Homework 8

Empirical Analysis of Micro Data
Fumio Ohtake and Daiji Kawaguchi

1.

Labor economists have long time wanted to estimate the causal effect of having kids on mother's labor supply behavior. A straight forward way to estimate the effect would be estimating the following equation:

work= b0 + b1 kids + other back ground variables + error,

where work is the dummy variable indicating if the mother works and kids is the variable for the number of kids. However, it has been noticed that number of kids is not exogenous from mother's working behavior since those women who plans to work may not have kids; the causality may run in the opposite direction.

Recently, Angrist and Evans (1998) came up with a cleaver idea. Our casual observation or introspection tells us that those couples who have 2 kids in the same sex are more likely to have third kid than those couples who have 2 kids in the different sex, probably because couples prefer to have mixed sex composition in their kids. Angrist and Evans claim that the sex composition of first two kids is randomly determined and is independent from mothers' working behavior. They accordingly suggested using the first two kids sex composition as an instrument for the dummy variable indicating if the mother has more than two kids (i.e. kids>=3), which is the endogenous variable in the equation of mother's labor supply. They estimated following equation

work= b0 + b1 1[kids>=3] + other back ground variables + error,

using 1[first two kids are of same sex] as instrument for 1[kids>=3]. This study is a very good example of clever empirical study with natural experimental flavor. In this problem, you are asked to replicate their study using data set "labsup.dta" available from Wooldridge's web page.

1. 1.

Estimate

work= b0 + b1 1[kids>=3] + other back ground variables + error,

neglecting that work and 1[kids>=3] are binary variables. You should estimate the equation by OLS. The choice of appropriate control variables are left to you.

1.2.

Now consider the endogeneity of 1[kids>=3]. Use 1[first two kids are of same sex] as the instrument for 1[kids>=3]. Liner IV estimation suffices.

1.3.

Repeat 1.1. by Probit estimation. You can neglect the endogeneity of 1[kids>=3].

1.4

Test the endogeneity of 1[kids>=3] by Rivers and Vuong test.

1.5

Repeat 1.2. by Probit. You should now consider the endogeneity of 1[kids>=3]. You may either go with two-step estimation with an appropriate standard error correction or a full-blown maximum likelihood. If you take the first avenue, write down the variance matrix of second stage estimator that is the function of variance matrix of the first stage estimator. If you take the second avenue, write down the log likelihood function for each observation.

1.6.

Compare and discuss the difference of the results obtained from liner models and non-linear models. You should compare the results that are comparable.

2.

Solve 15.4, 15.6, 15.13, 16.3, 16.6, 16.8 in Wooldridge (2002).

Reference

Joshua Angrist and Bill Evans (1998) “Children and their parents’ labor supply: Evidence from Exogenous Variation in Family Size,” *American Economic Review*, June.