An Introduction to Instrumental Variables

Instrumental variables (IVs) are used to control for confounding and measurement error in observational studies. They allow for the possibility of making causal inferences with observational data. Like propensity scores, IVs can adjust for both observed and unobserved confounding effects. Other methods of adjusting for confounding effects, which include stratification, matching and multiple regression methods, can only adjust for observed confounders. IVs have primarily been used in economics research, but have recently begun to appear in epidemiological studies.

Observational studies are often implemented as a substitute for or complement to clinical trials, although clinical trials are the gold standard for making causal inference. The main concern with using observational data to make causal inferences is that an individual may be more likely to receive a treatment because that individual has one or more co-morbid conditions. The outcome may be influenced by the fact that some individuals received the treatment because of their personal or health characteristics.

Consider the linear regression model:

$$y_i = b_0 + b_1 X_{1i} + b_2 X_{1i} + ... + b_k X_{ki} + u_i ,$$

where

y_i is the outcome for the *i*th individual;

X_{mi} is the *m*th explanatory variable (*m* = 1 ,…, *k)* for the *i*th individual;

b_m is the parameter associated with the *m*th explanatory variable;

u_i is the random error term for the *i*th individual.

Let *Z* denote a randomization assignment indicator variable in this regression model, such that *Z* = 1 when a treatment is received and *Z* = 0 when the control or placebo is received, and let X_1 be the treatment. Z is referred to as the instrumental variable because it satisfies the following conditions:

(i) *Z* has a casual effect on *X;*
(ii) *Z* affects the outcome variable *Y* only through *X* (*Z* does not have a direct influence on *Y* which is referred to as the exclusion restriction);
(iii) There is no confounding for the effect of *Z* on *Y*.

There are two main criteria for defining an IV:

(i) It causes variation in the treatment variable;
(ii) It does not have a direct effect on the outcome variable, only indirectly through the treatment variable.

A reliable implementation of an IV must satisfy these two criteria and utilize a sufficient sample size to allow for reasonable estimation of the treatment effect. If the first assumption is not satisfied, implying that the IV is associated with the outcome, then estimation of the IV effect may be biased. If the second assumption is not satisfied, implying that the IV does not affect the treatment variable then the random error will tend to have the same effect as the treatment. When selecting an IV, one must ensure that it only affects whether or not the treatment is received and is not associated with the outcome variable.

Although IVs can control for confounding and measurement error in observational studies they have some limitations. We must be careful when dealing with many confounders and also if the correlation between the IV and the exposure variables is small. Both weak instruments and confounders produce large standard error which results in imprecise and biased results. Even when the two key assumptions are satisfied and the sample size is large, IVs cannot be used as a substitute for the use of clinical trials to make causal inference, although they are often useful in answering questions that an observational study can not. In general, instrumental variables are most suitable for studies in which there are only moderate to small confounding effects. They are least useful when there are strong confounding effects.

Instrumental Variables: A Brief Annotated Bibliography

Angrist, J.D. & Krueger, A.B. 2001. Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments. *Journal of Economic Perspectives, 15(4)*, 69-85.

In economics, IVs are used to determine which factors influence demand without affecting cost, or the factors that influence cost without affecting demand. In the discipline of epidemiology, primarily with respect to natural experiments, IVs are used (1) to counteract issues with measurement error in explanatory variables which result from a lack of accurate information available for analysis and (2) to overcome the issue of omitted variables in order to make casual inference in observational studies when randomization is infeasible or unethical. The fewer the number of instruments incorporated into the model, the smaller the bias. If the number of instruments is equivalent to the number of treatment or endogenous variables then the bias is approximately zero.

There are two main issues that may arise in the application of IVs: (1) we may choose a bad instrument which would result from the IV being correlated with the omitted variables or (2) bias may result if the instruments are only weakly correlated with the treatment variable(s). IVs are often difficult to interpret because they do not affect the behavior of every subject, although this makes them valuable in providing estimates for a specific group, such as the subjects who received treatment. There is also the issue of the LATE (Local Average Treatment Effect) which refers to the bias associated with the fact that subjects who receive treatment are often those that would not have taken treatment otherwise. This estimate is usually unreliable because it is only representative if every subject has a similar response to the treatment.

