Using Instrumental Variables to Find Causal Effects in Public Health

Antonio Trujillo, PhD
John Hopkins Bloomberg School of Public Health
Department of International Health
Health Systems Program

October 30, 2013
Today’s Topics

- Motivation
- IV: Matching Through partial assignment
- Graphical representation of the IV approach
- The problem of endogeneity in the context of program evaluation
- The general idea behind the Instrumental Variables (IVs) approach in PE
- The conceptual framework
- Three fundamental characteristics of IVs
- Testing the statistical validity of IVs
- Additional issues
Empirical Work: Some examples

- The effect of treatment for acute myocardial infarction
  IV distance to facility (Newhouse and McClellan)

- Effect of health insurance on medical care use among people with diabetes
  IVs community coverage; community knowledge about disease

- Effect of informal care on labor force participation
  IV parents’ health

- Effect of cognitive skills on self-management among people with diabetes
  IV right/left hand

- Estimate birth outcome in teenagers
  IV mother’s schooling (Rosenzweig)
Matching to Partial Assignment

- In the case of IV estimations, there is an incentive to take the treatment. This incentive is random. In experimental design the treatment itself is randomized.

- The core of argument is to find a variable Z that affects X; but has not direct effect on Y.

- Then, $r(Y,Z)$ divided by $r(X,Z)$ provided an estimate of the parameter of interest.

- Because we reduce the variance of X, IV methods do not work well with small samples.
Graphical representation of IV

Figure 1

\[ \begin{align*}
X & \rightarrow Y & \leftarrow \varepsilon_y \\
\end{align*} \]

*In this case, \( r(Y,X) \) is causal effect of \( X \) on \( Y \)

Figure 2

\[ \begin{align*}
C & \rightarrow Y & \leftarrow \varepsilon_y \\
X & \leftarrow \varepsilon_x \\
\end{align*} \]

*In this case, \( r(Y,X) \) is NOT causal effect of \( X \) on \( Y \)

Figure 3

\[ \begin{align*}
Y & \leftarrow \varepsilon_y \\
\beta & \leftarrow \alpha \\
X & \leftarrow \varepsilon_x \\
\end{align*} \]

- IV method: \( r(Y,Z) = \alpha \beta \), and \( r(X,Z) = \alpha \)
- Ratio of both correlation gives the causal effect \( \beta \)
More formally

- Let assume that we are interested in the following causal relationship:
  \[ Y = \beta_0 + B_1 X + \epsilon \]  
  \[ (1) \]
- Taking the expected value of \( Y \) and rewrite this expected value as changes in \( Z \) will be:
  \[ E[Y] = E[\beta_0 + B_1 X + \epsilon] = \beta_0 + B_1 E[X] + E[\epsilon] \]
  \[ (2) \]
  \[ E[Y/Z=1] - E[Y/Z=0] = B_1 (E[X/Z=1] - E[X/Z=0]) + (E[\epsilon/Z=1] - E[\epsilon/Z=0]) \]
  \[ (3) \]
- Diving by changes in \( X \) produced by changes in \( Z \) yields the IV estimate:
  \[ \frac{E[Y/Z=1] - E[Y/Z=0]}{E[Y/Z=1] - E[Y/Z=0]} = B_1 + K \]
  \[ (4) \]
- Main takeaways: (i) the bias \( K \) will equal zero if \( Z \) is uncorrelated with \( \epsilon \); (ii) it is assumed that the parameters of interest are constant structural effects for the motivation of this IV estimation; (iii) the bias is big when \( Z \) and \( X \) are weakly correlated.
- Extending this result when \( Z \) takes more than two values lead to similar conclusions.
The Problem of Endogeneity in PE

- Endogeneity arises when:
  - program participation could be function of the outcome variable;
  - or when the error term in the outcome equation may be correlated with the error term in the participation equation. In other words, unobservable factors influence both outcome and program participation.

- In this case, OLS estimates of the outcome equation would be inconsistent.

- One solution to the problem is to model participation using IVs. Another solution would be to measure the unobserved factor.

- The problem is in finding good instruments, and evaluating what happened with bad instruments. Keep in mind that weak instruments would also bias the results.
The General Idea Behind IVs in PE

- The idea behind IVs is to replace the endogenous variable (e.g., program participation) with its predicted value using exogenous variables and the IV.

- However, the first step is to identify (at least one) variable that explains program participation:
  - Such that the only link between the IV and the outcome of interest happens through program participation

- Different statistical tests are used to identify good instruments and the optimal number of them. More about these tests later.

