Abstract Title Page

Title:

Weighting Methods for Assessing Policy Effects Mediated by Peer Change

Authors and Affiliations:

Guanglei Hong, the University of Chicago

Takako Nomi, Consortium on Chicago School Research at the University of Chicago

Abstract Body

Background / Context:

A major advancement in the recent statistics literature on mediation has been the clarification of conceptual distinctions between controlled direct effects and natural direct and indirect effects (Pearl, 2001, Robins & Greenland, 1992). The total effect of a treatment can be decomposed into the natural direct effect and the natural indirect effect. In contrast, a controlled direct effect of the treatment on the outcome would be conceivable if a second intervention could hypothetically hold the mediator at a fixed level. The controlled direct effect and the natural direct effect will be equal if the controlled direct effect does not depend on the mediator value, which is conventionally assumed in path analysis (Holland, 1988; Sobel, 2008). Recent extensions of the path analysis method relax this assumption by invoking model-based assumptions with regard to how the treatment, the mediator, and the covariates interact in the structural model for the outcome (Pearl, 2010; Petersen, Sinisi, & van der Lann, 2006; VanderWeele & Vansteelandt, 2009). The instrumental variable approach requires, among other identification assumptions, the exclusion restriction that constrains its applicability when a nonzero direct effect is hypothesized.

Purpose / Objective / Research Question / Focus of Study:

This study introduces a new set of weighting procedures for revealing the mediation mechanism in multi-level settings. These methods are illustrated through an investigation of whether the impact of a system-wide policy change on student outcomes is mediated by policy-induced peer composition change. When the policy changed not only lower-achieving students' course-taking but also the ability composition of math classes among other concurrent changes, unpacking the overall policy impact on math achievement is challenging. To illustrate, our causal questions focus on decomposing the total policy effect into the indirect effect mediated by peer composition change and the direct effect of the policy for the subpopulation of lower achieving students. Specifically, we ask: (1) Did the increase in peer ability mediate the policy effect on these students' math achievement? (2) Would the policy have a direct effect on these students' math achievement if their peer composition had remained unchanged by the policy?

Population / Participants / Subjects:

Chicago has the third largest public school system in the United States. The population of interest includes all 59 neighborhood high schools in existence before and after the algebra-for-all policy was introduced in 1997. We select one pre-policy cohort and one post-policy cohort of first-time ninth graders ineligible for special education services. The subpopulation of interest here includes students who would probably take remedial math in the pre-policy year and were expected to experience an improvement in peer ability in the post-policy year.

Intervention / Program / Practice:

The policy of interest here required algebra for all ninth graders enrolled in the Chicago Public Schools (CPS) in 1997 and thereafter. Among students with relatively lower math incoming skills, only half of them took algebra in the pre-policy years. As schools increased algebra enrollment and eliminated remedial math, they often created mixed-ability algebra classes by enrolling lower-achieving students in the same classes with higher-achieving peers.

Significance / Novelty of study:

Through predicting every student's probability of taking algebra and expected peer composition before and after the policy, we identify a principal stratum of lower-achieving students who would experience a change in course-taking and a rise in peer ability due to the policy change. We develop strategies to adjust for not only observed and unobserved between-cohort differences that were not a result of the policy but also between-cohort differences that could have been a result of the policy and could have confounded the peer effect on student outcomes. The total effect is then decomposed through an innovative weighting procedure.

Statistical, Measurement, or Econometric Model:

For student i in school j, let $Z_{ij} = 1$ if the student attended the ninth grade in the post-policy year and 0 otherwise. If the student's potential outcome values do not depend on how the treatment was assigned and what treatments were received by other students in the population, under this Stable Unit Treatment Value Assumption (SUTVA) (Rubin, 1986), the total effect of the policy for this student is simply the difference between the two potential outcomes: $Y_{ij}(Z_{ij} = 1) - Y_{ij}(Z_{ij} = 0)$. However, when treatments are delivered in schools and classrooms, a student's potential outcome may depend not only on the treatment assignment but also on the treatment setting (Hong, 2004; Hong & Raudenbush, 2006). Peer ability composition in a math class is an important feature of the treatment setting that may constrain instruction and may affect a student's math achievement. Moreover, peer ability itself is an immediate result of the policy. Hence student i in school j would experience $C_{ij}(Z_{ij} = 1)$ if attending the ninth grade in the post-policy year and would experience $C_{ij}(Z_{ij} = 0)$ if attending the ninth grade in the pre-policy year instead. A student's counterfactual math outcome under the algebra-for-all policy with peer ability counterfactually remaining unchanged by the policy is denoted by $Y[1, C(0)]$. The natural indirect effect is the expected change in student math outcome solely attributable to the change in peer ability induced by the policy, represented as $E\{Y[1, C(1)] - Y[1, C(0)]\}$. The natural direct effect is the expected change in math outcome associated with the policy change yet unrelated to the changes in peer ability, represented as $E\{Y[1, C(0)] - Y[0, C(0)]\}$. The natural direct effect of the policy would inform us of, for example, how the policy might have affected students' outcomes if the schools had not created mixed-ability algebra classes.

