Instrumental Variables (IV) Estimation: A Brief Introduction

Susan L. Ettner, Ph.D.
Professor
Division of General Internal Medicine and Health Services Research, UCLA

To Start, Some Definitions

- Exogenous variables are those determined outside the model ("pre-determined")
- Endogenous variables are those determined within the model (i.e., they depend on other variables in the model)
- Dependent variables ("outcomes") are endogenous
- Usually your independent variables ("predictors") are exogenous
- However, sometimes the predictors are also endogenous → problem

Endogeneity Bias

- Endogenous predictor → Potential endogeneity bias
- Sources of endogeneity
  - Omitted variables (or "treatment selection")
  - Reverse causality
  - Simultaneity
  - Measurement error
  - Auto-correlated errors
- Possible solution: Instrumental variables (IV)
- Today: Focus on treatment selection bias

The Treatment Selection Problem

- Suppose you wish to compare outcomes of:
  - Patients who receive an intervention (or are enrolled in a program, or get a treatment)
  - Patients in usual care (or who are not treated)
- Key point: Participation in the intervention, program or treatment is not randomized
- You believe patients either self-selected or were selected (e.g., by providers) into the treatment based on certain attributes, some unmeasured

The Problem (cont'd)

- Examples are "real-world" interventions or programs where it is not feasible to randomize
- Similar problem if you have randomization but are interested in estimating the "as-treated" effects instead of the intent-to-treat (ITT) effects
- Might be interested in as-treated effects when there is significant crossover due to non-adherence or contamination, since ITT conservative

The Problem (cont'd)

- Suppose unobserved attributes determining whether patients end up in intervention vs. usual care are also attributes affecting the outcome
- Then the indicator for treatment assignment is *endogenous*, leading to potential bias
- Easiest to think of this as a violation of the assumption that the treatment assignment indicator is uncorrelated with the error term
- Can lead to biased estimate of its causal effect on the outcome

Example

- Q: How does (good) glycemic control compare for patients with diabetes whose usual source of care (USC) is a diabetes specialist vs. PCP?
- Suppose patients who see specialists are unobservably sicker
- Indicator for specialist picks up effect of greater severity in addition to true causal effect
- Positive effect of seeing specialist is offset by the negative effect of being a sicker patient
- Beneficial effects of specialty care on glycemic control understated (biased towards zero).

Same Example, Different Omitted Variable

- Q: How does (good) glycemic control compare for patients with diabetes whose USC is a diabetes specialist vs. PCP?
- Suppose patients who see specialists are unobservably more motivated
- Indicator for specialist picks up effect of greater motivation in addition to true causal effect
- Positive effect of seeing a specialist reinforced by positive effect of being more motivated
- Beneficial effects of specialty care on glycemic control overstated (biased away from zero).

Treatment Selection Bias and the Residual

- Suppose you estimate a linear regression
- Recall that the residual (or "error") term captures the effects of all omitted and imperfectly measured variables
- Q: Are these omitted variables correlated with your covariates (the so-called X's)?
  - *If no*: Model will have lower explanatory power and estimates may be less precise
  - *If yes*: The coefficient estimates may also be biased for those X's

Implications of Omitted Variables

- If you can't measure a variable that predicts your outcome and is also correlated with your treatment → bias
  - Just as important to think about what you *aren't* controlling for as what you *are* controlling for
  - Conceptual model helps here
- Likely direction of bias depends on correlation of omitted variable with both regressor and outcome
  - Caveat: Sometimes competing effects → can't know direction of the bias

Determining the Direction of Bias (simple case – no 2nd order effects)

True model: $Y = \alpha X_1 + \beta X_2 + \varepsilon$

Model run: $Y = \varphi X_1 + \mu$

Think of running auxiliary regression of X_2 on X_1:

$X_2 = \theta + \rho X_1 + \eta$

Instead of α, estimated coefficient on X_1 will be $\varphi = \alpha + \rho\,\beta$, which could be $> \alpha$ or $< \alpha$, depending on the signs of α, ρ and β

Direction of Omitted-Variable Bias (Simple Case)

True model: $Y = \alpha X_1 + \beta X_2 + \varepsilon$ X_2 not measured	**X_1 and X_2 are positively correlated** ($\rho > 0$)	**X_1 and X_2 are negatively correlated** ($\rho < 0$)
$\alpha > 0, \beta > 0$	α^{hat} biased away from zero	α^{hat} biased toward zero
$\alpha > 0, \beta < 0$	α^{hat} biased toward zero	α^{hat} biased away from zero
$\alpha < 0, \beta > 0$	α^{hat} biased toward zero	α^{hat} biased away from zero
$\alpha < 0, \beta < 0$	α^{hat} biased away from zero	α^{hat} biased toward zero