Greenland, S. 2000. An Introduction to instrumental variables for epidemiologists. *International Journal of Epidemiologists, 29*, 722-729.

IVs have primarily been used in the economic discipline but have recently been integrated into the field of epidemiology. They are used to control for confounding and measurement errors in observational studies so that causal inferences can be made. This paper discusses the application of instrumental variables for confounding control, non-compliance and misclassification correction in non-experimental research.

Hernan, M.A. & Robins, J.M 2006. Instruments for Causal Inference- An Epidemiologists Dream? ***Epidemiology, 17*****(4), 360-372.**

This paper discusses the implementation of IVs to estimate the average causal effect of an exposure on the outcome of interest and the conditions that must be satisfied to achieve consistent estimates of the causal effect. The three main conditions that define an instrumental variable are: (i) *Z* has a casual effect on *X*, (ii) *Z* affects the outcome variable *Y* only through *X* (*Z* does not have a direct influence on *Y* which is referred to as the exclusion restriction), and (iii) There is no confounding for the effect of *Z* on *Y*. The variable *Z* is the randomization assignment indicator ($Z = 1$ when treatment is received and $Z = 0$ when the control or placebo is received), *X* is the actual treatment received and *Y* is the outcome variable.

The IV estimator appears to be an "epidemiologist's dream" because of its ability to facilitate causal inference from observational studies even if the confounders are unmeasured. However, there are limitations that may hinder the use of IVs. These include (1) the three specified conditions may not be satisfied, and (2) the issue of most epidemiologic exposures being time-varying (the treatment variable is not considered time-varying but in reality the subjects may discontinue or change treatment), which is not taken into consideration when using IV analysis.

Martens, E.P., Pestman, W.R., de Boer, A., Belitser, S.V., & Klungel, O.H. (2006). Instrumental Variables: Application and Limitations. ***Epidemiology, 17*****(3), 260-267.**

This paper discusses the application of instrumental variables to the field of epidemiology. Instrumental variables have the advantage of being able to adjust for all confounders including unobserved ones like propensity scores and unlike most other adjustment methods such as stratification, matching and multiple regression methods.

When implementing instrumental variable analyses, one must be careful when dealing with small sample sizes, a large number of confounders or weak instruments. When the correlation between the IV and the exposure variables is small, there will be increased standard error and bias even if the sample size is large. Therefore, the IV must satisfy the necessary conditions and be selected with caution in order to prevent the potential bias that may result. Generally, IV analysis is suitable when there is moderate confounding and less useful when there is strong confounding.

Newhouse, J.P & McClellan, M. 1998. Econometrics in Outcomes Research: The Use of Instrumental Variables. ***Annual Review of Public Health, 19,*** **17-24.**

IV analysis is often applied to outcomes research, which involves studying the consequences of a treatment using observational data in order to monitor and improve quality of care. The core underlying issue associated with the use of observational data is that a subject may be more likely to receive treatment because they have a certain characteristic that someone who did not receive treatment does, such as a co-morbidity. In order to address for the issue of confounding in observational studies, IVs come into use because of their ability to control for observed and unobserved confounders in an observational study. There are two main assumptions that IVs hold for reliable implementation: (1) They cause variation in the treatment variables and (2) they do not have a direct effect on the outcome variable (only indirectly through the treatment variable). This allows the researcher to determine the level of exogenous variation, which is how much the variation in the treatment variable affects the outcome variable. Although IVs are often useful in answering questions that an observational study cannot, they cannot be used as a substitute for clinical trials. Observational studies have the benefit of being able to include a more general population whereas clinical trials produce results that are not as generalizable and may be compromised because subjects dealing with serious co morbidities may be excluded from the trial. Therefore the results obtained from a clinical trial will have more internal validity and those obtained from an observational study will have more external validity.

Wunsch, H., Linde-Zwirble, W.T. & Derek, C.A. 2006. Methods to adjust for bias and confounding in critical care health services research involving observational data. *Journal of Critical Care, 21*, 1-7.

