Instrumental Variables: A Tool to Reduce Bias in Non-Randomized Studies

Mark Bounthavong, PharmD, MPH, Student Chapter President, University of Washington, Seattle, WA, USA; **Blythe Adamson, MPH**, Student Chapter Vice-President, University of Washington, Seattle, WA, USA; and **Anirban Basu, PhD**, Director, Pharmaceutical Outcomes Research and Policy Program, University of Washington, Seattle, WA, USA

KEY POINTS . . .

In standard ordinary least square, unbiased estimates of the average treatment effect can be calculated when the exposure variable is not correlated with the error term.

However, in non-experimental studies, the exposure variable is oftentimes correlated with the error term and yields a spurious conclusion driven not by the exposure variable but by the association between the exposure variable and error term.

An instrumental variable is used to reduce this bias, but two conditions must be met: (1) strong correlation with the exposure variable and (2) no correlation with the error term.

The following article represents the second in a series that highlights local student chapter activities and research talents. In this section, we present a primer on using instrumental variables in outcomes research for students and experienced professionals alike.

Introduction

When random assignment of an exposure *(D)* is not possible in a study, we are confronted with a challenge to draw a causal inference. Random assignment mitigates bias by balancing observed and unobserved characteristics at baseline across exposure groups. Non-experimental studies without randomization lead to selection bias, as the distribution of factors that affect outcomes may be imbalanced between the exposed and unexposed groups. These factors are called confounders.

In a simple linear regression framework estimated using ordinary least squares (OLS), such confounding is addressed by controlling for these factors as covariates *(X)* in the regression:

$$Y = \beta_0 + \beta_1 D + \beta_2 X + \varepsilon,$$

where, under some general assumptions, β_1 represents the average treatment effect (ATE) parameter. Alternatively, propensity score methods can balance the distribution of these observed confounders between groups [1]. However, it is almost impossible to justify that all possible confounders are observed in the data at hand and have been adjusted for. Consequently, the effects of an exposure estimated from non-experimental studies using traditional methods are always subject to bias from the unobserved confounders. In the regression setting, such unobserved confounding is indicated when the exposure variable *D* is correlated with the error term ε, and therefore, OLS produces biased estimates for β_1. The exposure variable *D* is consequently *endogenous*.

This problem is further illustrated with a path analysis in Figure 1. Panel A shows the direct effect of the explanatory variable *D* on outcome Y. Standard OLS produces unbiased estimates of β_1 if there is no association between all other factors that affect Y (i.e., error) and exposure D. But if we observe the correlation in Panel B between exposure and error, the OLS will generate biased estimates that are not equal to the ATE. The OLS estimate would pick up a spurious correlation between D and Y that is not the causal effect of D, but rather driven by the association of ε with D and Y.

Instrumental Variables

This endogeneity problem can be addressed with an instrumental variable (IV), which is a naturally occurring factor that directly affects the endogenous variable in question, but does not affect outcomes in any other way. An IV can act as a natural randomizer in the absence of artificial random assignment, and therefore, helps tease out the causal effect of the exposure even in the presence of unobserved confounding.

Figure 1. Path diagrams for the regression model if (A) following the Ordinary Least Squares (OLS) regression assumption of no correlation between the exposure D and error term ε, (B) a correlation between exposure and error that is common in non-randomized studies, and (C) instrument variable Z is added to reduce bias in non-randomized study design.