Note: Estimate is $\alpha + \rho\,\beta$ instead of α, so bias is $\rho\,\beta$

Unobserved Severity Example

True model: $Y = \alpha X_1 + \beta X_2 + \varepsilon$ X_2 not measured	**X_1 and X_2 are positively correlated** ($\rho > 0$)	**X_1 and X_2 are negatively correlated** ($\rho < 0$)
$\alpha > 0, \beta > 0$	α^{hat} biased away from zero	α^{hat} biased toward zero
$\alpha > 0, \beta < 0$	α^{hat} biased toward zero	α^{hat} biased away from zero
$\alpha < 0, \beta > 0$	α^{hat} biased toward zero	α^{hat} biased away from zero
$\alpha < 0, \beta < 0$	α^{hat} biased away from zero	α^{hat} biased toward zero

Y=glycemic control, X_1=specialist USC, X_2=severity

Unobserved Motivation Example

True model: $Y = \alpha X_1 + \beta X_2 + \varepsilon$ X_2 not measured	**X_1 and X_2 are positively correlated** ($\rho > 0$)	**X_1 and X_2 are negatively correlated** ($\rho < 0$)
$\alpha > 0, \beta > 0$	α^{hat} biased away from zero	α^{hat} biased toward zero
$\alpha > 0, \beta < 0$	α^{hat} biased toward zero	α^{hat} biased away from zero
$\alpha < 0, \beta > 0$	α^{hat} biased toward zero	α^{hat} biased away from zero
$\alpha < 0, \beta < 0$	α^{hat} biased away from zero	α^{hat} biased toward zero

Y=glycemic control, X_1=specialist USC, X_2=motivation

Some Possible Solutions (But No Panacea)

Possible Methods for Addressing Bias Due to Treatment Selection

- Propensity score methods
  - Designed to better exploit observed attributes and prevent out-of-sample prediction
  - Does not necessarily help with bias due to unobservables (in some cases exacerbates it)
  - Can use R&R method to test sensitivity
- Heckman "treatment effects" models
  - Simultaneous estimation of treatment and outcome equations
  - Usually requires strong assumptions about the joint distribution of the error terms

Methods for Addressing Bias (cont'd)

- Quasi-experimental designs using longitudinal data
  - Compare change over time among treatment vs. controls to "net out" heterogeneity
  - Can only "difference out" attributes that are constant over time
  - Assumes that secular time trends are the same in the absence of the intervention
  - Treatment selection could violate this assumption
  - Hard to check without a long "pre" period

Instrumental Variables Methods

IV in Action

Even after adjusting for a fairly extensive set of baseline severity measures...

- Depression treatment was associated with *worse* outcomes in the PIC study
- Medicaid insurance was associated with *higher* mortality in the HCSUS study

Both of these results reversed themselves when IV methods were used.

Instrumental Variables (IV)

- The idea is to achieve quasi-randomization using an instrumental variable (IV) that has a direct impact on the endogenous regressor but only an indirect impact on outcome
- In our original example, a valid IV would influence whether the patient's USC was a diabetes specialist vs. PCP but would affect glycemic control only through the patient's USC
- Exogenous variation in the instrument allows us to isolate the causal effect of the endogenous regressor on the outcome

A Simple Example of How IV Works

Q: Does having a diabetes specialist as USC improve glycemic control?

Stylized Fact #1: Suppose that glycemic control is better among patients whose USC is a diabetes specialist.

But…does this mean that having a diabetes specialist as USC improves glycemic control, or that patients who are more motivated to keep their diabetes under control are more likely to seek out a specialist as their USC?

A Simple Example (cont'd)

Stylized Fact #2: Suppose that patients who live in areas with a high density of diabetes specialists (relative to PCPs) are more likely to have a diabetes specialist as their USC.

Stylized Fact #3: Finally, suppose that patients who live in areas with a high density of diabetes specialists have better glycemic control.

A Simple Example (cont'd)

Implication:

If we can assume that the relative density of diabetes specialists in an area does not *directly* influence an individual patient's glycemic control…

Then the only way to explain the better glycemic control among patients living in areas with greater density of diabetes specialists is through whether the patient's USC is a diabetes specialist.

In Summary…

In this example, we "know":
High specialist density → patient has specialist as USC
High specialist density → patient has better glycemic control before controlling for whether patient has specialist as USC

We assume:
High specialist density does not → patient has better glycemic control after controlling for whether patient has specialist as USC

This means:
Having a specialist as USC *causally* → better glycemic control

Identification

The casual impact of having a diabetes specialist as USC on glycemic control is "identified" through our assumption that relative provider specialist density (the "instrument"):

(1) affects whether the patient has a specialist as USC

(2) does not directly affect the patient's glycemic control

If either assumption fails, we cannot draw this causal inference.

