Ratio-of-Mediator-Probability Weighting for Causal Mediation Analysis in the Presence of Treatment-by-Mediator Interaction

Guanglei Hong

Jonah Deutsch

Heather D. Hill

University of Chicago

(This is a working paper. Please do not cite or circulate without the authors' permission)

Correspondence:

Guanglei Hong

University of Chicago, Department of Comparative Human Development,

5736 S. Woodlawn Ave., Chicago, IL 60637

Email: ghong@uchicago.edu

Acknowledgement:

This research was supported by a major research grant funded by the Spencer Foundation, a Scholars Award from the William T. Grant Foundation, a US Department of Education Institute of Education Sciences "Statistical and Research Methodology in Education" grant, and start-up funds from the University of Chicago for the first author. The National Opinion Research Center (NORC) at the University of Chicago has provided research facilities.

Abstract

In many psychological studies, decomposition of the total effect into a direct effect and an indirect effect is challenging when the mediator-outcome relationship depends on the treatment condition. This article introduces an innovative ratio-of-mediator-probability weighting approach and its non-parametric counterpart that allows one to decompose the total effect in the presence of treatment-by-mediator interaction. The new weighting strategies greatly simplify model specification while minimizing reliance on assumptions with regard to the distribution of the outcome, the distribution of the mediator, and the functional form of the outcome model. Simulation results reveal satisfactory performance of the parametric and non-parametric weighting methods under the identification assumptions. We illustrate with an analysis of the impact of a welfare-to-work program on maternal depression mediated by employment experience when there is evidence that employment (the mediator) affects depression (the outcome) differently under different policy conditions (the treatment).

Keywords:

Causal inference; causal mechanism; direct effect; ignorability; indirect effect; instrumental variable; marginal structural model; path analysis; principal stratification; structural equation modeling

Causal mediation analysis is essential for testing psychological theories with regard to the mechanism through which a treatment exerts an impact on the outcome. To assess the role of a hypothesized mediator that could have been affected by the treatment and could have subsequently affected the outcome, the total effect of the treatment on the outcome is decomposed into two pieces: an "indirect effect" that channels the treatment effect through the hypothesized mediator and a "direct effect" that works through other unspecified mechanisms. Unfortunately, even in a randomized experimental study, the assignment of subjects to different mediator values is typically not randomized. Therefore, mediation analysis is challenging due to potential confounding of the mediator-outcome relationship by covariates. More challenges arise when the mediator-outcome relationship is additionally moderated by the treatment condition.

Path analysis (Duncan, 1966; Wright, 1934) has been of conventional use for mediation analysis. Most researchers in psychology have followed the analytic procedures outlined in Baron and Kenny (1986). Mediated effects are estimated through combining two or more regression models. A Sobel test has been used for determining whether there is a nonzero indirect effect (Sobel, 1982). However, treatment effect decomposition through path analysis requires a series of strong assumptions (Bullock, Green, & Ha, 2010; Holland, 1988; Sobel, 2008). In addition to linearity and additivity, path analysis requires that, given the observed covariates, the treatment assignment and the mediator value assignment are not contaminated by unobserved selection and therefore can be regarded "ignorable". Moreover, there should be no treatment-by-mediator interaction in a standard path analysis model.

Researchers in various disciplines have employed alternative analytic strategies for mediation analysis, many of which are encompassed by the general framework of structural equation modeling (Bollen, 1987; Jo, 2008; Jöreskog, 1970; MacKinnon, 2008). The validity of their results often relies on model-based assumptions in addition to their respective identification assumptions. These alternative strategies include the instrumental variable (IV) method (Heckman & Robb, 1985; Kling, Liebman & Katz, 2007; Raudenbush, Reardon, & Nomi, 2012), the principal stratification method (Elliott, Raghunathan, & Li, 2010; Frangakis & Rubin, 2002; Gallop, Small, Lin, Elliott, Joffee, & Ten Have, 2009; Page, 2012; Rubin, 2004), modified regression approaches (Pearl, 2010; Petersen, Sinisi, & van der Lann, 2006) and direct effect models (van der Lann & Petersen, 2008), marginal structural models (Coffman & Zhong, 2012; Robins, 2003; Robins & Greenland, 1992) and conditional structural models (Valeri & VanderWeele, forthcoming; VanderWeele, 2009), and a resampling approach (Imai, Keele, and Yamamoto, 2010; Imai, Keele, & Tingley, 2010).

This paper contributes to the literature by introducing parametric and nonparametric ratio-of-mediator-probability weighting (RMPW) strategies for decomposing the total effect into a direct effect and an indirect effect. The proposed method relaxes important constraining assumptions required in conventional mediation analyses that are potentially implausible in many applications. Most importantly, the RMPW methods allow for treatment-by-mediator interaction (Hong, 2010a). Applicable to binary and multi-valued mediators, these new weighting methods minimize the need for specifying the outcome model and simplify the computation of standard errors. Hence they are relatively easy to implement with standard software packages by applied researchers. Through a series of Monte Carlo simulations, we assess the performance of the parametric and non-parametric RMPW procedures and compare with that of conventional, model-based approaches such as path analysis and the instrumental variable (IV) method. The results reveal important strengths of the new weighting strategies.

We illustrate these methods with an analysis of the impact of a welfare-to-work program on maternal depression mediated by employment experience when there is evidence that employment (the mediator) affects depression (the outcome) differently under different policy conditions (the treatment). The application example is described in the next section, followed by definitions of the causal parameters, a review of the existing methods for causal mediation analysis, the theoretical rationale for using RMPW to identify the causal effects of interest, the identification assumptions, and the parametric and non-parametric weighting procedures applied to binary mediators. After presenting the simulation results, we discuss further extensions to causal mediation moderated by pretreatment characteristics and to multi-category and continuous mediators. The last section concludes and raises issues for future research.

Application Example

In the late-1990s, the US government's six decades-long welfare cash assistance program (i.e., Aid to Families with Dependent Children, AFDC) was replaced nationwide by a new program (i.e., Temporary Assistance for Needy Families, TANF). The old program entitled low-income single parents with dependent children to collect cash benefits without working. In contrast, the new program, which was designed to move low-income mothers from welfare to work, put time limits on the receipt of cash assistance and required parents to seek and secure employment as a condition for receiving assistance and for avoiding financial penalties. The political debate surrounding welfare reform focused on the potential for employment to benefit welfare recipients psychologically as well as financially (Jagannathan, Camasso, & Sambamoorthi, 2010). Low-income single mothers often experience clinical depression among other barriers to securing employment (Ahluwalia et al., 2001). In this context, successful experience in meeting the challenges of the work mandate under the new policy might possibly alleviate depressive symptoms; while the failure to find work when it is required for welfare receipt might potentially heighten depressive symptoms.

Research evidence obtained from the National Evaluation of Welfare-to-Work Strategies (NEWWS), which randomly assigned welfare recipients to work mandates, has shown that treatment group members had higher employment rates and earnings than comparison group members. We investigate whether the increase in employment induced by work mandates had an impact on maternal depression. As we reasoned above, the impact of employment on depression may depend on the policy. Hence the policy can be viewed as a *moderator*. Additionally, employment is considered to be a potential *mediator* of the policy effect on depression. We reason that depressive symptoms might possibly subside among those who became employed due to the policy change and might worsen among those who remained unemployed regardless of the policy change. When there is a treatment-by-mediator interaction, even if the average policy effect on depression is zero, questions about the mediation mechanism remains relevant.

We analyze data selected from the NEWWS experiment in Riverside, California (Hamilton, 2002). Applicants eligible for welfare were assigned at random to either a Labor Force Attachment (LFA) program representing the new policy or a control condition representing the old policy. The LFA program required participation in work or work-related activities as a condition of welfare receipt. Participants assigned to the control condition were eligible for welfare without being subject to the employment requirement. Our sample includes 208 LFA group members and 486 control group members with a child aged 3 to 5 years. Unemployment Insurance records maintained by the State of California provide quarterly administrative data on *employment* for each study participant. All participants were surveyed shortly before the randomization and again at the two-year follow-up. The self-administered questionnaire at the two-year follow-up included twelve items measuring *depressive symptoms* (e.g., I could not get going) on a frequency scale from 1 (rarely, less than 1 day during the past week) to 4 (most of the time, 5-7 days during the past week). The summary score ranged from 0 to 34 with a mean equal to 7.49 and a standard deviation equal to 7.74. The data had complete information on policy assignment and employment record.

The baseline survey provided rich information about participant background, including measures of: (a) maternal psychological well-being; (b) history of employment and welfare use; (c) human capital, employment status, earnings, and income; (d) personal attitudes toward employment, including the preference to work, willingness to accept a low-wage job, shame to be on welfare; (e) perceived social support and barriers to work; (f) practical support and barriers to work including childcare arrangement and extra family burden; (g) household composition, including number and age of children and marital status; (h) teen parenthood; (i) public housing residence and residential mobility; and (j) demographic features.

Causal Parameters

We define the person-specific and population-average causal effects of interest in terms of the counterfactual outcomes (Holland, 1986, 1988; Pearl, 2001; Robins & Greenland, 1992; Rubin, 1978). Let A denote random assignment, Z for employment experience during the two years after randomization, and Y for depressive symptoms at the two-year follow-up. Let $A = 1$ if a welfare mother was assigned to the LFA program and $A = 0$ if assigned to the control condition. For simplicity we are confining out interest to mediators measured on a binary scale, that is, $Z = 1$ if ever employed and $Z = 0$ if never employed during the two-year period, though our logic applies to multi-valued mediators as well.

