Paper 1

Abstract Title Page

Not included in page count.

Title: Under What Assumptions do Site-by-Treatment Instruments Identify Average Causal Effects?

Author(s): Sean F. Reardon (Stanford University, contact author)
sreardon@suse.stanford.edu
Stephen W. Raudenbush (University of Chicago)
sraudenb@uchicago.edu

Abstract Body

Background / Context:

In canonical applications of the instrumental variable method, exogenously determined exposure to an instrument (such as random assignment to a treatment condition) induces exposure to a mediating process that in turn causes a change in a later outcome. A crucial assumption known as the exclusion restriction is that the hypothesized instrument can influence the outcome only through its influence on exposure to the mediator of interest (Heckman & Robb, 1985; Imbens & Angrist, 1994). It may be the case, however, that multiple mediators operate jointly to influence the outcome, in which case a single instrument will not suffice to identify the causal effects of interest.

To cope with this problem, analysts have recently exploited the fact that a causal process is often replicated across multiple sites, generating the possibility of multiple instruments in the form of site-by-instrument interactions. These multiple instruments can, in principle, enable the investigator to identify the impact of multiple programs or treatments regarded as the mediators of the effect of an instrument. Kling, Liebman, and Katz (2007), for example, used random assignment in the Moving to Opportunity ("MTO") study as an instrument to estimate the impact of neighborhood poverty on health, social behavior, education, and economic self-sufficiency of adolescents and adults. Reasoning that the instrument might affect outcomes through mechanisms other than neighborhood poverty, they control for a second mediator, use of the randomized treatment voucher. To do so, they capitalize on the replication of the MTO experiment in five cities, generating ten instruments (the five sites generate ten site-by-treatment interactions as instruments because there were three randomly assigned treatment conditions per site) to identify the impact of the two mediators of interest, neighborhood poverty and experimental compliance. Using a similar strategy, Morris, Duncan, and Rodriguez (2010) used data from 16 implementations of welfare-to-work experiments to identify the impact of family income, average hours worked, and receipt of welfare as mediators.

Clearly, this strategy for generating multiple instruments has potentially great appeal in research on causal effects in social science. For example, Spybrook (2009) found that, among 75 large-scale experiments funded by the US Institute of Education Sciences over the past decade, the majority were multi-site studies in which randomization occurred within sites. In principle, these data could yield a wealth of new knowledge about causal effects in education policy. It is essential, however, that researchers understand the assumptions required to pursue this strategy successfully. To date, we know of no complete account of these assumptions.

Purpose / Objective / Research Question / Focus of Study:

Our purpose is to clarify the assumptions that must be met if this "multiple site, multiple mediator" strategy, hereafter referred to as "MSMM," is to identify the average causal effects (ATE) in the populations of interest.

Setting:

N/A

Population / Participants / Subjects:

N/A

Intervention / Program / Practice:

N/A

Significance / Novelty of study:

We are aware of only two studies that rely on site-by-treatment interactions to generate multiple instruments to estimate the impacts of multiple potential mediators (Kling et al., 2007; Morris et al., 2010). Neither of these studies discusses the assumptions underlying the models they fit. Given the large number of multi-site randomized trials that have been conducted in education and in other fields (Spybrook, 2009), such models are likely to become increasingly appealing as a means to estimating the effects of multiple hard-to-randomize potential mechanisms. Ours is the first paper that systematically describes the identifying assumptions of such models.

Statistical, Measurement, or Econometric Model:

We begin by delineating the assumptions required for identification in the case of a single instrument and a single mediator within a single-site study. Unlike Angrist, Imbens, and Rubin (1996) (hereafter AIR), we consider the general case where the mediator may be continuous or multi-valued. In this general case, the assumptions required for identification of the average treatment effect differ somewhat from those AIR (1996) describe for the binary mediator case. We link our discussion to recent papers describing the correlated random coefficient model, and show that the CRC model is identified by instrumental variables using a weaker assumption than that described by Heckman and Vytlacil (1998) and Wooldridge (2003).

Following a discussion of the single, site, single mediator case, we then consider the case of multiple sites with a single mediator before turning to the case of primary interest: MSMM. Finally, we generalize these results to any setting in which multiple instruments identify the impact of multiple mediators.

Usefulness / Applicability of Method:

Researchers often want to understand the mechanisms through which a specific treatment, program, or policy operates. Although it may be feasible to randomize individuals to specific treatments or programs, it is often not feasible to assign individuals to processes that are hypothesized to be mediators. For example, in the Reading First Impact Study (Gamse et al.,

2008), a regression discontinuity design provided exogenous variation in Reading First study, enabling the researchers to estimate the effects of the Reading First program as a whole. However, the program was hypothesized to work through its impacts on five distinct dimensions of teachers' reading instruction practices. These practices could not be randomly assigned, however (because the researchers cannot control what teachers do in the classroom). One could imagine using the site-by-instrument interactions in an IV model to estimate the effects of specific instructional practices. There are many such cases in educational and social science research. Thus, a clearer understanding of the necessary assumptions (and the consequences of their failure) may improve the quality of research on hard-to-randomize educational processes and mechanisms.