There are several methods that can adjust for the confounding and bias in observational studies. These include matching, stratification, multivariate adjustment, propensity scores and IVs. Each method has strengths and limitations. The IV approach has the advantage of requiring only a single variable. As well, it can be applied to problems where other types of adjustment are impossible or difficult to implement. IVs have had their primary application in economics but in the health field have been used to study quality of care. The key difference between confounding variables and IVs is that IVs do not directly influence the outcome variable whereas confounding variables do.

Some Examples of Research Papers that Use Instrumental Variables

Stukel, T. A., Fisher, E. S., Wennberg, D. E., Alter, D. A., Gottlieb, D. J. & Vermeulen, M. J. (2007). Analysis of Observational Studies in the Presence of Treatment Selection Bias: Effects of Invasive Cardiac Management on AMI Survival Propensity Score and Instrumental Variable Methods. *The Journal of the American Medical Association, 297*(3), 278-285.

Web Link:
http://jama.ama-assn.org.proxy2.lib.umanitoba.ca/cgi/reprint/297/3/278
Pdf Link:

Adobe Acrobat Document

This study compares the outcomes of applying four different methods to remove the selection bias that results from using observational data: multivariate model risk adjustment, propensity score risk adjustment, propensity-based matching and instrumental variables analysis. In conclusion, instrumental variable analysis was proven to be the most effective in producing the most unbiased estimates of the treatment effects whereas the remaining methods had similar restrictions with respect to removing selection bias. In this particular study, the treatment variable is invasive cardiac management, the instrumental variable is regional catheterization rate and the outcome variable is AMI survival rate. The instrumental variable plays the role of randomizing the subjects by region into so called "treatment groups," then compares groups of subjects with respect to their likelihood of receiving treatment. They measure the estimated treatment effect on the marginal population which is all the subjects with the exception of those that would never or always receive cardiac catheterization. Instrumental variable analysis was the most effective in removing bias and produced less biased estimates, smaller SEs, and provided closer approximations to the average population effects than RCTs. The method is the most applicable in making policy-relevant decisions but often does not address issues related to clinical studies.

Pilote, L., Beck, C.A., Eisenberg, M.J., Humphries, K., Joseph, L., Penrod, J.R. et al. 2008. Comparing invasive and noninvasive management strategies for acute myocardial infarction using administrative databases. *The American Heart Journal, 155*(1), 42-48.

Web Link:
http://www.mdconsult.com.proxy2.lib.umanitoba.ca/das/article/body/92619328-2/jorg=journal&source=&sp=20206435&sid=0/N/622909/s0002870307008290.pdf?issn=00028703
Pdf Link:

Adobe Acrobat Document

Earle, C. C., Tsai, J. S., Gleber, R. D., Weinstein, M. C., Neumann, P .J. & Weeks, J.C. (2001). Effectiveness of Chemotherapy for Advances Lung Cancer in the Elderly: Instrumental Variables and Propensity Analysis. *Journal of Clinical Oncology, 19*(4), 1064-1070.

Web Link:
http://jco.ascopubs.org/cgi/reprint/19/4/1064
Pdf Link:

Adobe Acrobat Document

Beck, C. A., Penrod, J., Gyrokos, T. W., Shapiro, S. & Pilote, L. (2003). Does Aggressive Care Following Acute Myocardial Infarction Reduce Mortality? Analysis with Instrumental Variables to Compare Effectiveness in Canadian and United States Patient Populations. *Health Services Research, 38* (6, Pt 1), 1423- 1440.

Web Link:
http://www.pubmedcentral.nih.gov.proxy2.lib.umanitoba.ca/picrender.fcgi?artid=1360957&blobtype=pdf
Pdf Link:

Adobe Acrobat Document

Brooks, J. M., Chrischilles, E. A., Scott, S. D. & Chen-Hardee, S. S. (2003). Was Breast Conserving Surgery Underutilized for Early Stage Breast Cancer? Instrumental Variables Evidence for Stage II Patients for Iowa. *Health Services Research, 38* (6, Pt 1), 1385-1402.

Web Link:
http://www.pubmedcentral.nih.gov.proxy2.lib.umanitoba.ca/picrender.fcgi?artid=1360955&blobtype=pdf

Pdf Link:

Adobe Acrobat Document