Usefulness / Applicability of Method:

This study utilizes marginal mean weighting through stratification (MMW-S) (Hong, 2010a, 2011) to adjust for between-cohort differences associated with the observed pretreatment covariates in estimating the average potential outcomes $E\{Y[1, C(1)]\}$ and $E\{Y[0, C(0)]\}$ where the weight is non-parametrically related to the propensity of policy exposure. In addition, we assess between-cohort differences in the math outcome among students attending schools that offered algebra to all ninth graders even prior to the policy and therefore were unaffected by the policy. Such differences are to be attributed to concurrent historical events associated with unobserved covariates. After matching students attending schools unaffected by the policy with those affected by the policy on observed pretreatment covariates summarized in a pair of prognostic scores (Hansen, 2008), we remove the between-cohort difference unrelated to the policy. The prognostic scores greatly reduce the dimension of covariates to be controlled for and thereby easing the matching of individual units on the potential outcome $E\{Y[0, C(0)]\}$. We then estimate two sets of propensity scores representing the conditional probabilities of experiencing a

range of peer ability levels in the pre-policy year and the post-policy year. These propensity scores are functions of the observed pretreatment and post-treatment covariates. This is followed by ratio-of-mediator-probability weighting (RMPW) applied to the post-policy cohort to obtain an estimate of the average counterfactual outcome $E\{Y[1, C(0)]\}$ (Hong, 2010b; Hong, Deutsch, Hill, 2011). This series of weighting makes it possible to estimate the total policy effect and to decompose the total effect into the natural direct effect and the natural indirect effect.

The identification assumptions include: (1) nonzero probability of policy exposure conditioning on the observed student and school pretreatment covariates; (2) no confounding of the relationship between policy exposure and math outcome conditioning on the observed student and school pretreatment covariates; (3) no confounding of the relationship between policy exposure and peer ability conditioning on the observed student and school pretreatment covariates; (4) nonzero probability of peer ability assignment under a given policy conditioning on the observed student pretreatment covariates and the observed school pretreatment and post-treatment covariates; (5) no confounding of the relationship between peer ability and math outcome under a given policy and across different policies conditioning on the observed student pretreatment covariates and the observed school pretreatment and post-treatment covariates.

Research Design:

The theoretical relationships among the algebra-for-all policy, peer ability, student math learning, and the covariates are represented in Figure 1 (please insert figure 1 here). Valid causal inference is threatened by two major sources of confounding. Observed pretreatment student characteristics $\mathbf{X}$ and school characteristics $\overline{\mathbf{X}}$ along with unobserved historical confounding $\mathbf{U}_X$ may confound the policy effect on math learning, the policy effect on peer ability, and the peer ability effect on math learning. Observed post-treatment covariates $\mathbf{W}(z)$ and unobserved post-treatment covariates $\mathbf{U}_W(z)$ may confound the peer ability effect on math learning.

Data Collection and Analysis:

Data. Student math achievement score comes from the Tests of Academic Proficiency (TAP) administered at the end of the ninth grade. Every student's latent math incoming ability is assessed on the basis of the Iowa Tests of Basic Skills (ITBS) achievement trajectory from the third to the eighth grade. This measure has been standardized across six cohorts of students with a mean of 0 and a standard deviation of 1. We use the class median math incoming ability to represent the average ability level of the peers within a math class. The entire sample in each cohort is divided into five peer ability levels denoted by $c = 0, 1, 2, 3, 4$. Table 1 shows descriptive statistics for the covariates for each cohort (please insert Table 1 here).

Weighted analysis. To remove between-cohort differences associated with observed student composition, we estimate propensity scores $\theta_{Z=z} = pr(Z = z|\mathbf{X})$ for $z = 0,1$. After dividing the analytic sample into $m = 1, \dots, M$ strata on the basis of the estimated $\theta_{Z=z}$, we compute the marginal mean weight $\omega_z = [n_m \times pr(Z = z)]/n_m^{Z=z}$, where n_m is the total number of students in stratum m; and $pr(Z = z)$ is the proportion of students enrolled under policy z. The denominator $n_m^{Z=z}$ is the number of students enrolled under policy z in stratum m. To adjust for selection of peer ability under each policy, we estimate two sets of propensity scores: one for the pre-policy year and the other for the post-policy year. We analyze an ordinal logistic regression model with students nested within schools in each cohort weighted by ω_z. To estimate the counterfactual outcome $E\{Y[1, C(0)]\}$, the RMPW for a post-policy student experiencing peer ability level c is computed as $\omega_{c0/c1} = \phi_{C(0)=c}/\phi_{C(1)=c}$. Here $\phi_{C(0)=c}$ is the

