

Worksheet week # 8

1. In this exercise, you replicate the paper *Instrumental Variables Methods in Experimental Criminological Research: What, Why, and How?* by Joshua Angrist (2006), in which the results from Minneapolis Domestic Violence Experiment (MDVE) are reexamined. You should use the *MDVE.gdt* data.

The main idea of the experiment run in years 1981 and 1982 is described by Angrist in the paper:

In response to a politically charged policy debate as to the wisdom of making arrests in response to domestic violence, the MDVE was conceived as a social experiment that might provide a resolution. The research design incorporated three treatments: arrest, ordering the offender off the premises for 8 hours (separation), and some form of advice that might include mediation. The research design called for one of these three treatments to be randomly selected each time participating Minneapolis police officers encountered a situation meeting the experimental criteria (some kind of apparent misdemeanor domestic assault where there was probable cause to believe that a cohabitant or spouse had committed an assault against the other party in the past 4 hours). Cases of life-threatening or severe injury, i.e., felony assault, were excluded. Both suspect and victim had to be present upon the officer's arrival.

The randomization device was a pad of report forms that were randomly color-coded for each of the three possible response. Officers who encountered a situation that met the experimental criteria were to act according to the color of the form on top of the pad. The police officers who participated in the experiment had volunteered to take part, and were therefore expected to comply with the research design. On the other hand, deviations from random assignment were allowed and even anticipated by the experimenters.

Suspects were randomly assigned to treatment, but in some cases police officers did not comply to the assignment, especially if weapons were present and if the suspect was under influence of drugs. In some cases, suspects were arrested when random assignment called for separation or advice. Officers would arrest in these cases when a suspect attempted to assault an officer, a victim persistently demanded an arrest, or if both parties were injured.

Further, Angrist simplifies the setup of the experiment by dividing the actions of police officers into two groups: arrest and “coddling” (meaning separation or advice). In the dataset, we can find two dummy variables:

$$z_{\text{coddled}} = \begin{cases} 1 & \text{if the offender should have been coddled} \\ 0 & \text{if the offender should have been arrested} \end{cases}$$

$$d_coddled = \begin{cases} 1 & \text{if the offender was coddled} \\ 0 & \text{if the offender was arrested} \end{cases}$$

In other words, $z_coddled$ informs us about what the police officer's reaction in the case should have been and $d_coddled$ about what it really was.

The point of this exercise is to determine what is the effect of the police action on the re-offense status given by the dummy variable y :

$$y = \begin{cases} 1 & \text{if there was re-offense} \\ 0 & \text{if there was no re-offense} \end{cases}$$

- (a) Estimate the effect of the police action on the re-offense status using OLS, first when only the treatment dummy $d_coddled$ is included in the regression, second with other covariates (race, year and quarter dummies, drug influence and presence of weapons).
- (b) Interpret the coefficient of treatment dummy. Is treatment dummy exogenous to the re-offense status? Explain.
- (c) Is the dummy $z_coddled$ a valid instrument for the treatment dummy? State the two properties of an IV, and apply them to $z_coddled$.
- (d) Estimate the effect using 2SLS, first when only the treatment dummy $d_coddled$ is included as explanatory variable in the regression, second with other covariates as well (use the dummy $z_coddled$ as the instrument).
- (e) Compare the results of OLS and 2SLS and explain why OLS underestimates the effect of the treatment.