A. OLS assumption

B. Correlation between error ε and exposure D

C. Adding instrument variable Z to reduce bias
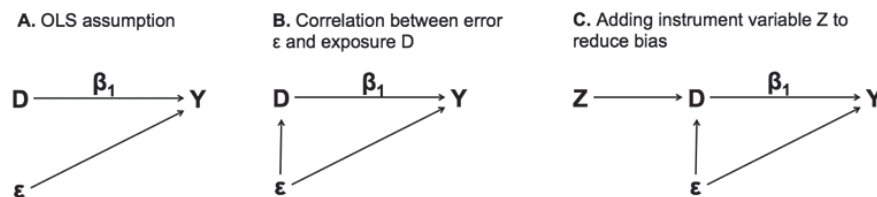
For an IV method to be a consistent estimator for β_1, two critical conditions should be met. First, the IV must be strongly correlated with the exposure variable given X,

$$Cov(Z, D|X) \neq 0,$$

and second the IV should be uncontaminated, (i.e., it cannot be correlated with the error term),

$$Cov(Z, \varepsilon|X) = 0.$$

In Panel C, the added IV Z is associated with the exposure but not the error term. The first assumption is testable. In fact, weak instruments usually create more bias than it helps alleviate. Typically, a rule of thumb is that the F-stat for an IV predicting D should be at least 10 [2]. The second assumption is not testable, and one must rely on theoretical rationales and some necessary statistical tests to support this assumption. However, if unobserved confounding is large, even moderately contaminated IVs can produce results that are more robust than naïve regression [3].

Why Is This So Important?

Unobserved variables correlated with both the exposure and outcomes variables can strongly influence the estimated relationship between the two and mask the magnitude of the causal effect. Good instrument variables help limit that biased inference.

Some final points to consider about IV analyses: IV should never be used to fix bad study designs; bad IV can cause more problems than solutions; and IV methods vary based on the type of regression methods used, such as two-stage predictor substitution and two-stage residual inclusion [4].

Conclusions

Studies taking advantage of IV have illuminated the causal relationships in Vietnam veterans with decreased wages [5], early birth season with lower educational attainment [6], and the effect of cardiac catheterization on mortality. Although challenging, selecting the best IV while meeting all the criteria established above can result in unbiased estimators of the ATE. More importantly, establishing causal relationships with non-experimental studies provides meaningful evidence to guide decision makers and policy makers.

References

[1] Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. Biometrika 1983;70:41–55. [2] Staiger D, Stock JH. Instrumental variables regression with weak instruments. Econometrica 1997;65:557–86. [3] Basu A, Chan KCG. Can we make smart choices between ols and contaminated IV methods? Health Econ 2014;23:462–72. [4] Terza JV, Basu A, Rathouz PJ. Two-stage residual inclusion estimation: Addressing endogeneity in health econometric modeling. J Health Econ 2008;27:531–43. [5] Angrist JD. Lifetime earnings and the Vietnam era draft lottery: Evidence from social security administrative records. Am Econ Rev 1990;80:313–36. [6] Angrist J, Keueger AB. Does compulsory school attendance affect schooling and earnings? Q J Econ 1991;106:979–1014. [7] Stukel TA, Fisher ES, Wennberg DE, et al. Analysis of observational studies in the presence of treatment selection bias: Effects of invasive cardiac management on ami survival using propensity score and instrumental variable methods. JAMA 2007;297:278–85. ■

Additional information:
The preceding article was part of the ongoing effort to share best practices and communicate lessons learned from ongoing education and service activities. Go to page 45 for this issue's Student Corner for more student news.

ISPOR STUDENT & MEETING TRAVEL GRANTS

Applications for ISPOR Student & Meeting Travel Grants are now being accepted for the ISPOR 21st Annual International Meeting in Washington, DC; ISPOR 7th Asia Pacific Conference in Singapore, and the ISPOR 19th Annual European Congress in Vienna, Austria.

APPLICATION DEADLINES:
ISPOR 21ST ANNUAL INTERNATIONAL MEETING: **MARCH 11, 2016**

ISPOR 7TH ASIA PACIFIC CONFERENCE: **JUNE 17, 2016**

ISPOR 19TH ANNUAL EUROPEAN CONGRESS: **19 AUGUST 2016**

> *To apply for a Student Travel Grant*
  http://www.ispor.org/student/Travel/grantApp.asp.

> *To apply for an ISPOR Meeting Travel Grant*
  http://www.ispor.org/awards/MeetingTravelScholarship.asp.