Supplemental Materials to Lab 4 of QM3

Shane Xinyang Xuan
ShaneXuan.com

Department of Political Science
University of California, San Diego

April 23, 2018

Exercise: 2SLS Estimation

- Load `grilic.dta` into Stata

- Construct time dummies for the 9 possible values for the `YEAR` variable of the form:

$$yr66 = \begin{cases} 1 & \text{if } \texttt{YEAR}_t = 66 \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

$$\vdots$$

$$yr73 = \begin{cases} 1 & \text{if } \texttt{YEAR}_t = 73 \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

- Estimate `LW` on `S`, `IQ`, and $\texttt{h} = (\texttt{expr}, \texttt{tenure}, \texttt{rns}, \texttt{smsa}, \texttt{yr66}, ..., \texttt{yr71})^T$ by 2SLS using $\{\texttt{med}, \texttt{kww}, \texttt{mrt}, \texttt{age}\}$ as instruments for `S`.

- Conduct an F test of the relevance of the instruments on the basis of the first-stage regression.

- Calculate Sargan's statistic for overidentification test of all instruments. Note that the statistic is asymptotically distributed as a chi-square variable with $(m - k)$ degrees of freedom, where m is the number of instruments and k is the number of endogenous variables.

- Do you have any concerns about either the relevance or the validity of the instruments?