Research Design:

We consider the case of a study design in which there are multiple sites, indexed by $s \in \{1,2,\dots,J\}$. In each site s, individuals are assigned to one of multiple possible values of a treatment, T, where T may be binary or continuous. The treatment T may affect a vector of P mediators, $\mathbf{M} = \{M_1, M_2, \dots, M_P\}$; each of the mediators may also be binary or continuous. We describe the person-specific effect of T on a mediator M_p as the person-specific "compliance" with respect to mediator M_p. Each of the mediators may each affect an outcome, Y. We are interested in estimating the average effect of each of the mediators on the outcome Y.

Data Collection and Analysis:

N/A

Findings / Results:

We identify 9 assumptions that must be met in order that the MSMM IV model identifies the average effects of each mediator in the population of interest. These are:

1. Stable unit treatment value assumption (SUTVA): each unit has one and only one potential outcome under each treatment condition: that is, for a population of size n, $Y_i(t_1, t_2, \dots, t_n) = Y_i(t_i)$ for all $i \in \{1,2,\dots,n\}$. In the IV model, this standard SUTVA assumption is actually composed of two distinct SUTVA assumptions:

    a. For each mediator p, each unit i has one and only one potential value of M_{pi} for each treatment condition t: that is, for a population of size n, $M_{pi}(t_1, t_2, \dots, t_n) = M_{pi}(t_i)$ for all $i \in \{1,2,\dots,n\}$.
    b. Each unit i has one and only one potential outcome value of Y_i for each vector of mediators $\mathbf{M}_i = \{M_{1i}, M_{2i}, \dots, M_{Pi}\}$: that is, for a population of size n, $Y_i(\mathbf{M}_1, \mathbf{M}_2, \dots, \mathbf{M}_n) = Y_i(\mathbf{M}_i)$ for all $i \in \{1,2,\dots,n\}$.

2. Exclusion restriction: the treatment T affects Y only through its impact on the set of P mediators, $\mathbf{M} = \{M_1, M_2, \dots, M_P\}$. That is, $Y(t) = Y\big(t, \mathbf{M}(t)\big) = Y\big(\mathbf{M}(t)\big)$.

3. Person-specific linearity of each mediator in T: the person-specific effect of T on each mediator M_p is linear. That is, $M_p(t) = M_p(0) + t\ \ _p$ for each p.

4. Person-specific linearity in $\mathbf{M}$: the person-specific effect of each mediator M_p is linear. That is, $Y\big(\mathbf{M}(t)\big) = Y\big(\mathbf{M}(0)\big) + \sum_{p=1}^{P}\ \ _p\big(t\ \ _p\big)$.

5. Ignorable within-site treatment assignment: the assignment of the instrument (the treatment, in our notation) must be independent of the potential outcomes within each site: $T \perp Y(t)|s,\ T \perp \mathbf{M}(t)|s, \forall\, t, s$.

6. No average within-site compliance-effect covariance: $E\left(Cov_s\big(\ \ _p,\ \ _p\big)\right) = 0. \forall p \in \{1,2,\dots,P\}$. A simpler, but stronger, assumption is that there is no within-site compliance-effect covariance in any site: $Cov_s\big(\ \ _p,\ \ _p\big) = 0, \forall p \in \{1,2,\dots,P\}, \forall s \in \{1,2,\dots,S\}$.

7. Site-by-mediator compliance matrix has sufficient rank. In particular, if G is the S x P matrix of the γ_{sp} s, then $rank(G) = P$. This implies three specific conditions:

    a. The compliance of at least $P - 1$ of the mediators varies across sites. That is, $Var\big(\gamma_{ps}\big) = 0$, for at most one $p \in \{1,2,\dots,P\}$.
    b. There are at least as many sites as mediators: $P \leq S$.
    c. There is some subset of Q site-specific average compliance vectors, $\mathbf{\Gamma}_s = \{\gamma_{1s}, \gamma_{2s}, \dots, \gamma_{Ps}\}$, where $S \geq Q \geq P$, that are linearly independent.

8. Parallel mediators: assignment to T does not influence a given mediator M_p through any other mediator $M_q, q \in \{1,2,\dots,P\}$. That is, the mediators do not influence one another. That is, $M_p\big(t, M_1, \dots, M_{p-1}, M_{p+1}, \dots, M_P\big) = M_p(t), \forall p \in \{1,2,,\dots,P\}$.