Two-Stage Least Squares (2SLS)

- How do you get the IV estimate?
- Most common IV method is 2SLS, which is a special case when you are using a linear regression for your outcome and your endogenous regressor is either continuous or can be linearized
- Other (more complex) methods available if your variables aren't continuous
- Here, we'll just treat everything as linear

1st-Stage (or "Reduced-Form") Regression

- Estimate regression of the endogenous regressor on all exogenous variables in the system
- This includes all predictors of the outcome (even those that normally don't belong in a treatment equation!) as well as the IVs
- Our example: Estimate the prob(patient had a diabetes specialist as USC) as a function of the density of diabetes specialists in the area and all predictors of glycemic control

Substituting Predicted for Actual Values

- Next, use the regression estimates from the 1st stage to construct a predicted value for the endogenous regressor
- Substitute this predicted value for the actual value of the endogenous regressor in the outcome equation
- Our example: Get predicted (linear) probability of having a diabetes specialist as USC for each patient, use instead of the original indicator for specialist USC

2nd-Stage (or "Structural") Regression

- After substituting predicted value for actual value of endogenous variable, estimate the outcome equation as you normally would
- Note: The IVs are excluded from this regression
- Standard software automatically adjusts SEs for the use of a predicted value
- Estimate glycemic control as a function of predicted prob(specialist USC), controlling for all other predictors of glycemic control (but not provider density)

Intuition (Such As It Is!)

- IV replaces the actual indicator for whether the patient had a specialist as USC with the predicted probability of having a specialist USC
- Use of prediction "breaks" its correlation with the error term in the glycemic control regression
- The predicted values are linear combinations of the exogenous variables, so by construction are not correlated with the error term.
- Thus you get an unbiased estimate of its causal effect on glycemic control.

Instruments and the "Exclusion" or "Identifying" Restriction

- Instruments for having a diabetes specialist as USC are variables included in the 1st- but not 2nd-stage regression.
- To "identify" the effect of diabetes specialty USC on glycemic control, we need at least one variable that:
  - affects specialty care
  - does not directly influence glycemic control after controlling for having specialty USC and other covariates.

Intuition Behind "Identification"

- To determine the effect of specialty care on glycemic control, we want to see how glycemic control changes when there is an exogenous shift in specialty care, *holding everything else determining glycemic control constant*.
- Thus, we need to find something that will shift around specialty care without changing any of the other predictors of glycemic control.

Importance of Instrument

- Our example: Instrument is relative density of diabetes specialists in the area.
- *If* provider density:
  - does not actually affect the patient's probability of seeing a specialist, or
  - has a direct influence on glycemic control (perhaps because areas with high density of diabetes specialists have more, or higher-quality, other resources for achieving glycemic control)
- *Then* we can't separate out independent causal effect of seeing a specialist.

Importance of Other Controls

- Usually easier to meet and test criterion that the IV (specialist provider density) predicts the endogenous regressor (whether specialist USC).
- More often, exclusion restriction fails (e.g., provider density can't affect glycemic control after controlling for specialist USC)
  - Also you can only test this assumption if you have >1 instrument
- More likely to meet this assumption if you control well for other confounders, but then can be harder to meet first assumption

Interpretation

- For which patients does IV estimation yield an unbiased estimate of the treatment effect?
- At a minimum, the IV estimate applies to the "marginal" people, i.e., those for whom a change in the instrument (provider density) would change the endogenous regressor (glycemic control).
- If the treatment effect is (assumed to be) homogeneous, then (by definition) the estimate generalizes to entire population.

Limitations of IV Analysis

- It's difficult to find a valid instrument
  - Sometimes instruments are so weak that IV is actually more biased than OLS, even in quite large samples
  - Even if your instrument is a strong predictor of treatment, often hard to argue that it's excludable from the second-stage outcome equation, again leading to bias
- If you don't have a valid instrument, then you also don't have a valid test for endogeneity

Limitations (cont'd)

- Key question: Is the correlation of the IV with the endogenous regressor high *relative* to its correlation with the outcome?
- The greater the correlation of the IV with the outcome, the stronger its correlation with the endogenous regressor needs to be.
- Crown et al. (2011) have a nice simulation paper showing that IV estimates have less error than OLS estimates only under circumstances close to ideal

Limitations (cont'd)

- If you have multiple potential IVs, with at least one known (or assumed to be) valid, can test the "overidentifying restrictions" on the model
- If only one instrument, can't test the exclusion restriction
- Even with a plausible instrument, IV estimates are less efficient than single-equation estimate
- Especially true if your endogenous regressor is dichotomous, so you lose precision/power

Future IV Topics

- Testing endogeneity (Hausman or augmented regression test)
- Testing the strength of the instrument
  - Staiger & Stock (1997): Need partial F statistic ≥ 5
- Testing the exclusion restriction (Sargan, 1984)
- The other three IV assumptions!
- IV with nonlinear outcomes (two-stage residual inclusion)
- IV with dichotomous endogenous regressors
- IV with multiple endogenous regressors