student's predicted probability of experiencing peer ability level c had the student counterfactually attended the ninth grade in the pre-policy year; $\phi_{C(1)=c}$ is the student's predicted probability of experiencing the same peer ability level c in the post-policy year. RMPW transforms the peer ability distribution of the post-policy cohort such that it approximates that of the pre-policy cohort. The outcome values of students in the post-policy cohort are adjusted on the basis of prognostic score matching. After creating a duplicate set of the post-policy cohort and combining this duplicate set with the original sample, we assign the weight as follows: to estimate $E\{Y[0, C(0)]\}$, let $\omega = \omega_z$ for pre-policy students in the original sample; to estimate $E\{Y[1, C(1)]\}$, let $\omega = \omega_z$ for post-policy students in the duplicate sample; to estimate $E\{Y[1, C(0)]\}$, let $\omega = \omega_z \times \omega_{c0/c1}$ for post-policy students in the original sample. Let D be a dummy indicator that takes value 1 for the duplicate post-policy students and 0 otherwise. A weighted two-level outcome model is simply:

$$Y_{ij} = \gamma_0 + Z\delta_1^{(NDE)} + D\delta_2^{(NIE)} + \Psi_{g=1}\gamma_1 + u_{0j} + e_{ij};\ e_{ij} \sim N(0, \sigma^2);\ u_{0j} \sim N(0, \tau).$$

Here $\delta_1^{(NDE)}$ estimates the average natural direct effect and $\delta_2^{(NIE)}$ estimates the average natural indirect effect. Hypotheses testing are based on robust standard errors.

Findings / Results:

The results suggest that, although the algebra-for-all policy did not raise lower achieving ninth graders' math achievement on average, the policy apparently improved math achievement in some schools while worsening math achievement in some other schools. We have found evidence that the average policy effect was partly mediated by peer ability change. We reason that, for lower-achieving students, a rise in peer ability in class may produce a negative effect due to unfavorable social comparisons. This negative effect may counteract the potential positive effect of participating in academic discourse involving higher-ability peers.

Conclusions:

In general, to reveal the causal mechanism of a system-wide policy is challenging not only because those who actually experienced the policy were not identical to those who did not experience the policy. Moreover, those who experienced different mediator values under a given policy tend to be systematically different. Using cohort data and a non-equivalent comparison group, this study has illustrated a series of analytic strategies that help us to overcome some of the important limitations of the existing methods. (1) By using RMPW to decompose the total effect into the natural direct effect and the natural indirect effect, our approach requires fewer identification assumptions and model-based assumptions than do most other existing approaches. (2) When the direct effect is allowed to depend on mediator values, all the existing methods require that post-treatment covariates do not confound the mediator-outcome relationship given the observed pretreatment covariates. The time series data from the CPS schools provide us with a pair of post-treatment observations of each school, one before the policy was introduced and one after. Including both observations in the propensity score models allows us to adjust for the selection of peer ability associated with the post-treatment covariates without introducing bias. (3) By matching students attending schools affected by the policy with their counterparts in schools unaffected by the policy on a pair of prognostic scores, we are able to assess and remove the impact of historical confounding in the post-policy math outcome. The prognostic score based adjustment also enables us to adjust for the vast differences between the two cohorts in student incoming math skills that the propensity score adjustment was unsuitable for.

Appendices

Appendix A. References

Hansen, B. B. (2008). The prognostic analogue of the propensity score. *Biometrika*, *95*(2), 481-488.

Holland, P. W. (1988). Causal inference, path analysis, and recursive structural equation models (with discussion). In C. C. Clogg (Ed.), *Sociological Methodology* (pp.449-493). Washington, DC: American Sociological Association.

Hong, G. (2004), "Causal Inference for Multi-Level Observational Data with Application to Kindergarten Retention," unpublished Ph.D. dissertation, University of Michigan, School of Education.

Hong, G. (2010a). Marginal mean weighting through stratification: Adjustment for selection bias in multi-level data. *Journal of Educational and Behavioral Statistics*, *35*(5), 499-531.

Hong, G. (2010b). Ratio of Mediator Probability Weighting for Estimating Natural Direct and Indirect Effects. *2010 Proceedings of the American Statistical Association*, Biometrics Section [pp.2401-2415], Alexandria, VA: American Statistical Association.