9. Mean independence of the site-average compliances and effects. Within each site s, let $\gamma_{ps} = E[\ \ _p|s]$ and $\delta_{ps} = E[\ \ _p|s]$ be the site-average compliance and effect of mediator p, respectively. Likewise, let $\gamma_p = E[\ \ _p]$ and $\delta_p = E[\ \ _p]$ be the average compliance and average effect of mediator p in the population. Then we assume that the site-average effects are independent of the site average compliances. That is, $E\big[\delta_{ps}|\gamma_{1s}, \gamma_{2s}, \dots, \gamma_{Ps}\big] = E\big[\delta_{ps}\big] = \delta_p\ \forall p \in \{1, \dots, P\}$.

Some of these are familiar assumptions from the standard IV model in the case where there is a single site ($J = 1$) and a single mediator ($P = 1$). In this case, assumptions 1-7 are sufficient for the IV model to identify the average effect of the mediator on Y in the population. Assumptions 8 and 9 pertain only when there are multiple mediators (assumption 8) or multiple sites (assumption 9). Moreover, in this case, if both T and M are binary, then assumptions 3 and

4 are unnecessary; assumptions 1, 2, 5, 6 and 7 are sufficient to identify the average effect of M on Y in the population.

We note that assumptions 1, 2, 5, and 7 are generalizations of four of the five IV assumptions identified by Angrist, Imbens, and Rubin (1996) [note that in the case when $S = P = 1$, the sufficient rank assumption is equivalent to the ―nonzero average causal effect of T on M" assumption made by AIR]. Their fifth assumption, montonicity, is notably absent from our set of assumptions, however. We show that the monotonicity assumption is relevant only in the case where both T and M are binary and where we wish to identify the average effect of M on Y among the population of ―compliers", in Angrist, Imbens, and Rubins' (1996) terminology. If we wish to identify the average effect of M on Y in the population (arguably an estimand of more general interest), then no montonicity assumption is needed. Instead, we require assumption 6, the no compliance-effect covariance assumption. Moreover, this assumption is relevant even when T and M are not binary. Thus, our set of assumptions is more generally applicable than the AIR assumptions, and identifies an estimand that is of more general interest.

We note also that assumption 6 (no within-site compliance-effect covariance) is weaker than the assumption that Heckman and Vytlacil (1998) and Wooldridge (2003) make in order to identify average effects in the correlated random coefficients model. We show that the assumption of no compliance effect covariance (along with the other standard IV assumptions) is sufficient to identify the average effects using an IV model when the data are generated from a correlated random effects process.

Conclusions:

Our investigation of the assumptions of the multiple-mediator, multiple-site IV model demonstrates that such models rely on a large number of non-trivial assumptions. Of most importance are the assumptions regarding the relationship between compliance and effect. Any correlation—whether within or between sites—between compliance (with respect to any meditator) and effect (again, of any mediator) will potentially bias the estimates of the effect of any mediator. Because individuals or groups (sites) with the most to gain from a specific practice (mediator) may be more likely to comply with treatment assignment, there is good reason to worry that compliance-effect covariances are non-zero in many cases.

Appendices

Not included in page count.

Appendix A. References

Angrist, J. D., Imbens, G. W., & Rubin, D. B. (1996). Identification of Causal Effects Using Instrumental Variables. *Journal of the American Statistical Association, 91*(434), 444-455.

Gamse, B. C., Jacob, R. T., Horst, M., Boulay, B., & Unlu, F. (2008). *Reading First Impact Study Final Report*. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Heckman, J. J., & Robb, R. (1985). Using Longitudinal Data to Estimate Age, Period and Cohort Effects in Earnings Equations. In W. M. Mason & S. E. Feinberg (Eds.), *Cohort Analysis in Social Research Beyond the Identification Problem*. New York: Springer-Verlag.

Heckman, J. J., & Vytlacil, E. (1998). Instrumental Variables Methods for the Correlated Random Coefficient Model: Estimating the

Average Rate of Return to Schooling When the Return is Correlated with Schooling. *The Journal of Human Resources, 33*(4), 974-987.

Imbens, G. W., & Angrist, J. D. (1994). Identification and Estimation of Local Average Treatment Effects. *Econometrica, 62*(2), 467-475.

Kling, J. R., Liebman, J. B., & Katz, L. F. (2007). Experimental Analysis of Neighborhood Effects. *Econometrica, 75*(1), 83-119.

Morris, P., Duncan, G. J., & Rodrigues, C. (2010). Does Money Really Matter? Estimating Impacts of Family Income on Young Children's Achievement with Data from Random-Assignment Experiments. *Unpublished manuscript*.

Woolridge, J. M. (2003). Further results on instrumental variables estimation of average treatment effects in the correlated random coefficient model. *Economics Letters, 79*, 185-191.