Estimating Causal Direct and Indirect Effects in the Presence of Post-Treatment Confounders: A Simulation Study

by

Yasemin Kisbu Sakarya

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved November 2013 by the
Graduate Supervisory Committee:

David P. MacKinnon, Chair
Leona S. Aiken
Stephen G. West
Roger E. Millsap

ARIZONA STATE UNIVERSITY

December 2013

ABSTRACT

In investigating mediating processes, researchers usually use randomized experiments and linear regression or structural equation modeling to determine if the treatment affects the hypothesized mediator and if the mediator affects the targeted outcome. However, randomizing the treatment will not yield accurate causal path estimates unless certain assumptions are satisfied. Since randomization of the mediator may not be plausible for most studies (i.e., the mediator status is not randomly assigned, but self-selected by participants), both the direct and indirect effects may be biased by confounding variables. The purpose of this dissertation is (1) to investigate the extent to which traditional mediation methods are affected by confounding variables and (2) to assess the statistical performance of several modern methods to address confounding variable effects in mediation analysis. This dissertation first reviewed the theoretical foundations of causal inference in statistical mediation analysis, modern statistical analysis for causal inference, and then described different methods to estimate causal direct and indirect effects in the presence of two post-treatment confounders. A large simulation study was designed to evaluate the extent to which ordinary regression and modern causal inference methods are able to obtain correct estimates of the direct and indirect effects when confounding variables that are present in the population are not included in the analysis. Five methods were compared in terms of bias, relative bias, mean square error, statistical power, Type I error rates, and confidence interval coverage to test how robust the methods are to the violation of the no unmeasured confounders assumption and confounder effect sizes. The methods explored were linear regression with adjustment, inverse propensity weighting, inverse propensity weighting with

truncated weights, sequential g-estimation, and a doubly robust sequential g-estimation. Results showed that in estimating the direct and indirect effects, in general, sequential g-estimation performed the best in terms of bias, Type I error rates, power, and coverage across different confounder effect, direct effect, and sample sizes when all confounders were included in the estimation. When one of the two confounders were omitted from the estimation process, in general, none of the methods had acceptable relative bias in the simulation study. Omitting one of the confounders from estimation corresponded to the common case in mediation studies where no measure of a confounder is available but a confounder may affect the analysis. Failing to measure potential post-treatment confounder variables in a mediation model leads to biased estimates regardless of the analysis method used and emphasizes the importance of sensitivity analysis for causal mediation analysis.

To Eser

ACKNOWLEDGMENTS

I would like to thank David MacKinnon, Leona Aiken, Stephen West, and Roger Millsap for serving as committee members and for their invaluable input into my work. I am especially grateful to my advisor David MacKinnon for his support and mentorship throughout my studies at ASU. I thank my family and friends for all their support. Finally, I am truly grateful to my husband, Eser, for his love, patience, and support.

TABLE OF CONTENTS

Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

Figure Page

Figure Page

Figure Page

Chapter 1

INTRODUCTION

Testing mediating processes is important in social and medical sciences. Mediation analysis allows researchers to investigate the underlying mechanisms of a treatment and to address competing explanations, whereas a randomized experiment focusing on group differences in outcomes is generally not adequate to reveal the causal processes underlying how a treatment achieved its effects. In a typical mediation model, an independent variable (X) causes a mediator (M), and then the mediator causes an outcome (Y) (Judd & Kenny, 1981; Baron & Kenny, 1986; MacKinnon, 2008). Examples of mediation from psychological sciences include the following: leader expectations influence performance through subordinate self-efficacy (Davidson & Eden, 2000); dietary social norms influence healthy eating through intentions (Ranby et al., 2011); workload influences job satisfaction through employees' perceived control over time (Claessens, et al., 2004); perceived justice influences health through sleep quality (Elovainio et al., 2009); and transformational leadership influences organizational citizenship behavior through leader-member exchange (Wang et al., 2005).