Treatment effect on the mediator. We use Z_1 to denote a mother's potential employment experience if assigned to the LFA program and Z_0 for the mother's potential employment experience if assigned to the control condition. Hence the person-specific treatment effect on a mother's employment of being assigned to the LFA program versus control is $Z_1 - Z_0$.

Treatment effect on the outcome. Let Y_{1Z_1} denote a mother's potential psychological outcome if assigned to the LFA program and Y_{0Z_0} for the potential outcome if assigned to the control condition. The person-specific treatment effect on a mother's depression is $Y_{1Z_1} - Y_{0Z_0}$.

Controlled direct effects of the treatment. The treatment may take on values *a* = 0 or 1; while the mediator may also take on values *z* = 0 or 1. Hence a mother's depression level denoted by Y_{az} would be Y_{00} if the mother was assigned to the control group and unemployed, would be Y_{01} if assigned to the control group and employed, would be Y_{10} if assigned to the LFA program and unemployed, and finally, would be Y_{11} if assigned to the LFA program and employed. The controlled direct effect of the treatment if one is unemployed regardless of the treatment assignment is defined as $Y_{10} - Y_{00}$; analogously, the controlled direct effect of the treatment if one is employed regardless of the treatment assignment is defined as $Y_{11} - Y_{01}$. At the population level, the average controlled direct effect of the treatment $E(Y_{10} - Y_{00})$ represents the effect of policy change when all welfare recipients are hypothetically deprived of employment opportunities; while $E(Y_{11} - Y_{01})$ is the treatment effect if, hypothetically, all welfare recipients are offered and take jobs.

Natural direct effect of the treatment. We use Y_{1Z_0} to denote a mother's counterfactual outcome if assigned to the LFA program yet experiencing employment as she would have had under the control condition. The natural direct effect of the policy on depression is defined by $Y_{1Z_0} - Y_{0Z_0}$, representing the effect of the policy on maternal depression if the policy, perhaps counterfactually, failed to change one's employment experience. If one would be unemployed under the control condition, the natural direct effect would be equal to $Y_{10} - Y_{00}$; conversely, if one would be employed under the control condition, the natural direct effect would be equal to $Y_{11} - Y_{01}$. Because Z_0 is a random variable that can take on values 0 or 1, the population average natural direct effect is

$$E\left(Y_{1Z_0} - Y_{0Z_0}\right) = E\left(Y_{1Z_0} - Y_{0Z_0} | Z_0 = 0\right)pr(Z_0 = 0) + \left(Y_{1Z_0} - Y_{0Z_0} | Z_0 = 1\right)pr(Z_0 = 1).$$

When $Y_{10} - Y_{00}$ and $Y_{11} - Y_{01}$ are equal, that is, when there is no treatment-by-mediator interaction, the population average natural direct effect is equal to the population average controlled direct effect.

Natural indirect effect of the treatment. This is defined by $Y_{1Z_1} - Y_{1Z_0}$, representing the change in a mother's depressive symptoms attributable to the policy-induced change in her employment experience.

The total effect of treatment assignment $Y_{1Z_1} - Y_{0Z_0}$ is the sum of the natural direct effect and the natural indirect effect. The decomposition of the total effect is not unique because, alternatively, one may define the natural direct effect as $Y_{1Z_1} - Y_{0Z_1}$ and the natural indirect effect as $Y_{0Z_1} - Y_{0Z_0}$.

A Review of the Existing Methods for Mediation Analysis

In this section we compare the existing approaches to causal mediation analysis, discuss their key assumptions, and reveal their potential limitations in attempt to decompose the total policy effect on maternal depression into a direct effect and an indirect effect mediated by employment.

Path analysis. Following Baron and Kenny's (1986) procedure, one would analyze a pair of regression models with covariance adjustment for a vector of pretreatment covariates **X** that may confound the mediator-outcome relationship. The first model would regress mediator *Z* on policy assignment *A* and covariates **X**; the second model would regress outcome *Y* on mediator *Z*, policy assignment *A*, and covariates **X**. The indirect effect of the policy is then computed as the product of the estimated policy effect on the mediator and the estimated mediator effect on the outcome conditioning on the policy. The above method would not apply if, conditioning on covariates **X**, there is evidence for a nonzero policy-by-employment interaction effect on depression. Because there is evidence in the NEWWS data that employment would reduce depression under the LFA condition but not under the control condition, conventional path analysis does not have an easy solution for decomposing the total effect. In addition, when there are a relatively large number of pretreatment covariates, the outcome model and the mediator model in path analysis are prone to specification errors that may lead to bias (Drake, 1993).

Instrumental variable (IV). Widely used in econometrics, this method would choose the random assignment of participants to either LFA or the control group as the instrumental variable and involve two-stage least squares estimation. At the first stage, employment *Z* is regressed on policy assignment *A*; while at the second stage, depression *Y* is regressed on predicted employment $\hat{Z}$, the latter being a function of policy assignment *A* obtained from the first-stage analysis. The regression coefficient for predicted employment in the second model is then interpreted as the employment effect on depression or “the local average treatment effect on the compliers” (Angrist, Imbens, & Rubin, 1996). In the current application, the compliers would be defined as those who would be employed only if assigned to the LFA program and would be unemployed on if assigned to the control condition. The IV method requires the “exclusion restriction,” that is, the policy would affect depression only through changing one’s employment and hence the direct effect must be zero, an assumption that contradicts our earlier reasoning that the policy would likely heighten one's depression if one remained unemployed. Hence its usage is unwarranted when the direct effect is a key parameter of interest. Some latest development of the IV method has made possible an investigation of multiple mediators when multiple instruments are available typically in a multi-site randomized trial (Kling, Liebman & Katz, 2007; Raudenbush, Reardon, & Nomi, 2012).

Principal stratification. According to Frangakis & Rubin (2002), individual units within a principal stratum are homogeneous in how they would respond to alternative treatment assignments at the intermediate stage. In the current application, those in the same principal stratum would have the same set of potential mediator values (Z_1 and Z_0). The principal effect is the total effect of the treatment on the outcome defined within a principal stratum. Applying this framework and Bayesian estimation techniques, researchers have been able to estimate principal stratum-specific causal effects. Clearly, for units whose mediator values cannot be changed by the treatment, the principal effect is equivalent to the natural direct effect (Elliott, Raghunathan, & Li, 2010; Gallop, Small, Lin, Elliott, Joffee, & Ten Have, 2009; Page, 2012). Yet for those whose mediator values can be affected by the treatment, decomposing the total effect into a natural direct effect and a natural indirect effect requires the strong assumption of a constant direct effect (VanderWeele, 2011, 2012).

Modified regression approaches and direct effects models. Petersen, Sinisi, and van der Laan (2006) modified the path analysis approach to estimating natural direct effects by directly incorporating an interaction between treatment and mediator in the outcome model. They estimate the average direct effect as $\hat{E}\left(Y_{1Z_0} - Y_{0Z_0}\right) = \beta_1 + \beta_2\hat{E}(\mathbf{X}) + \beta_3\hat{E}(Z_0) + \beta_4\hat{E}(Z_0\mathbf{X})$. The analytic procedure involves (1) analyzing a multiple regression of the outcome on the mediator, treatment, pretreatment covariates, and their interactions and obtain the corresponding regression coefficient estimates, and (2) analyzing a multiple regression of the mediator on the treatment, pretreatment covariates, and their interactions. After obtaining the predicted Z_0 as a function of $\mathbf{X}$ in the control group, by taking the sample average of the predicted Z_0 over the distribution of $\mathbf{X}$, one obtains an estimate of $E(Z_0)$. By taking the sample average of the product of the predicted Z_0 and $\mathbf{X}$, one obtains an estimate of $E(Z_0\mathbf{X})$. Similar to the standard approach of path analysis, this modified regression approach requires linearity of the outcome model and of the mediator model. The functional forms of these models may have direct consequences for identification. Moreover, standard error computation becomes cumbersome for the estimate of a causal effect represented as a function of multiple parameters. More recently, van der Laan and Petersen (2008) outlined a series of models in which the natural direct effect is defined as a function of the underlying parameters and, in each case, can be directly estimated through solving a corresponding estimating function. These models require a user-supplied conditional distribution of the mediator representing the data generating function under the control condition. Also required are specifications of the outcome model as a function of the treatment, the mediator, and the covariates. The distributional assumptions and the modeling assumptions are necessary for obtaining estimators with good practical performance.

Marginal structural models and conditional structural models. Marginal structural models influential in epidemiology apply an inverse-probability-of-treatment weighting (IPTW) strategy to remove selection bias (Coffman & Zhong, 2012; Robins, 2003; Robins & Greenland, 1992). For estimating the natural direct and indirect effects, this approach becomes inapplicable when there is treatment-by-mediator interaction. VanderWeele (2009) proposed a solution that employs two conditional structural models similar to those specified by Petersen et al (2006) in the above modified regression approach. Assuming that the outcome is a linear function of the mediator given the treatment and the pretreatment covariates, one can obtain an estimate of the counterfactual outcome $E(Y_{1Z_0}|\mathbf{X}=\mathbf{x})$ by replacing the mediator value in the outcome model with the predicted value of Z_0 given the pretreatment covariates. According to VanderWeele (2009), the conditional structural models are unique in that they are conditional on a subset of pretreatment covariates that confound the mediator-outcome relationship across the treatment conditions (Pearl, 2001). In the NEWWS example, these are covariates associated with employment experience under the control condition and with depressive symptoms under the experimental condition. In the meantime, IPTW is used to adjust for the confounding effects of the entire set of pretreatment covariates. However, Robins (2003) argued that it is unlikely that one would accept the assumption of no mediator-outcome confounding across the treatment conditions after conditioning on the pretreatment covariates unless one believed that the mediator value assignment was randomized by nature within levels of the pretreatment covariates. In that sense, the set of covariates to be adjusted for removing mediator-outcome confounding across the treatment conditions is the same as the set of covariates to be adjusted for removing mediator-outcome confounding within each treatment condition. As the number of such covariates becomes large, the issue of model specification again raises concern.

