Week 5 – Problem Set

1.

Consider the two-equation structural model

$$y_{1i} = \gamma_1 \, y_{2i} + \mathbf{z}_{1i}\boldsymbol{\delta}_1 + u_{1i}$$
$$y_{2i} = \gamma_2 \, y_{1i} + \mathbf{z}_{2i}\boldsymbol{\delta}_2 + u_{2i}$$

Solve for the reduced forms. What must be true for the reduced form parameters to be identified? What must be true for the structural coefficients to be identified?

2.

This problem considers some practical ways to develop instruments for nonlinear endogenous regressors using interaction terms and fitted values. Consider the model estimated in Problem 5c of Problem Set 3 using the card.xlsx data. In that model, $log(wage)$ is regressed on $educ$, $exper$, $exper^2$, $black$, $south$, $smsa$, $smsa66$, and all but one of the regional dummies using 2SLS; $educ$ is endogenous and instrumented with $near4c$.

(a) Add $educ^2$ as an endogenous regressor to the model in 5c and, in addition to $nearc4$ alone, use interaction terms with $nearc4$ as the instruments, i.e., $nearc4 * exper$, $nearc4 * exper^2$, $nearc4 * black$, etc. Perform a robust t-test of the significance of $educ^2$ in the model.

(b) Add $educ^2$ as an endogenous regressor to the model in 5c and use $(\widehat{educ})^2$ as its instrument, where $\widehat{educ}$ is the fitted value from a reduced form regression of $educ$ on all exogenous variables (including $nearc4$). Perform a robust t-test of the significance of $educ^2$ in the model.

(c) Add $black * educ$ as an endogenous regressor to the model in 5c and use interaction terms with $black$ as the instruments, i.e., $black * nearc4$, $black * exper$, $black * exper^2$, $black * south$, etc. Perform a robust t-test of the significance of $black * educ$ in the model.

(d) Add $black * educ$ as an endogenous regressor to the model in 5c and use $black * \widehat{educ}$ as its instrument, where $\widehat{educ}$ is defined as in (b). Perform a robust t-test of the significance of $black * educ$ in the model.

3.

This problem uses data from openness.xlsx, which contains 114 observations about import and inflation rates by country. These data were used by Romer (1993) ("Openness and Inflation: Theory and Evidence," *Quarterly Journal of Economics* 108, 869-903) to examine whether the openness of an economy leads to lower inflation rates. The variables include:

Variable Name	Variable Label
open	imports as a percentage of GDP, 1973-
inf	average annual inflation, 1973-
pcinc	per capita income in US dollars, 1980
land	land area in square miles
oil	=1 if major oil producer
good	=1 if "good" data

Assuming that $pcinc$, $land$, and oil are exogenous, consider the two-equation structural model

$$
\begin{aligned}
inf_i &= \delta_{10} + \gamma_{12}\, open_i + \delta_{11}\, log(pcinc_i) + \delta_{12}\, oil_i + u_{1i} \\
open_i &= \delta_{20} + \gamma_{21}\, inf_i + \delta_{21}\, log(pcinc_i) + \delta_{22}\, log(land_i) + u_{2i}
\end{aligned}
$$

(a) Is Equation 1 identified?

(b) Estimate the reduced form for $open$ and verify whether $log(land)$ is a valid instrument.

(c) Compare the OLS and 2SLS estimates of γ_{12} in Equation 1 and assess whether the results comport with economic theory.

(d) Add the term $\gamma_{13}\, open^2$ to Equation 1 as an endogenous regressor and estimate using 2SLS with $log(land)^2$ as its instrument. Perform a robust t-test of its statistical significance.

(e) Add the term $\gamma_{13}\, open^2$ to Equation 1 and perform 2SLS in two stages: (1) regress $open$ on a constant, $log(pcinc)$, $log(land)$, and oil to obtain the fitted values $\widehat{open}$; (2) then regress inf on a constant, $\widehat{open}$, $(\widehat{open})^2$, $log(pcinc)$, and oil. How does your estimate compare with that in (d)? Is this estimation strategy valid?