Review of "Identification of Causal Effects Using Instrumental Variables"

Yichen Ji

Abstract

The seminal paper "Identification of Causal Effects Using Instrumental Variables" by [2] provides a comprehensive framework for the identification and estimation of causal effects using instrumental variables (IV) in the presence of endogeneity. This work is pivotal in econometrics and statistics, offering clarity on the application of IV methods, particularly in observational studies where randomized experiments are impractical. The paper highlights the formal introduction of the Local Average Treatment Effect (LATE), which details the conditions under which IV estimates can be interpreted as causal effects for a specific subpopulation defined by the instrument. This literature review aims to briefly unpack the proposed framework's motivation, assumptions, mathematical foundations, and implications.

1 Motivation: Structural Equation Models in Economics

The pursuit of understanding causal relationships, as opposed to mere associative observations, occupies a central place in the field of econometrics and economics in general. This quest encompasses a broad spectrum of inquiries, from evaluating the ramifications of educational attainment on earnings to dissecting the outcomes of employment training programs on labor market trajectories and analyzing the influence of inputs on firm outputs. Traditionally, the economics community has relied on structural equation models, underpinned by instrumental variables (IV), to navigate through these complex causal effects. For example, the dummy endogenous variable model infers the effect of veteran status on a health outcome:

$$\begin{aligned} Y_i &= \beta_0 + \beta_1 \cdot D_i + \varepsilon_i, \\ D_i^* &= \alpha_0 + \alpha_1 \cdot Z_i + \nu_i \\ D_i &= \begin{cases} 1 & \text{if } D_i^* > 0 \\ 0 & \text{if } D_i^* \leq 0. \end{cases} \end{aligned} \tag{1}$$

where for person i, Y_i is the observed health outcome, D_i is the observed treatment (i.e., veteran status), and Z_i is the observed draft status, β_1 is the causal effect of D on Y, D^* reflects the reality that compliance is a decision based on weighing the anticipated benefits of serving against those of not serving (compile or deny).

Notably, the identification of β_1 requires two critical assumptions for Z_i to be the instrumental variable (IV) in this example, namely exogeneity and relevance:

$$E[Z_i \cdot \varepsilon_i] = 0, \quad E[Z_i \cdot \nu_i] = 0 \tag{2}$$

$$\operatorname{cov}(D_i, Z_i) \neq 0 \tag{3}$$

Intuitively speaking, exogeneity assumption (2) means that the instrument should be uncorrelated with the error term in the outcome equation, which implies that the instrument affects the outcome only through its impact on the treatment variable, not through any direct or unobserved paths. Relevance assumption (3) infers that the instrument must be associated with the treatment variable, that is, changes in the instrument must lead to changes in the treatment, ensuring that the instrument has a substantial impact on the exposure of interest.

However, these critical assumptions also turn out to be the reasons why the structural equations approach is not widely used in the statistics community. First, structural equation models are critically sensitive to their underlying assumptions and have shown difficulties in replicating experimental findings, raising concerns about their reliability in accurately capturing causal relationships. Second, the assumptions in structural equation models often rely on abstract disturbances from unspecified regression functions, making it hard to interpret these models and communicate findings due to the lack of directly observable variables.

Parallel developments in statistics, particularly the Rubin Causal Model (RCM) based on the potential outcomes framework, offer a complementary viewpoint [3]. [2] proposes to reconcile these parallel tracks, thereby enriching the methodological toolbox for causal inference, especially in scenarios under noncompliance and nonignorable treatment allocation conditions.

2 Framework

2.1 Rubin Causal Model (RCM) $\cap$ Instrumental Variable (IV)

The Rubin causal model (RCM) provides a rigorous methodology for estimating the population-level causal effects of treatments or interventions on outcomes by focusing on potential outcomes. The RCM posits that each unit (e.g., individual, school, district) has a set of potential outcomes associated with each possible treatment level. The causal effect for a unit is defined as the difference between the potential outcomes under the treatment and control conditions. However, since only one of these outcomes can be observed for each unit, the true causal effect for any single unit is fundamentally unobservable.

2.1.1 Notation

Let $\mathbf{Z}$ be the N-dimensional vector of assignments with i th element Z_i, and let $D_i(\mathbf{Z})$ be an indicator for whether person i would serve given the randomly allocated vector of draft assignments $\mathbf{Z}$. Both Z and D take binary values, that is, there is no partial compliance. [2] defines the causal effect for individual i of Z on D as

$$D_i(1) - D_i(0)$$

and the causal effect of Z on Y as

$$Y_i\left(1, D_i(1)\right) - Y_i\left(0, D_i(0)\right)$$

and the causal effect of D on Y as

$$Y_i(1) - Y_i(0)$$

We focus on average causal effects in groups of people who can be induced to change treatments. Inferences about such average causal effects are made using changes in treatment status induced by treatment assignment, provided the assignment does affect the treatment.