Resampling approach. Instead of relying on a user-supplied conditional distribution of Z_0 as required by the direct effect models, Imai and colleagues (Imai, Keele, and Yamamoto, 2010; Imai, Keele, & Tingley, 2010) employed a simulation strategy to empirically generate such a distribution. They developed a computationally intensive algorithm that requires fitting a mediator model and an outcome model followed by repeatedly resampling the potential values of the mediator Z_0 and the potential outcomes given the simulated values of the mediator Y_{1Z_0}. The analysis nonetheless depends on correct specifications of both the outcome and the mediator models.

Within the general SEM framework, the analytic strategies discussed above share a number of considerations. To obtain a point estimator of the natural indirect effect, most of these strategies require either computing the product of a parameter estimate obtained from the outcome model and one from the mediator model or computing the difference between the total effect estimate and the natural direct effect estimate. Extra programming is often required for obtaining the asymptotic standard error and the confidence interval of the natural indirect effect estimate. In addition, the product approach requires a linear mediator-outcome relationship. All of these approaches are prone to specification errors in the functional forms of the outcome model, the mediator model, or both.

Ratio-of-Mediator-Probability Weighting: Intuitive Rationale

To decompose the total effect into a natural direct effect and a natural indirect effect, our primary goal is to obtain consistent estimates of three population average potential outcomes: the average potential outcome associated with the control condition $E\left(Y_{0Z_0}\right)$, the average potential outcome associated with the experimental condition $E\left(Y_{1Z_1}\right)$, and the average potential outcome associated with the experimental condition when the mediator counterfactually takes on values associated with the control condition $E\left(Y_{1Z_0}\right)$. In this section, we use a simple hypothetical example to clarify the logic of total effect decomposition in the context of a sequential randomized experiment. We then develop a strategy to approximate the sequential experimental design.[1]

Suppose that, after welfare recipients had been assigned at random either to the LFA program $(A = 1)$ with a 0.5 probability or to the control condition $(A = 0)$ with the same probability, those in the LFA program would subsequently be assigned at random to be employed with probability .7 or to be unemployed with probability .3, while those in the control group would be assigned at random to be employed with probability .4 or to be unemployed with probability .6.

In this hypothetical experiment, the LFA group and the control group are expected to have the same pretreatment composition. Hence the average observed outcome of the LFA group estimates $E(Y_{1Z_1})$ while that of the control group estimates $E(Y_{0Z_0})$. Under sequential randomization, the employed in the LFA group are expected to have the same pretreatment composition as the employed in the control group, the unemployed in the LFA group, and the unemployed in the control group. To estimate $E(Y_{1Z_0})$, which can be written as $E(Y_{11}) \times pr(Z_0 = 1) + E(Y_{10}) \times pr(Z_0 = 0)$, we may simply transform the employment rate in the LFA group to resemble that in the control group through weighting. Each employed individual in the LFA group may receive a weight that is equal to $pr(Z_0 = 1)/pr(Z_1 = 1) = 0.4/0.7 = 4/7$, while each unemployed individual in the LFA group may receive a weight that is equal to $[1 - pr(Z_0 = 1)]/[1 - pr(Z_1 = 1)] = (1 - 0.4)/(1 - 0.7) = 2$. The weighted LFA group would then show an employment rate 0.4 equivalent to that in the control group. The average weighted outcome of the LFA group therefore estimates $E(Y_{1Z_0})$. In general, for an LFA participant displaying mediator value z for $z = 0,1$, the weight is a ratio of the individual's probability of having mediator value z under the control condition to that under the LFA condition such that

$$E(Y_{1z}) \times pr(Z_0 = z) = E\left(\frac{pr(Z_0 = z)}{pr(Z_1 = z)} \times Y_{1z}\right) \times pr(Z_1 = z).$$

The ratio-of-mediator-probability weighting (RMPW) strategy is more broadly applicable than the conventional path analysis method. Most importantly, the RMPW strategy does not require the assumption that there is no treatment-by-mediator interaction. In contrast, when this assumption is violated, path analysis will generate biased estimates of the natural direct effect and the natural indirect effect. For the natural indirect effect estimate, the bias is

$$-E\{(Y_{11}-Y_{01})-(Y_{10}-Y_{00})\}\times\{pr(Z_1=1)-pr(Z_0=1)\}\times pr(A=0).$$

Appendix A derives the bias.

The above sequential randomized experiment is likely infeasible because some individuals might be already employed at the time of randomization and might be highly motivated to continue working while others might have to overcome multiple barriers in seeking employment. In theory, an individual's probability of employment under a given policy condition can be predicted by a set of important motivators and inhibitors. A relatively realistic experiment would involve identifying subgroups of participants who were homogeneous in the predicted probability of employment under a given policy and assigning individuals to be employed according to their predicted probabilities (Hong, 2006). For individuals with pretreatment characteristics $\mathbf{x}$, let the probability of being employed if assigned to the LFA program be $pr(Z_1=1|\mathbf{X}=\mathbf{x})$ and the probability of being employed if assigned to the control condition be $pr(Z_0=1|\mathbf{X}=\mathbf{x})$. This would be equivalent to a sequential randomized experiment within levels of $\mathbf{X}=\mathbf{x}$. To estimate $E(Y_{1Z_0})$, again we use weighting to transform the employment rate in the LFA group. The weight is $pr(Z_0=1|\mathbf{X}=\mathbf{x})/pr(Z_1=1|\mathbf{X}=\mathbf{x})$ for the employed and $[1-pr(Z_0=1|\mathbf{X}=\mathbf{x})]/[1-pr(Z_1=1|\mathbf{X}=\mathbf{x})]$ for the unemployed in the LFA group.

In the current study, only the policy assignment was random. Yet by taking advantage of the rich pretreatment information, we may nonetheless approximate a sequential randomized experiment within levels of the observed pretreatment characteristics. Specifically, we may predict each LFA participant's probability of being employed under the LFA condition $pr(Z_1 = 1|A = 1, \mathbf{X} = \mathbf{x})$ and that under the control condition $pr(Z_0 = 1|A = 0, \mathbf{X} = \mathbf{x})$ through analyzing two propensity score models for employment under the two respective policy conditions. By virtue of the random policy assignment, the propensity score model specified under the control condition would apply to the LFA participants had they been counterfactually assigned to the control condition instead. Hence applying the weight $pr(Z_0 = 1|A = 0, \mathbf{X} = \mathbf{x})/pr(Z_1 = 1|A = 1, \mathbf{X} = \mathbf{x})$ to the employed and $[1 - pr(Z_0 = 1|A = 0, \mathbf{X} = \mathbf{x})]/[1 - pr(Z_1 = 1|A = 1, \mathbf{X} = \mathbf{x})]$ to the unemployed in the LFA group, we obtain an estimate of $E(Y_{1Z_0})$.

Theoretical Results and Identification Assumptions

This section derives the theoretical results and clarifies the identification assumptions under which the RMPW strategy removes selection bias in estimating the natural direct effect and the natural indirect effect.

Following van der Laan and Petersen (2008), we represent the joint distribution of the observed data $O = (\mathbf{X}, A, Z_A, Y_{AZ_A})$ in general as follows:

$$f^{(a,z)}(Y_{az}|A = a, Z_a = z, \mathbf{X}) \times q^{(a)}(Z_a = z|A = a, \mathbf{X}) \times p(A = a|\mathbf{X}) \times h(\mathbf{X}),$$

where $f^{(a,z)}(\cdot)$, $q^{(a)}(\cdot)$, $p(\cdot)$, and $h(\cdot)$ are probability or density functions. For simplicity, we use $f(\cdot)$ to represent $f^{(a,z)}(\cdot)$ henceforth. The theorem below presents the general result for identifying the population average counterfactual outcome $E(Y_{aZ_{a'}})$ associated with treatment a when the mediator counterfactually takes on values associated with an alternative treatment a'.

Using observed outcome data from units in treatment a, we estimate $E\left(Y_{aZ_{a'}}\right)$ through weighting where $W_{(aZ_{a'})}$ denotes the weight assigned to the units in treatment a (Hong, 2010a).

THEOREM. $E(Y^*|A=a) \equiv E\left(W_{(aZ_{a'})}Y|A=a\right)$ is an observed data estimand for $E\left(Y_{aZ_{a'}}\right)$, where

$$W_{(aZ_{a'})} = \frac{q^{(a')}(Z_{a'}=z|A=a',\mathbf{X})}{q^{(a)}(Z_a=z|A=a,\mathbf{X})} \times \frac{p(A=a)}{p(A=a|\mathbf{X})} \qquad (1)$$

for all possible values of a and z.