- Since IV estimates are consistent but not unbiased, analysts using IV should use large databases. In addition, one needs to be very careful with the use of correct standard errors when using IV estimates.
IVs key issues:
- exogenous variation in PP
- IV allows us to estimate the coefficient of interest without knowing what the omitted variables are

IVs are used to solve:
- omitted variable problem in causal relationship.
- bias in estimates from measurement problems
The conceptual framework

- If program participation is endogenous, we can model it as:

\[ Y_{ij} = X_{ij} \beta + P_j \alpha + e_{ij} \]

of interest to the evaluator is \( \alpha \).

\[ P_{ij} = X_{ij} \delta + Z_{ij} \kappa + \mu_{ij} \]

where \( Z_{ij} \) are instruments. Both errors could be correlated because the omitted variables influence both the outcome and program participation.

- The evaluator needs to distinguish between endogenous and exogenous variables in the case of simultaneous equations. Endogenous variables may have both direct and indirect effects.
The conceptual framework (II)

- To have a fully specified a structural model, be able to derive a set of reduced form equations from the structural equations.

- The key issue of identification is recovery of the structural parameters from the reduced form equation.

- The estimation methods usually implemented are:
  - Substitute the IVs for program participation and run OLS on outcome
  - 2SLQ: estimate the participation equation with exogenous variables and IVs and then substitute the estimated value in the outcome equation (this would work as an IV)

- Be careful with the standard errors when using IVs. Nowadays most statistical packages adjust the standard errors in the second stage to use the corrected ones.
Weak Instruments

- If you discover that you only have weak instruments, it is better to use OLS (assume program participation exogenous)

- Instrumental variables are weak when:
  
  - there is weak correlation between the IVs and the program participation;
  - strong correlation between IVs and the outcome of interest
  - strong correlation between IVs and the error term in the outcome equation
More on Weak Instruments

- One needs to keep in mind that it is difficult to find a variable (IV) that does not have a net direct effect on the outcome of interest. But even if one finds this variable, IV estimators are biased in finite samples.

- From equation (4) previously presented, it is easy to derived that the biased will be big when the denominator is close to zero (i.e, weak instrument). This relationship is independent of sample size. In this case, the IV would not contain relevant information about the causal effect even in cases with small standard errors.

- In empirical work, it is hard to sell the non-net direct effect assumptions. For this reason, IV that comes from naturally occurring variation are becoming more common in the literature.

- Yet, naturally occurring events may also have an impact in other causes of the outcome.
Characteristics of Good IVs

- Highly correlated with Program Participation (i.e., IV’s are not correlated with included variables)
- Uncorrelated with the error term in both equations (i.e. uncorrelated with outcome)
- Only link with outcome is through program participation
Testing the Validity of IVs

- Steps to test the validity of IVs:
  - Test exogeneity of program participation: problematic due to measurement characteristics.
  - Test for good instruments
    - $R^2$ test
    - IVs explain program participation
    - IV does not belong to the outcome of interest equation after controlling for program participation (over-identification test)
Test of Exogeneity of PP

- There are two common tests for exogeneity of PP:
  - Hausman test (most commonly used)
  - Spencer–Berk test
- Hausman test: (null hypothesis: PP is exogenous)
  - Get the reduced form of program participation (based on exogenous variables and IVs)
  - Get predicted PP
  - Run the outcome variable with PP and predicted PP and exogenous variables. The test of the null is whether the coefficient of the predicted PP is significantly different from zero.
- It is important to keep in mind that the test for exogeneity depends on the model specification (i.e., definition of exogenous variables and IVs)
**R² test**

- One should run the PP equation with only IVs:
  - If the R² in participation equation is high when one uses IVs, estimates from 1SLQ and 2SLQ would be similar.
  - If R² is low; then IV estimations would be practically meaningless. An OLS would be a better option.
  - There is not a definite cut off point; but the literature suggests 0.03-0.05. Below this, instrumental variables are not necessary, and weak.
One should implement the following steps:

- Run the PP equation with exogenous variables and IVs
- Run the PP equation with exogenous variables (not including IVs)
- Conduct a t-test for the IV coefficient
- In case of multiple IVs, one may conduct a Wald test or likelihood ratio test using the results from the two previous regressions

One would like to reject the null hypothesis (IVs are statistically significant)
Over-identification Test

- One should implement the following steps:
  - Run the outcome equation with exogenous variables and program participation
  - Run the outcome equation with exogenous variables, program participation and IVs
  - Conduct a T-test for the IV coefficient
  - In case of multiple IVs, one may conduct a Wald test or likelihood ratio test using the results from the two previous regressions

- One would like to accept the null hypothesis (i.e., after controlling for program participation, IVs do not explain the outcome of interest)
Additional issues

- IV provides the program effect for those for whom the IV influences their decision to participate.

- Yet, we may have people with different program impact. Therefore, IV estimates provide the Local Average Treatment Effect (LATE) instead of the Average Treatment Effect.

- It is important to keep in mind functional form issues for the first and the second stage. In practice, it is better to use a linear function in the first stage and use the correct non-linear function in the second stage.