2.1.2 Assumptions

The formal definition of an instrument in the RCM needs 5 assumptions:

1. Stable Unit Treatment Value Assumption (SUTVA) [4]:
   - If $Z_i = Z_i'$, then $D_i(\mathbf{Z}) = D_i\left(\mathbf{Z}'\right)$.
   - If $Z_i = Z_i'$ and $D_i = D_i'$, then $Y_i(\mathbf{Z}, \mathbf{D}) = Y_i\left(\mathbf{Z}', \mathbf{D}'\right)$.
2. Random Assignment: The treatment assignment Z_i is random $\Pr(\mathbf{Z} = \mathbf{c}) = \Pr\left(\mathbf{Z} = \mathbf{c}'\right)$ for all $\mathbf{c}$ and $\mathbf{c}'$ such that $\iota^T \mathbf{c} = \iota^T \mathbf{c}'$, where ι is the N dimensional column vector with all elements equal to one:
3. Exclusion Restriction: $\mathbf{Y}(\mathbf{Z}, \mathbf{D}) = \mathbf{Y}\left(\mathbf{Z}', \mathbf{D}\right)$ for all $\mathbf{Z}, \mathbf{Z}'$ and for all $\mathbf{D}$.
4. Nonzero Average Causal Effect of Z on D: The average causal effect of Z on D, $E\left[D_i(1) - D_i(0)\right]$ is not equal to zero.
5. Monotonicity [1]: $D_i(1) \geq D_i(0)$ for all $i = 1, \ldots, N$.

2.2 Local Average Treatment Effect (LATE)

Under SUTVA and Exclusion Restriction assumptions, the causal effect of Z on Y for person i is the product of the causal effect of D on Y and the causal effect of Z on D:

$$
\begin{aligned}
&Y_i\left(1, D_i(1)\right)-Y_i\left(0, D_i(0)\right)\\
&=Y_i\left(D_i(1)\right)-Y_i\left(D_i(0)\right)\\
&=\left[Y_i(1)\cdot D_i(1)+Y_i(0)\cdot\left(1-D_i(1)\right)\right]-\left[Y_i(1)\cdot D_i(0)+Y_i(0)\cdot\left(1-D_i(0)\right)\right]\\
&=\left(Y_i(1)-Y_i(0)\right)\cdot\left(D_i(1)-D_i(0)\right)
\end{aligned}
$$

As a consequence, in addition to the monotonicity assumption, we can write the average causal effect of Z on Y as:

$$
E\left[Y_i\left(D_i(1),1\right)-Y_i\left(D_i(0),0\right)\right]=E\left[\left(Y_i(1)-Y_i(0)\right)\mid D_i(1)-D_i(0)=1\right]\cdot P\left[D_i(1)-D_i(0)=1\right] \quad (4)
$$

Accordingly, the concept of the Local Average Treatment Effect (LATE) is introduced to establish the relationship between the IV estimand and the causal effect of D on Y. Under the aforementioned assumptions, the LATE as the instrumental variables estimand is defined as:

$$
LATE=\frac{E\left[Y_i\left(D_i(1),1\right)-Y_i\left(D_i(0),0\right)\right]}{E\left[D_i(1)-D_i(0)\right]}=E\left[\left(Y_i(1)-Y_i(0)\right)\mid D_i(1)-D_i(0)=1\right] \quad (5)
$$

Note that the monotonicity assumption implies that $E\left[D_i(1)-D_i(0)\right]$ equals $P\left[D_i(1)-D_i(0)=1\right]$, and the nonzero average causal effect assumption implies that $E\left[D_i(1)-D_i(0\right]$ differs from zero.

(5) targets a more specific estimate of causal effects compared to what is typically derived from traditional analyses, focusing on the subgroup of individuals who comply with their treatment assignment due to an instrumental variable. This group, known as compliers, consists of those whose treatment status is directly influenced by the IV, allowing for a clear delineation of causality in situations where not all participants adhere to their assigned treatment.

3 Discussion

[2] has broad implications for empirical research, particularly in economics and social sciences, where endogeneity and unobservable variables frequently bias causal estimates. The paper elucidates how IV methods, under the LATE framework, can provide more accurate and interpretable estimates of causal effects, emphasizing the importance of carefully selecting and justifying instrumental variables.

The work of [2] has profoundly influenced the field of econometrics and beyond, offering a robust framework for addressing endogeneity through instrumental variables. By introducing and formalizing the concept of the Local Average Treatment Effect (LATE), they provide a clearer interpretation of IV estimates, contributing to more accurate and reliable causal inference in observational studies. [2] remains a cornerstone in the methodology of causal inference, guiding researchers in the application and interpretation of instrumental variable techniques.

References

[1] Joshua Angrist and Guido Imbens. *Identification and estimation of local average treatment effects*. 1995.

[2] Joshua D Angrist, Guido W Imbens, and Donald B Rubin. "Identification of causal effects using instrumental variables". In: *Journal of the American statistical Association* 91.434 (1996), pp. 444–455.

[3] Paul W Holland. "Statistics and causal inference". In: *Journal of the American statistical Association* 81.396 (1986), pp. 945–960.

[4] Donald B Rubin. "Bayesian inference for causal effects: The role of randomization". In: *The Annals of statistics* (1978), pp. 34–58.