Problem set 13: week 17

Please go through the following alongside the Stata answer file.

1. Y_1 and Y_2 are jointly determined by the following system:

$$
\begin{aligned}
Y_{i1} =& \alpha_1 + \delta_1 Y_{i2} + u_{i1} \ (A) \\
Y_{i2} =& \alpha_2 + \delta_2 Y_{i1} + \beta_2 Z_{i1} + u_{i2} \ (B)
\end{aligned}
$$

Solving for Y_2 we get:

$$
\begin{aligned}
Y_2 =& \alpha_2 + \delta_2(\alpha_1 + \delta_1 Y_2 + u_1) + \beta_2 Z1 + u_2 \\
Y_2 =& \alpha_2 + \delta_2\alpha_1 + \delta_2\delta_1 Y_2 + \delta_2 u_1 + \beta_2 Z1 + u_2 \\
Y_2 =& \frac{\alpha_2 + \delta_2\alpha_1}{1-\delta_1\delta_2} + \frac{\beta_2}{1-\delta_1\delta_2} Z_1 + \frac{\delta_2 u_1 + u_2}{1-\delta_1\delta_2} \text{ simplifying} \\
Y_2 =& \pi_0 + \pi_1 Z_1 + v_2
\end{aligned}
$$

- Since $v_2 = \frac{\delta_2 u_1 + u_2}{1-\delta_1\delta_2}$ we can see that Y_2 will depend directly on u_1 if $\delta_2 \neq 0$.
- Therefore, $E[u_1|Y_2] \neq 0$: if we estimate (A) using OLS it will be a biased estimator of δ_1.

2. Solving for Y_1 we get:

$$
\begin{aligned}
Y_1 =& \alpha_1 + \delta_1(\alpha_2 + \delta_2 Y_1 + \beta_2 Z_1 + u_2) + u_1 \\
Y_1 =& \alpha_1 + \delta_1\alpha_2 + \delta_1\delta_2 Y_1 + \delta_1\beta_2 Z_1 + \delta_1 u_2 + u_1 \\
Y_1 =& \frac{\alpha_1 + \delta_1\alpha_2}{1-\delta_1\delta_2} + \frac{\delta_1\beta_2}{1-\delta_1\delta_2} Z_1 + \frac{\delta_1 u_2 + u_1}{1-\delta_1\delta_2} \text{ simplifying} \\
Y_1 =& \pi_0 + \pi_1 Z_1 + v_1
\end{aligned}
$$

- Since $v_1 = \frac{\delta_1 u_2 + u_1}{1-\delta_1\delta_2}$ we can see that Y_1 will depend directly on u_2 if $\delta_1 \neq 0$.
- Therefore, $E[u_2|Y_1] \neq 0$: if we estimate (B) using OLS it will be a biased estimator of δ_2.

3.

- Given Z_2 is in model (B) but not included in the model (A), we can theoretically use it as an IV for Y_2. That is, Z_2 is relevant as is correlated with Y_2, and it is exogenous as it is uncorrelated with the error in Y_1.

- There are no terms in the model (A) that are not in (B), i.e. there are no theoretically available instruments for Y_1.

4.

- Increasing the number of cigarettes smoked a day by one increases income by 0.1% on average. Note, however, this is very insignificant, i.e. there is no statistical effect.
- Increasing the wage by 1% increased the number of cigarettes smoked per day by 0.0088, on average. Again the coefficient is very insignificant, i.e. there is no statistical effect.
- Note, if there these variables are jointly determined these interpretations are based on biased estimators.

5.

- 1st mode1, example reason: smoking may effect income through days or productivity lost through poor health.
- 2nd model, example reason: the higher ones income the higher the demand for cigs (assuming they are a normal good among smokers).
- From the reduced form for cigarettes we may potentially use cigprice and smoking restrictions as IVs. We can test for relevance with an F-test.
- They are jointly significant at the 5% level. Exogeneity of the instruments is discussed below.

6. Now the coefficient on cigs is negative and almost significant at the 10% level against a two-sided alternative. The estimated effect is very large: each additional cigarette someone smokes lowers predicted income by about 4.2%. Note: the 95% CI for the coefficient on cigs is very wide. Do we believe this estimate: We may question the exogeneity of our instruments. Assuming that state level cigarette prices and restaurant smoking restrictions are exogenous in the income equation may be problematic. Incomes are known to vary by region, as do restaurant smoking restrictions. It could be that in states where income is lower (after controlling for education and age), restaurant smoking restrictions are less likely to be in place.