We prove in Appendix B that the above results hold under the following assumptions:

Assumption 1 (Nonzero probability of treatment assignment). $0 < p(A=a \mid \mathbf{X}) < 1$.

Assumption 2 (Independence of treatment assignment and potential outcomes). $Y_{aZ_a}, Y_{aZ_{a'}} \amalg A \mid \mathbf{X}$.

Assumption 3 (Nonzero probability of mediator value assignment). $0 < p(Z_a = z|A = a\,, \mathbf{X}) < 1$.[2]

Assumption 4 (No confounding of mediator-outcome relationship within the treatment assigned). $Y_{az} \amalg Z_a \mid A = a, \mathbf{X}$.

Assumption 5 (No confounding of treatment-mediator relationship). $Z_a \amalg A \mid \mathbf{X}$.

Assumption 6 (No confounding of mediator-outcome relationship across treatment conditions). $Y_{az} \amalg Z_{a'} \mid A = a, \mathbf{X}$.

Under Assumptions 1 and 2, the sample mean outcome of the control group members provides an unbiased estimate of $E\left(Y_{0Z_0}\right)$ while the sample mean outcome of the LFA participants provides an unbiased estimate of $E\left(Y_{1Z_1}\right)$. Additionally under Assumptions 3, 4, 5, and 6, the RMPW adjusted sample mean outcome of the LFA participants provides an unbiased estimate of $E\left(Y_{1Z_0}\right)$.

In the NEWWS study, because of the randomized policy assignment, each unit had a nonzero probability of being assigned to either the LFA program or the control group. The policy assignment was independent of all the potential outcome values and the potential mediator values. Hence Assumptions 1, 2, and 5 are satisfied. Assumption 4 requires that the observed covariates adequately account for all the potential confounding of the mediator-outcome relationship under each treatment condition. Assumption 6 additionally requires that, within levels of the observed pretreatment covariates, the mediator value assignment under the control condition is independent of the potential outcome under the LFA condition. Pearl (2001) interpreted this as an assumption requiring no post-treatment covariates given the observed pretreatment covariates. The above six assumptions constitute “sequential ignorability” (Imai, Keele, and Yamamoto, 2010; Imai, Keele, & Tingley, 2010), that is, the treatment assignment and the mediator value assignment under each treatment are ignorable within levels of the pretreatment covariates. Unlike many of the existing methods, the RMPW strategy does not require the no treatment-by-mediator interaction assumption. Hence Assumptions 1~6 are adequate for identifying the natural effects.

Here we discuss the implications of Assumption 3 in the NEWWS application. About 35% of the welfare recipients assigned to the LFA program were unemployed during the two years after randomization while as many as 40% of those assigned to the control group had employment during the same time period. Under the “principal stratification” framework (Angrist, Imbens, & Rubin, 1996; Frangakis & Rubin, 2002; Rubin, 2004), the former would be regarded as “never-takers” and the latter “always-takers.” This deterministic view of potential mediator values may not apply well in a job market that is often full of uncertainty. Assumption 3 represents a probabilistic view of mediator value assignment. That is, many of those who were unemployed under the LFA condition may have had a nonzero probability of being employed; similarly, many of those who were employed under the control condition may have had a nonzero probability of becoming unemployed.

RMPW Application to NEWWS Data

Applying the above theoretical results to an analysis of the NEWWS data, we illustrate the parametric and nonparametric analytic procedures for a binary mediator and report the results in this section. According to the results of intent-to-treat analysis, assignment to the LFA program (that is, $A = 1$ rather than 0) increased the employment rate from 39.5% to 65.4%. However, the average policy effect on maternal depression cannot be statistically distinguished from zero (coefficient = 0.11, SE = 0.64, t = 0.18, p = 0.86). We have measures of 86 pretreatment covariates that are theoretically associated with maternal depression or with employment experience. After creating a missing category for each categorical covariate with missing information and then assuming missing at random, we conduct maximum likelihood-based imputation in the outcome and the continuous covariates (Little & Rubin, 2002).

Applying the marginal structure model approach non-parametrically to adjusting for the observed pretreatment covariates (Hong, 2010, 2012a), we regress the outcome on the binary treatment, the binary mediator, and their interaction. The estimated controlled direct effect of employment under LFA (coefficient = -2.49, SE = 1.20, t = -2.07, $p < 0.05$) and that under the control condition (coefficient = 0.74, SE = 0.76, t = 0.97, p = 0.33) are significantly different (coefficient = 3.23, SE = 1.42, t = 2.27, $p < 0.05$). According to these results, had all welfare mothers continued to be covered by the old policy, employment would not have affected maternal depression by a significant amount. However, once employment became one of the primary qualifications for welfare receipt, employment success apparently would lead to a reduction in depressive symptoms. The above result serves as empirical evidence for the existence of policy-by-employment interaction in the NEWWS data.

Below we discuss how to decompose the total effect into a natural direct effect and a natural indirect effect. For a binary mediator, we first present a parametric approach that estimates RMPW directly as a ratio of the estimated conditional probability of mediator value assignment under the control condition to that under the LFA condition. We then provide a non-parametric alternative approach in which each conditional probability is computed on the basis of propensity score stratification.

Mediator on a Binary Scale: Parametric Approach

When the policy assignment was randomized and with a binary mediator, Equation (1) takes the following form:

$$W_{(1Z_0)} = \frac{q^{(0)}(Z_0 = z|A = 0, \mathbf{X})}{q^{(1)}(Z_1 = z|A = 1, \mathbf{X})} \text{ for } z = 0,1. \tag{2}$$

We predict each individual's conditional probability of being employed if assigned to the control condition and the conditional probability of being employed if assigned to the LFA program each as a function of $\mathbf{x}$ denoted by $\theta_{Z_0}(\mathbf{x})$ and $\theta_{Z_1}(\mathbf{x})$, respectively.

To proceed, we reconstruct the data set to include the sampled control group members, the sampled LFA members, and a duplicate set of the sampled LFA members. Let D be a dummy indicator that takes value 1 for the duplicate LFA members and 0 otherwise. We assign the weight as follows to units in the analytic sample:

If $A = 0$, and $D = 0$, then $RMPW = 1.0$;

If $A = 1$, and $D = 1$, then $RMPW = 1.0$;

If $A = 1, D = 0,$ and $Z = 1$, then $RMPW = \theta_{Z_0}/\theta_{Z_1}$;

If $A = 1, D = 0,$ and $Z = 0$, then $RMPW = (1 - \theta_{Z_0})/(1 - \theta_{Z_1})$.

In the RMPW adjusted LFA group, the proportion of employment is 0.376, similar to that in the control group in which 39.5% of the individuals were employed. We then analyze an outcome model through ordinary least squares and obtain either robust cluster standard errors or bootstrapped standard errors in Stata.[3]

$$Y = \gamma^{(0)} + \gamma^{(NDE)}A + \gamma^{(NIE)}D + e. \quad (3)$$

Here $\gamma^{(0)}$ estimates $E(Y_{0Z_0})$; $\gamma^{(0)} + \gamma^{(NDE)}$ estimates $E(Y_{1Z_0})$; and $\gamma^{(0)} + \gamma^{(NDE)} + \gamma^{(NIE)}$ estimates $E(Y_{1Z_1})$. Hence we obtain $\gamma^{(NDE)}$ as an estimator of the natural direct effect $E(Y_{1Z_0} - Y_{0Z_0})$ and $\gamma^{(NIE)}$ as an estimator of the natural indirect effect $E(Y_{1Z_1} - Y_{1Z_0})$.

The estimated natural direct effect is 1.02 (robust *SE* = 0.93; *Wald* χ^2 = 1.20, *p* = 0.27), about 14% of a standard deviation of the outcome; the estimated natural indirect effect is -0.70 (robust *SE* = 1.04; *Wald* χ^2 = 0.46, *p* = 0.50). The natural direct effect estimate indicates that, if the employment experience of all welfare mothers would have counterfactually remained unchanged by the welfare-to-work policy, maternal depression would have increased only by an insignificant amount on average. The natural indirect effect estimate indicates that apparently the policy-induced change in employment was not great enough to produce a significant reduction in maternal depression on average.

Mediator on a Binary Scale: Non-Parametric Approach

After ranking the units in the analytic sample by θ_{Z_1}, that is, the conditional probability of being employed if assigned to the LFA program as a function of pretreatment covariates, we divide the sample into three even portions. Within each of these three subclasses, we then rank and subdivide again by θ_{Z_0}. Let $S = 1, \ldots, 9$ denote the nine strata. We generate a duplicate set of the LFA group as before and use $D = 1$ to denote the duplicates and 0 otherwise. The non-parametric weight is computed as follows:

If $A = 0$, and $D = 0$, then $NRMPW = 1.0$;

If $A = 1$, and $D = 1$, then $NRMPW = 1.0$;

If $A = 1, D = 0,$ and $Z = z$, then

$$NRMPW = \frac{\sum_{i \in C}(1 - A_i) I_i(Z_0 = z) I_i(S = s)}{\sum_{i \in C}(1 - A_i) I_i(S = s)} \times \frac{\sum_{i \in E} A_i I_i(S = s)}{\sum_{i \in E} A_i I_i(Z_1 = z) I_i(S = s)}.$$

Here $i \in C$ represents an individual unit in the control group, and $i \in E$ is an individual unit in the LFA group. Let the indicator $I_i(S = s)$ takes on value 1 if participant i is in stratum s and 0 otherwise. Let $I_i(Z_a = z)$ for $a = 0{,}1$ be an indicator that takes on value 1 when the mediator value of participant i is z. Hence the first element of the weight above computes the proportion of control group units in stratum s who displayed mediator value z under the control condition, while the second element is the inverse of the proportion of LFA group units in stratum s who displayed mediator value z under the LFA condition.

