insignificant coefficients are not always a sign of multicollinearity. Effects may be truly zero.
- Insignificant results may often be interesting.
- Report insignificant coefficients for full disclosure.