Hong, G. (2011). Marginal Mean Weighting through Stratification: A Generalized Method for Evaluating Multi-Valued and Multiple Treatments with Non-Experimental Data. To appear in *Psychological Methods*.

Hong, G., Deutsch, J., & Hill, H. (2011). Parametric and non-parametric weighting methods for estimating mediation effects: An application to the National Evaluation of Welfare-to-Work Strategies. *2011 Proceedings of the American Statistical Association*, Social Statistics Section, Alexandria, VA: American Statistical Association

Hong, G., & Raudenbush, S. W. (2006). Evaluating kindergarten retention policy: A case study of causal inference for multi-level observational data. *Journal of the American Statistical Association*, *101*(475), 901-910.

Pearl, J. (2001). Direct and indirect effects. In *Proceedings of the American Statistical Association Joint Statistical Meetings*. Minn, MN: MIRA Digital Publishing, 1572-1581, August 2005.

Pearl, J. (2010). The mediation formula: A guide to the assessment of causal pathways in non-linear models. Los Angeles, CA: University of California, Los Angeles. Technical report R-363, July 2010.

Peterson, M. L., Sinisi, S. E., & van der Laan, M. J. (2006). Estimation of direct causal effects. *Epidemiology, 17*(3), 276-284.

Robins, J. M. & Greenland, S. (1992). Identifiability and exchangeability for direct and indirect effects. *Epidemiology, 3*(2), 143-155.

Rubin, D. B. (1986). "Statistics and causal inference: Comment: Which ifs have causal answers," *Journal of the American Statistical Association, 81*(396), 961-962.

Sobel, M. E. (2008). Identification of causal parameters in randomized studies with mediating variables. *Journal of Educational and Behavioral Statistics, 33*(2), 230–251.

VanderWeele, T. & Vansteelandt, S. (2009). Conceptual issues concerning mediation, interventions and composition. *Statistics and Its Interface*, *2*, 457-468.

Appendix B. Tables and Figures

Table 1. Descriptive Statistics

	Pre-policy		Post-policy	
	Mean	SD	Mean	SD
Outcome (*Y*)				
TAP math scores	22.57	10.91	30.12	11.28
Mediator (*M*)				
Class median ability (cohort mean adjusted)	-0.68	0.34	-0.06	0.34
Class median ability (cohort mean unadjusted)	-0.61	0.37	-0.24	0.33
Post-Treatment Covariates ($W(z)$)				
Student algebra course-taking	0.42	0.49	1	0
School algebra enrollment rate	0.72	0.15	0.99	0.03
Algebra enrollment rate for disabled students	0.36	0.20	0.92	0.16
Algebra enrollment rate for low ability students without disability	0.56	0.19	1.00	0.01
% 9th grade students enrolled in advanced math	0.01	0.02	0.02	0.03
% new teachers teaching 9th grade math	0.35	0.18	0.34	0.20
# of 9th grade math teachers	12.34	4.96	11.22	4.42
% large classes	0.39	0.23	0.39	0.22
% small classes	0.18	0.16	0.27	0.13
Student Pretreatment Covariates (*X*)				
Incoming math ability	-0.94	0.29	-0.72	0.26
Male	0.46	0.50	0.50	0.50
White	0.08	0.27	0.06	0.23
Hispanic	0.34	0.47	0.38	0.49
Asian	0.01	0.11	0.01	0.10
Old for grade	0.15	0.36	0.23	0.42
Social status index	0.09	0.86	0.06	0.80
Poverty index	-0.07	0.87	-0.15	0.84
From the attendance zone	0.62	0.48	0.67	0.47
Moved once	0.26	0.44	0.28	0.45
Moved twice or more	0.13	0.34	0.13	0.33
Receiving ELL services in 8th grade	0.18	0.38	0.23	0.42
Previously received ELL services	0.19	0.39	0.20	0.40
School Pretreatment Covariates: ($\bar{X}$)				
Average incoming math skills	-0.26	0.37	-0.03	0.45
Standard deviations in incoming skills	0.73	0.12	0.85	0.15
% special education students	0.13	0.05	0.18	0.09
% White	0.09	0.12	0.08	0.12
% Asian	0.02	0.04	0.02	0.05
% Hispanics	0.25	0.30	0.27	0.30
% students who are old for grade	0.13	0.03	0.21	0.03

% students who moved once	0.27	0.07	0.27	0.07
% students who moved twice or more	0.12	0.04	0.10	0.04
Average social status index	-0.05	0.56	-0.03	0.55
Average poverty index	0.15	0.64	0.11	0.66
% students from attendance zone	0.51	0.30	0.51	0.32
% students who received ELL services in 8^{th} grade	0.11	0.14	0.12	0.14
School size in hundreds	4.78	1.76	3.86	1.56

Figure 1. Causal Model with Potential Confounders