Applying the above non-parametric weight, we analyze Equation (3) again through ordinary least squares and obtain either robust cluster standard errors or bootstrapped standard errors. The estimated natural direct effect is 1.07 (robust *SE* = 0.88, *Wald* $\chi^2 = 1.50$, $p = 0.22$); the estimated natural indirect effect is -0.76 (robust *SE* = 0.99, *Wald* $\chi^2 = 0.58$, $p = 0.45$). To increase bias reduction and improve precision, we additionally control for the fixed effects of the nine cells in analyzing Equation (5). The estimated natural direct effect becomes 0.96 (robust *SE* = 0.85, *Wald* $\chi^2 = 1.25$, $p = 0.26$); the estimated natural indirect effect is still -0.76 (robust *SE* = 0.97, *Wald* $\chi^2 = 0.61$, $p = 0.44$). As before, we conclude that the policy-induced change in employment appears to play a minimal role in mediating the policy impact on maternal depression. In general, the point estimates of non-parametric weighting are similar to those of parametric weighting. However, non-parametric weighting appears to be slightly more efficient than parametric weighting, a feature that we will investigate further through simulations.

Simulations

Comparisons between Parametric and Non-Parametric RMPW Strategies

We conduct a series of Monte Carlo simulations to assess the performance of the non-parametric weighting procedures relative to the parametric weighting procedures in the case of a binary randomized treatment indicator, a binary mediator, and a continuous outcome. The simulated data resemble the structure of the NEWWS Riverside data. We select two different sample sizes: N = 800 represents a relatively small sample size similar to the NEWWS Riverside data; N = 5,000 represents a large sample size seen in some national evaluations. With non-parametric weighting, we also compare three-by-three strata with four-by-four strata in estimating the natural direct and indirect effects. For each given sample size, we generate 1,000 random samples.

In our baseline model, potential outcomes Y_{az} for $a = 0{,}1$ and $z = 0{,}1$ are each a linear additive function of three standard normal independent covariates X_1, X_2, and X_3. Let the logit of propensity for employment under each treatment be a linear additive function of these same covariates. The controlled direct effects are determined by three parameters $\delta^{(A)}$, $\delta^{(Z)}$, and $\delta^{(AZ)}$. Specifically, we have that $E(Y_{10} - Y_{00}) = \delta^{(A)}$, $E(Y_{01} - Y_{00}) = \delta^{(Z)}$, $E(Y_{11} - Y_{01}) = \delta^{(A)} + \delta^{(AZ)}$ and $E(Y_{11} - Y_{10}) = \delta^{(Z)} + \delta^{(AZ)}$. The above three parameters together with $E\left(\theta_{Z_0}\right)$ and $E\left(\theta_{Z_1}\right)$ determine the natural direct effect and the natural indirect effect. Specifically, $E\left(Y_{1Z_0} - Y_{0Z_0}\right) = \gamma^{(NDE)} = \delta_A + \delta_{AZ} E\left(\theta_{Z_0}\right)$, and $E\left(Y_{1Z_1} - Y_{1Z_0}\right) = \gamma^{(NIE)} = (\delta_Z + \delta_{AZ})\left[E\left(\theta_{Z_1}\right) - E\left(\theta_{Z_0}\right)\right]$.

We compare across three sets of parameter value specifications shown in Table 1. In simulation (a), the natural direct effect and the natural indirect effect are both set to be zero. Simulations (a) and (b) approximate the employment rate in the LFA program and that under the control condition similar to what have been observed in the NEWWS data. Simulation (c) increased the employment rate in the LFA program and decreased that under the control condition.

The evaluation criteria for causal effect estimate $\hat{\lambda}$ include the following: (1) bias in the point estimate: $E(\hat{\lambda}) - \lambda$; (2) sampling variability of the point estimate: $\sigma^2(\hat{\lambda}) = E[\hat{\lambda} - E(\hat{\lambda})]^2$; (3) mean square error: $E\left[(\hat{\lambda} - \lambda)^2\right] = \sigma^2(\hat{\lambda}) + [E(\hat{\lambda}) - \lambda]^2$; and (4) approximate bias in the standard error estimate: $E\left[\widehat{\sigma(\hat{\lambda})}\right] - \sigma(\hat{\lambda})$.

Results from a naïve analysis serve as the baseline for assessing the performance of the weighting procedures. To estimate the natural direct and indirect effects, the naïve procedure is similar to the RMPW procedure except that the weight is equal to $pr(Z = z|A = 0)/pr(Z = z|A = 1)$ for estimating $E(Y_{1Z_0})$.

Table 2 summarizes the key results corresponding to the three sets of baseline parameter values. Both parametric weighting and non-parametric weighting perform generally well in all three cases. RMPW removes nearly 100% of the bias; NRMPW with three-by-three strata removes 85% or more of the initial bias. Yet NRMPW estimates consistently show a higher efficiency and often display a smaller MSE than do RMPW estimates. As the number of strata increases, the percentage of bias removal tends to increase when the sample size is relatively large. For example, NRMPW with four-by-four strata removes 90% or more of the initial bias. However, the advantage of increasing strata apparently disappears in small samples when $E(\theta_{Z_0})$

and $E(\theta_{Z_1})$ are shifting away from 0.5. Besides, an increase in the number of strata sometimes results in a loss in efficiency. The weighted results do not display a considerable amount of bias in the estimation of standard errors. Comparing the standard error estimates with their corresponding sampling standard deviations approximated on the basis of 1,000 samples, we found the discrepancy close to zero across all the parameter estimates. For any single estimate, the discrepancy never exceeds 0.047 standard deviations of a potential outcome. Details of the simulation results are displayed in Tables S1, S2, and S3 in the online supplement.

Comparisons with Conventional Approaches to Mediation Analysis

Comparisons with path analysis. To evaluate the performance of the weighting methods relative to path analysis, we analyze naïve path analysis models with no statistical adjustment for any covariates and with no treatment-by-mediator interaction. The naïve path analysis results now serve as the basis for comparing across the adjustment methods. Adjusted path analysis makes linear covariance adjustment for all three covariates yet still excludes treatment-by-mediator interaction. In analyzing baseline model (a) in which the treatment-by-mediator interaction is zero, the parametric RMPW results replicate the adjusted path analysis results showing nearly 100% bias removal in all causal effect estimates. The NRMPW method with three-by-three strata removes nearly 85% of the bias in estimating the natural direct and indirect effects. However, the weighting methods are not as efficient as the adjusted path analysis method. In analyzing baseline models (b) and (c) where a nonzero treatment-by-mediator interaction exists, the weighting methods greatly outperform the adjusted path analysis in terms of bias reduction and mean square error despite the loss in efficiency. The same pattern holds regardless of sample size. Table S4 in the online supplement summarizes the results.

Comparisons with the IV method. To compare the weighting methods with the IV method, we assume that the exclusion restriction holds and therefore let $\delta_A = \delta_{AZ} = 0$ while keeping all other parameter values the same as before. The natural direct effect is now zero; the natural indirect effect equals $\delta_Z[E(\theta_{Z_1}) - E(\theta_{Z_0})]$. The parameter value specifications are modified as follows:

(a) $E(\theta_{Z_0}) = .3918$, $E(\theta_{Z_1}) = .6609$, $\delta^{(A)} = \delta^{(Z)} = \delta^{(AZ)} = 0$, and hence $\gamma^{(NDE)} = \gamma^{(NIE)} = 0$;

(b*) $E(\theta_{Z_0}) = .3918$, $E(\theta_{Z_1}) = .6609$, $\delta^{(A)} = 0$, $\delta^{(Z)} = 1.5$, $\delta^{(AZ)} = 0$, hence $\gamma^{(NDE)} = 0$ and $\gamma^{(NIE)} = 0.40365$;

(c*) $E(\theta_{Z_0}) = .1938$, $E(\theta_{Z_1}) = .8066$, $\delta^{(A)} = 0$, $\delta^{(Z)} = 2.5$, $\delta^{(AZ)} = 0$, hence $\gamma^{(NDE)} = 0$ and $\gamma^{(NIE)} = 1.532$.

Specifically, we compare the RMPW and three-by-three-stratum NRMPW estimates of the natural indirect effect with the ITT effect estimates with covariance adjustment for all the covariates. Under the exclusion restriction and the assumption that the treatment effect on the mediator is independent of the mediator effect on the outcome (Raudenbush, Reardon, & Nomi, 2012), IV estimates of the natural indirect effect is expected to be equivalent to the ITT effect. In all cases, we find that the RMPW estimates are comparable to the ITT estimates. The non-parametric weighting results show a slightly larger bias as expected. Table S5 in the online supplement shows the results.

Extensions of the RMPW Method

In this section we discuss extensions of the RMPW method to an analysis of causal mediation mechanism that may vary across subpopulations of units, to mediation analysis with multi-category or continuous mediators, and to non-experimental studies. Although challenges may arise to the non-parametric RMPW method as the mediator takes an increasing number of values, we will show that the parametric RMPW method is flexible for handling multi-valued mediators.

RMPW Procedure for Analyzing Moderated Mediation

We have found in the NEWWS data that welfare recipients who had been teen parents in the past appeared to respond differently to the policy than those who had never become teen parents. Non-teen mothers assigned to the LFA program displayed heightened depressive symptoms two years after randomization in comparison with their counterparts assigned to the control group. In contrast, although teen mothers started with a higher average level of depression at the baseline, two years later, assignment to the LFA program did not seem to increase their depressive symptoms more than assignment to the control group did. It is also noteworthy that, while non-teen mothers assigned to the LFA program showed a higher employment rate (63%) than those assigned to the control group (41%), an even larger difference in employment rate emerged among teen mothers between the LFA group (70%) and the control group (37%). We suspect that many teen mothers may have accrued important life skills and secured some social support over the years as they endured the hardship associated with teen parenthood. Such skills and social support may enable them to adjust to the employment requirement under the new policy. Yet as shown in earlier research, job-related stress is sometimes comparable to unemployment-related stress for single mothers with young children. Hence we ask whether the policy-induced increase in employment contributed to teen mothers' psychological well-being during the two years of program participation. We also ask the extent to which the development of depressive symptoms among non-teen mothers was attributable to the lack of adequate increase in employment during the same time period.

Let V be a dummy indicator for teen parent status, that is, $V = 1$ for a teen parent and $V = 0$ for a non-teen parent. To answer the above research questions, researchers conventionally conduct multi-group comparisons in structural equation modeling. However, this is no longer an option when there is policy-by-employment interaction in either or both subpopulations. Here we extend the RMPW procedure to investigate teen parenthood as a potential moderator in the mediation analysis. We conduct the same set of analysis as described earlier within each subpopulation. Equation (3) is now modified to include two sub-models, one for teen mothers and the other for non-teen mothers:

$$Y = V\left(\gamma_{V1}^{(0)} + \gamma_{V1}^{(ND)}A + \gamma_{V1}^{(NI)}D\right) + (1 - V)\left(\gamma_{V0}^{(0)} + \gamma_{V0}^{(ND)}A + \gamma_{V0}^{(NI)}D\right) + e.$$

Here $\gamma_{V1}^{(ND)}$ and $\gamma_{V1}^{(NI)}$ estimate the natural direct effect and the natural indirect effect, respectively, for teen mothers; $\gamma_{V0}^{(ND)}$ and $\gamma_{V0}^{(NI)}$ estimate the natural direct effect and the natural indirect effect, respectively, for non-teen mothers.

We find that, for teen mothers, the natural direct effect is -0.53 (SE = 1.23, *Wald* χ^2 = 0.18, p = .67) and the natural indirect effect is -0.42 (SE = 1.35, *Wald* χ^2 = 0.10, p = .76). There is conceivably a ceiling effect in the extent to which the new policy could elevate teen mothers' depression level should their employment level remain unchanged. In the meantime, despite the considerable amount of policy-induced increase in employment among teen mothers, the success did not seem to reduce their depressive symptoms. In contrast, for non-teen mothers, the natural direct effect is 1.90 (SE = 1.33, *Wald* χ^2 = 2.04, p = .15) and the natural indirect effect is -0.81 (SE = 1.47, *Wald* χ^2 = 0.30, p = .58). The amount of increase in depressive symptoms is noteworthy (Effect Size = 0.25) should non-teen mothers be subject to the new policy requirement without improvement in their employment rate. However, due to the limited statistical power in the data, the estimated differences between teen mothers and non-teen mothers cannot be statistically distinguished from zero.

RMPW Procedure for Multi-Category Mediators

The binary measure of employment does not distinguish among people who were employed to varying degrees. To overcome this limitation, we examine a three-category measure of employment: Never employed (z = 0), low employment (z = 1) (i.e., employed for no more than 50% of the two-year period), and high employment (z = 2) (i.e., employed for more than 50% of the two-year period). In the NEWWS data, the proportion of participants with low employment was larger in the LFA group (34.6%) than in the control group (23.3%). An even larger difference in high employment lies between the LFA group (30.8%) and the control group (16.3%).

We define the natural direct effect and the natural indirect effect of policy assignment on maternal depression the same as before. Below we highlight the modifications in the analytic procedure for accommodating a multi-valued mediator. Each individual now has three propensity scores under each treatment condition. Specifically, $\theta_{Z_1=0}$, $\theta_{Z_1=1}$, and $\theta_{Z_1=2}$ represent the conditional probability of having zero employment, low employment, and high employment under LFA; $\theta_{Z_0=0}$, $\theta_{Z_0=1}$, and $\theta_{Z_0=2}$ represent the conditional probabilities of having these three levels of employment under the control condition. A comparison between a multinomial logistic regression model and an ordinal logistical regression model shows that, in this case, the latter fits the data as adequately as the former.

As before, we reconstruct the data set to include the sampled control group members, the sampled LFA members, and a duplicate set of the sampled LFA members. Let D be a dummy indicator that takes value 1 for the duplicate LFA members and 0 otherwise. Applying Equation (3) to the analytic sample, the parametric RMPW can be computed as follows:

If $A = 0$, and $D = 0$, then $RMPW = 1.0$;

If $A = 1$, and $D = 1$, then $RMPW = 1.0$;

If $A = 1, D = 0,$ and $Z_1 = z$, then $RMPW = \theta_{Z_0=z}/\theta_{Z_1=z}$. (4)

Regardless of the distribution of the multi-category mediator, the outcome model is specified the same as that in model (3) and is analyzed through generalized least squares with robust standard errors. The estimated natural direct effect is 1.19 (*SE* = 0.90, *Wald* χ^2 = 1.76, *p* = 0.19); the estimated natural indirect effect is -.62 (*SE* = 1.03, *Wald* χ^2 = 0.37, *p* = 0.55). These results are similar to the decomposition of the total effect when employment is measured on a binary scale.

RMPW Procedure for Continuous Mediators

NEWWS Administrative records contain information about the number of quarters a welfare recipient was employed during the two years after randomization. A continuous mediator typically provides richer information than the related categorical measures. Conventional mediation analysis (e.g., SEM) requires specifying the functional form of the relationship between the outcome and the continuous mediator. We propose the RMPW method as an alternative that builds on a distributional assumption of the mediator under each treatment condition.

Suppose that the mediator values are normally distributed, that is, $Z_a \sim N(\mu_a, \sigma_a^2)$ for $a = 0, 1$, within levels of pretreatment covariates $\mathbf{X} = \mathbf{x}$. For each individual unit, we may predict $\mu_1 = E(Z_1|\mathbf{X})$ and $\mu_0 = E(Z_0|\mathbf{X})$ through regressing Z on $\mathbf{X}$ in each treatment group and applying the same model to the alternative treatment group. Additionally, through regressing Z^2

on **X** in each treatment group and applying the same model to the alternative group, we obtain $\varphi_1 = E(Z_1|\mathbf{X})^2$ and $\varphi_0 = E(Z_0|\mathbf{X})^2$ for each unit. Hence the conditional variance can be computed as $\sigma_a^2 = \varphi_a - \mu_a^2$ for $a = 0, 1$. We then estimate the density for a given mediator value z as follows:

$$\theta^{(Z_a=z)} = q^{(a)}(Z_a = z|A = a, \mathbf{X}) = (2\pi\sigma_a^2)^{-\frac{1}{2}} exp\left[-\frac{(z-\mu_a)^2}{2\sigma_a^2}\right], \text{for } a = 0, 1.$$

Following Equation (4), we compute the weight for LFA units displaying mediator value z: $W_{(1Z_0)} = \theta^{(Z_0=z)}/\theta^{(Z_1=z)}$, which can be simplified as

$$W_{(1Z_0)} = \frac{\sigma_1}{\sigma_0} exp\left[\frac{(z-\mu_1)^2}{2\sigma_1^2} - \frac{(z-\mu_0)^2}{2\sigma_0^2}\right].$$

The subsequent procedure is the same as that for analyzing a multi-category mediator.

In the NEWWS data, however, the number of quarters employed under each treatment condition is apparently not normally distributed. As mentioned earlier, as many as 60% of the welfare recipients in the control group and about 35% of those in the LFA group were never employed during the two years after randomization. Hence we assume that the continuous mediator has a joint distribution of a Bernoulli random variable indicating whether one was ever employed during the two-year period and a binomial random variable recording a nonzero number of quarters employed. In treatment group a and within levels of pretreatment covariates $\mathbf{X} = \mathbf{x}$, let $Z_1 = 1$ if an individual unit was ever employed and 0 otherwise with distribution $Z_1 \sim Bernoulli(p_1)$, where $p_1 = pr(Z_1 = 1|A = a, \mathbf{X})$. When $Z_1 = 1$, let $Z_2 = z$ be the number of quarters employed over a total of eight quarters with distribution $Z_2 \sim Binomial(8, p_2)$, where $p_2 = pr(M_t = 1|A = a, \mathbf{X})$ for $t = 1, \ldots, 8$ is the average conditional probability that one was employed (i.e., $M_t = 1$) in a given quarter. If one was unemployed in quarter t, then $M_t = 0$. The joint distribution of Z_1 and Z_2 can be written as

$$\theta^{(Z_a=z)} = pr(Z_1, Z_2|A = a, \mathbf{X}) = pr(Z_1|A = a, \mathbf{X})pr(Z_2|A = a, Z_1 = 1, \mathbf{X}).$$

We have that,

when $Z_1 = 0$, $pr(Z_1, Z_2|A = a, \mathbf{X}) = pr(Z_1 = 0|A = a, \mathbf{X}) = 1 - p_1$;

when $Z_1 = 1$ and $Z_2 = z$, $pr(Z_1, Z_2|A = a, \mathbf{X}) = p_1 \binom{8}{z} {p_2}^z (1 - p_2)^{1-z}$.

As we have discussed earlier for a binary mediator, p_1 can be estimated as the propensity score for ever being employed under treatment a. Analyzing the pooled quarterly employment data of those who were ever employed in each treatment group, we obtain an estimate of p_2 as the propensity score for being employed in a given quarter under treatment a. Following Equation (4), we compute the weight for LFA units displaying mediator value z for $z = 0, 1, \ldots, 8$: $W_{(1Z_0)} = \theta^{(Z_0=z)}/\theta^{(Z_1=z)}$. The subsequent procedure is the same as that for analyzing a multi-category mediator.

Conclusion and Discussion

This paper has presented a new weighting approach to causal mediation analysis. The conventional path analysis/SEM approach assumes that there is no treatment-by-mediator interaction. Recent extensions of the linear regression method relax this assumption typically by invoking model-based assumptions with regard to how the treatment, the mediator, and the covariates interact in the structural model for the outcome. Many existing methods also require combining multiple parametric models in estimating the indirect effect. In contrast, the RMPW method specifies the outcome simply as a function of the natural direct effect and the natural indirect effect. Generalized least squares analysis of the weighted data generates robust standard errors for both the natural direct effect and the natural indirect effect estimates and therefore provides direct tests of the null hypotheses. Because the outcome model does not involve model-based assumptions, this new strategy is suitable for handling a large number of pretreatment covariates without a need to specify multi-way interactions among the treatment, the mediator, and the covariates. Due to the simplicity and the non-parametric nature of the outcome model, the RMPW strategy applies regardless of the distribution of the outcome, the distribution of the mediator, or the functional relationship between the outcome and the mediator.

According to the simulation results with a binary treatment and a binary mediator, both parametric and non-parametric weighting demonstrate satisfactory performance under the identification assumptions. When there is no treatment-by-mediator interaction in the simulated data, the weighting results replicate the results from path analysis. When there is a zero direct effect in the simulated data, the weighting results replicate the IV results. As expected, the weighting methods outperform the conventional methods especially when the assumption of no treatment-by-mediator interaction or the exclusion restriction does not hold. Hence the RMPW method provides a viable alternative to the existing methods for causal mediation analysis. We have shown its extension to analyses of moderated mediation effects as well as to multi-category and continuous mediators. Future research may extend this new approach to studies of multiple concurrent mediators, time-varying mediators, and mediation problems in multi-level data.

Similar to some of the latest advancements in causal mediation analysis, the RMPW method identifies the natural direct effect and the natural indirect effect under the untestable assumption of sequential ignorability. Even though the ignorability of treatment assignment can be warranted by a randomized experiment, mediator value assignment is typically not randomized. Similar to most other existing methods, the RMPW method removes selection bias associated the observed pretreatment covariates. The result will be biased if the mediator-outcome relationship within a treatment condition or across treatment conditions is confounded by omitted covariates. Sensitivity analysis may be employed to assess the consequence of such an omission (Imai, Keele, & Tingley, 2010; Imai, Keele, & Yamamoto, 2010; VanderWeele, 2010). Additionally, as Imai (2012) pointed out in his commentary on an application of the RMPW method (Hong & Nomi, 2012), whenever propensity scores are employed to remove selection bias associated with the observed pretreatment covariates, it is a major challenge to correctly specify the parametric propensity score models. Imai and Ratkovic (2012) have proposed using the generalized method of moments estimation to obtain covariate balancing propensity scores. This new strategy will likely improve the performance of propensity score-based weighting methods including the RMPW method.

References

Angrist, J. D., Imbens, G. W., & Rubin, D. B. (1996). Identification of causal effects using instrumental variables (with discussion), *Journal of the American Statistical Association*, *91*, 444-472.

Austin, P. C. (2011). Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies. *Pharmaceutical Statistics*, **10**(2), 150-161.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*, 1173-1182.

Bullock, J. G., Green, D. P., & Ha, S. E. (2010). Yes, But What's the Mechanism? (Don't Expect an Easy Answer). *Journal of Personality and Social Psychology*, **98**, 550-558.

Coffman, D. L. & Zhong, W. (2012). Assessing mediation using marginal structural models in the presence of confounding and moderation. *Psychological Methods*. Advance online publication. Doi:10.1037/a0029311.

Derr, M. K., Hill, H., & Pavetti, L. (2000). *Addressing mental health problems among TANF recipients: A guide for program administrators*. Washington, DC: Mathematica Policy Research.

Drake, C. (1993). Effects of misspecification of the propensity score on estimators of treatments effects. *Biometrics*, **49**, 1231-1236.

Duncan, O. D. (1966). "Path analysis: Sociological examples," *American Journal of Sociology*, *72*, 1-16.

Elliott, M. R., Raghunathan, T. E., & Li, Y. (2010). Bayesian inference for causal mediation effects using principal stratification with dichotomous mediators and outcomes. *Biostatistics*, *11*, 353-372.

Frangakis, C. E., & Rubin, D. B. (2002). "Principal stratification in causal inference," *Biometrics*, *58*, 21-29.

Gallop, R., Small, D. S., Lin, J. Y., Elliott M. R., Joffee M., & Ten Have, T. R. (2009). Mediation analysis with principal stratification. *Statistics in Medicine*, *28*, 1108-1130.

Hamilton, G. (2002). *Moving people from welfare to work: Lessons from the National Evaluation of Welfare-to-Work Strategies*. Manpower Demonstration Research Corporation.

Heckman, J. J., & Robb, R. (1985). Alternative methods for evaluating the impact of an intervention. *Journal of Econometrics*, *30*, 239-267.

Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association*, **81**, 945-960.

Holland, P. (1988). "Causal inference, path analysis, and recursive structural equations models," *Sociological methodology*, **18**, 449-484.

Hong, G. (2006). Multilevel experimental designs and quasi-experimental approximations for studying intervention implementation as a mediator. Paper presented at the 2006 Annual Meeting of the American Educational Research Association, San Francisco, CA.

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

Hong, G. (2012a). Marginal mean weighting through stratification: A generalized method for evaluating multi-valued and multiple treatments with non-experimental data. *Psychological Methods*, *17*(1), 44-60.

Hong, G. (2012b). Covariate-informed parallel design. *Journal of the Royal Statistical Society, Series A (Statistics in Society)*. Forthcoming.

Hong, G., & Nomi, T. (2012). Weighting methods for assessing policy effects mediated by peer change. *Journal of Research on Educational Effectiveness* special issue on the statistical approaches to studying mediator effects in education research, *5*(3), 261-289.

Imai, K. (2012). Comments: Improving weighting methods for causal mediation analysis. *Journal of Research on Educational Effectiveness* special issue on the statistical approaches to studying mediator effects in education research, *5*(3), 293-295.

Imai, K., Keele, L., & Tingley, D. (2010). A general approach to causal mediation analysis. *Psychological Methods*, **15**(4), 309-334.

Imai, K., Keele, L., & Yamamoto, T. (2010). Identification, inference and sensitivity analysis for causal mediation effects. *Statistical Science*, **25**(1), 51-71.

Imai, K., & Ratkovic, M. (2012). Covariate balancing propensity score (Working paper). Available at http://imai.princeton.edu/research/CBPS.html.

Imai, K., Tingley, D., & Yamamoto, T. (2012). Experimental designs for identifying causal mech-anisms (with discussions). *Journal of the Royal Statistical Society, Series A (Statistics in Society)*. Forthcoming.

Jagannathan, R., Camasso, M. J., & Sambamoorthi, U. (2010). Experimental evidence of welfare reform impact on clinical anxiety and depression levels among poor women. *Social Science & Medicine*, **71**(1), 152-160.

Jo, B. (2008). Causal inference in randomized experiments with mediational processes. *Psychological Methods*, *13*, 314-336.

Jöreskog, K. G. (1970). A general method for analysis of covariance structures. *Biometrika*, *57*, 239-251.

Kling, J. R., Liebman, J. B., & Katz, L. F. (2007). Experimental analysis of neighborhood effects. *Econometrica*, *75*(1), 83-119.

Little, R. J. A., & Rubin, D. B. (2002). *Statistical analysis with missing data*. New York: Wiley.

Mattei, A. and Mealli, F. (2011). Augmented designs to assess principal strata direct effects. *Journal of the Royal Statistical Society, Series B (Methodological)*, *73*(5), 729-752.

Page, L. C. (2012). Principal stratification as a framework for investigating mediational processes in experimental settings. *Journal of Research on Educational Effectiveness* Special Issue on the Statistical Approaches to Studying mediator Effects in Education Research, *5*(3), 215-244.

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.

Raudenbush, S. W., Reardon, S. F., Nomi, T. (2012). Statistical analysis for multisite trials using instrumental variables with random coefficients. *Journal of Research on Educational Effectiveness* Special Issue on the Statistical Approaches to Studying mediator Effects in Education Research, *5*(3), 303-332.

Robins, J. M. (1999). Marginal structural models versus structural nested models as tools for causal inference. In M. Elizabeth Halloran and Donald Berry (Eds.), *Statistical Models in Epidemiology, the Environment, and Clinical Trials* (pp.95-134). New York: Springer.

Robins, J. M. (2003). Semantics of causal DAG models and the identification of direct and indirect effects. In P. J. Green, N. L. Hjort, & S. Richardson (Eds.), *Highly structured stochastic systems* (pp. 70–81). New York, NY: Oxford University Press.

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

Rosenbaum, P. R. (1987). Model-based direct adjustment. *Journal of the American Statistical Association*, **82**, 387-394.

Rubin, D. B. (1978), Bayesian inference for causal effects: The role of randomization. *The Annals of Statistics*, **6**, 34-58.

Rubin, D. B. (2004). Direct and indirect causal effects via potential outcomes. *Scandinavian Journal of Statistics*, *31*, 161-170.

Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. In S. Leinhart (Ed.), *Sociological methodology* (pp.290-312). San Francisco: Jossey-Bass.

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

Sobel, M. E., & Stuart, E. A. (2012). Comments. *Journal of Research on Educational Effectiveness* Special Issue on the Statistical Approaches to Studying mediator Effects in Education Research, *5*(3), 290-292.

Valeri, L., & VanderWeele, T. J. (forthcoming). Mediation analysis allowing for exposure-mediator interactions and causal interpretation: Theoretical assumptions and implementation with SAS and SPSS macros. *Psychological Methods*.

van der Laan, M. J., & Peterson, M. L. (2008). Direct effect models. *International Journal of Biostatistics*, **4**(1), Article 23.

VanderWeele, T. J. (2009). Marginal structural models for the estimation of direct and indirect effects. *Epidemiology*, **20**, 18-26.

VanderWeele, T. J. (2011). Principal stratification: Uses and limitations. *International Journal of Biostatistics*, *7*, Article 28, 1-14.

VanderWeele, T. J. (2012). Comments: Should principal stratification be used to study mediational processes? *Journal of Research on Educational Effectiveness* Special Issue on the Statistical Approaches to Studying mediator Effects in Education Research, *5*(3), 245-249.

Wright, S. (1934). The method of path coefficients. *Annals of Mathematical Statistics*, *5*, 161-215.

Appendix A

Bias in Path Analysis Estimates in the Presence of Treatment-by-Mediator Interaction

In a hypothetical sequential randomized experiment, suppose that $pr(Z_0 = 1) = \beta_0$ and that $pr(Z_1 = 1) = \beta_0 + \beta_1$. Also suppose that the data generation function for the potential outcomes is $Y_{az} = \theta_0 + \theta_1 a + \theta_2 z + \theta_3 az + \varepsilon$. Hence the total effect is $\theta_1 + \theta_3\beta_0 + (\theta_2 + \theta_3)\beta_1$, the natural direct effect is $\theta_1 + \theta_3\beta_0$, and the natural indirect effect is $(\theta_2 + \theta_3)\beta_1$.

Path analysis invokes the assumption of linearity and additivity (Holland, 1988) and specifies the observed outcome model as $Y = \gamma_0 + \gamma_1 A + \gamma_2 Z + e$. We can show that $\gamma_2 = \theta_2 + \theta_3 \times pr(A = 1)$. The natural indirect effect estimate is

$$\gamma_2\beta_1 = \theta_2\beta_1 + \theta_3\beta_1 \times pr(A = 1) = (\theta_2 + \theta_3)\beta_1 - \theta_3\beta_1 \times pr(A = 0).$$

Hence the bias in the natural indirect effect estimate is $-\theta_3\beta_1 \times pr(A = 0)$, which is equivalent to $-E\{(Y_{11} - Y_{01}) - (Y_{10} - Y_{00})\} \times \{pr(Z_1 = 1) - pr(Z_0 = 1)\} \times pr(A = 0)$.

Appendix B

Proof of Theorem

We derive a weight $W_{(aZ_{a'})}$ such that $E\left(Y_{aZ_{a'}}\right)$ can be consistently estimated by $E\left(W_{(aZ_{a'})}Y|A=a\right)$.

$$E\left(Y_{aZ_{a'}}\right) \equiv E\left\{E\left(Y_{aZ_{a'}}|\mathbf{X}\right)\right\}.$$

By Assumptions 1 and 2, the above is equal to

$$E\left\{E\left(Y_{aZ_{a'}}|A=a,\mathbf{X}\right)\right\}$$

$$\equiv \iiint_{\mathbf{x},z,y} y \times f(Y_{az}=y|A=a,Z_{a'}=z,\mathbf{X}=\mathbf{x}) \times q^{(a')}(Z_{a'}=z|A=a,\mathbf{X}=\mathbf{x})$$

$$\times h(\mathbf{X}=\mathbf{x})dydzd\mathbf{x},$$

which, by Assumptions 4, 5, and 6, is equal to

$$\iiint_{\mathbf{x},z,y} y \times f(Y_{az}=y|A=a,Z_a=z,\mathbf{X}=\mathbf{x}) \times q^{(a')}(Z_{a'}=z|A=a',\mathbf{X}=\mathbf{x}) \times h(\mathbf{X}=\mathbf{x})dydzd\mathbf{x}$$

which, by Bayes Theorem and by Assumptions 1, 2, and 3* is equal to

$$\iiint_{\mathbf{x},z,y} y \times f(Y_{az}=y|A=a,Z_a=z,\mathbf{X}=\mathbf{x}) \times q^{(a)}(Z_a=z|A=a,\mathbf{X}=\mathbf{x}) \times h(\mathbf{X}=\mathbf{x}|A=a)$$

$$\times \frac{q^{(a')}(Z_{a'}=z|A=a',\mathbf{X}=\mathbf{x})}{q^{(a)}(Z_a=z|A=a,\mathbf{X}=\mathbf{x})} \times \frac{p(A=a)}{p(A=a|\mathbf{X}=\mathbf{x})} dydzd\mathbf{x} = E(Y^*|A=a),$$

where $Y^* = W_{(aZ_{a'})}Y$ and

$$W_{(aZ_{a'})} = \left\{q^{(a')}(Z_{a'}=z|A=a',\mathbf{X})/q^{(a)}(Z_a=z|A=a,\mathbf{X})\right\} \times \{p(A=a)/p(A=a|\mathbf{X})\}.$$

This concludes the proof. □

Footnotes

[1] There is an emerging literature on experimental designs for investigating causal mediation mechanisms (Hong, 2012b; Imai, Tingley, & Yamamoto, 2012; Mattei & Mealli, 2011; Sobel & Stuart, 2012). A review and comparison across these different designs is beyond the scope of this paper.

[2] This assumption can be replaced by a weaker assumption that the conditional support for mediator values under the control condition does not exceed that under the experimental condition. Hence the analytic sample may include participants who would always be employed or would always be unemployed regardless of the policy assignment.

[3] The Stata command for analyzing the outcome model is simply:

reg Y A D [weight=RMPW], vce(cluster clustvar);

or to obtain standard errors through bootstrapping, the command is:

reg Y A D [weight=RMPW], vce(bootstrap).

Table 1

Parameter Values for the Three Sets of Simulations

	$E(\theta_{Z_0})$	$E(\theta_{Z_1})$	$\delta^{(A)}$	$\delta^{(Z)}$	$\delta^{(AZ)}$	$\gamma^{(NDE)}$	$\gamma^{(NIE)}$
(a)	.3918	.6609	0	0	0	0	0
(b)	.3918	.6609	2.5500	1.5000	-4.5000	0.7869	-0.8073
(c)	.1938	.8066	1.5000	2.5000	-3.7500	0.77325	-0.7660

Table 2

Summary of Simulation Results

	Baseline Model	N = 5,000			N = 800		
		RMPW	*NRMPW 3×3*	*NRMPW 4×4*	*RMPW*	*NRMPW 3×3*	*NRMPW 4×4*
Natural Direct Effect							
% Bias removal ($\hat{\gamma}^{(NDE)}$)	(a)	0.999	0.854	0.903	0.984	0.863	0.918
	(b)	0.980	0.873	0.922	0.999	0.868	0.891
	(c)	0.995	0.858	0.907	0.988	0.857	0.877
Relative efficiency ($\hat{\gamma}^{(NDE)}$)	(a)	0.950	1.583	1.727	0.885	1.389	1.389
	(b)	0.972	1.207	1.250	0.945	1.132	1.090
	(c)	0.741	1.162	1.194	0.656	1.004	1.030
MSE($\hat{\gamma}^{(NDE)}$)	(a)	0.002	0.002	0.002	0.011	0.008	0.008
	(b)	0.004	0.004	0.003	0.022	0.019	0.020
	(c)	0.006	0.010	0.006	0.037	0.030	0.028
Natural Indirect Effect							
% Bias removal ($\hat{\gamma}^{(NIE)}$)	(a)	0.998	0.856	0.904	0.998	0.857	0.910
	(b)	0.991	0.865	0.913	0.990	0.865	0.886
	(c)	0.999	0.856	0.905	0.994	0.855	0.875
Relative efficiency ($\hat{\gamma}^{(NIE)}$)	(a)	1.333	2.000	2.000	1.000	1.353	1.211
	(b)	0.579	0.688	0.688	0.561	0.670	0.616
	(c)	0.559	0.950	0.905	0.491	0.779	0.820
MSE($\hat{\gamma}^{(NIE)}$)	(a)	3×10^{-4}	0.001	0.001	0.002	0.003	0.002
	(b)	0.002	0.003	0.002	0.012	0.011	0.012
	(c)	0.003	0.008	0.005	0.022	0.020	0.018