A key goal in scientific research is causal inference. Randomized experiments are often seen as the best method to achieve causal inference because they balance possible confounding variables in a way that observed differences between treatment groups can be attributed to the treatment (that is, randomization ensures the strong ignorability of the treatment assignment). Researchers using a mediation approach also typically conduct randomized experiments to investigate the mediated (indirect) effect of a randomized

treatment on the targeted outcome through the hypothesized mediator. However, work on causal inference in mediation shows that only randomizing the treatment does not ensure obtaining accurate causal path estimates for the relation of M to Y, so that both the direct and indirect effects are still subject to potential confounding variables (Holland, 1988). Therefore, it is critical to consider causal estimation issues when testing mediating mechanisms.

I first describe the single mediation model from a linear regression approach and then the potential outcomes framework. The assumptions for causal inference in mediation are outlined. Next, modern causal inference methods in the presence of a mediator are described.

1.1 Linear Regression Approach to Mediation

The most common approach to mediation employs OLS regression (Baron & Kenny, 1986; MacKinnon, 2008). The basic mediation model can be summarized in three equations including three variables: X, the treatment variable, Y, the dependent variable, and M, the mediator (MacKinnon & Dwyer, 1993):

$$(1) \quad E(Y \mid X = x) = i_1 + c\,X + e_1$$

$$(2) \quad E(M \mid X = x) = i_2 + a\,X + e_2$$

$$(3) \quad E(Y \mid X=x, M=m) = i_3 + c'\,X + b\,M + e_3$$

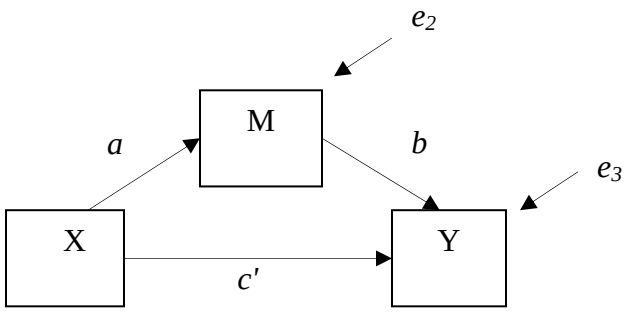


Figure 1. Single mediator model

Equation 1 gives the expected Y given X, where X can take on values *x.* In other words, it estimates the effect of the treatment X on outcome Y (the *c* regression coefficient). Equation 2 predicts the effect of X on the mediator (the *a* path). Equation 3, where X can take on values *x* and M can take values *m*, estimates the effect of treatment X on the outcome Y adjusting for the effects of the mediating variable M (the *c'* path). i_1, i_2, and i_3 are intercepts; and e_1, e_2, and e_3 are errors which are assumed to be independent across equations. Note that lower case letters, x, m, and y represent values of variables X, M, and Y, respectively. This distinction between the variables and the values of the variables defines causal effects at different values, X=x, M=m, and Y=y and allows for the possibility that different causal effects may be obtained at different values, x, m, and y of variables X, M, and Y, respectively.

Three approaches to test for mediation are (1) the causal steps approach also known as the Baron and Kenny test of mediation (1986; Hyman, 1955; Judd & Kenny,

1981), (2) difference in coefficients method, and (3) the product of coefficients method (MacKinnon et al., 2002). In the Baron and Kenny method, a series of statistical tests are conducted to decide if the data are consistent with mediation. Specifically the approach requires the following four steps to be conducted:

i) The effect of X on Y in equation (1) must be significant (i.e., the *c* path). In other words, there must be a significant total effect of independent variable on outcome.

ii) The effect of X on M in equation (2) must be significant (i.e., the *a* path). In other words, the independent variable should significantly affect the mediator.

iii) The effect of M on Y adjusted for X must be significant (i.e., the *b* path). In other words, mediator should be significantly related to the outcome variable even after controlling the effect of the independent variable.

iv) The relation between X and Y should be weaker when the mediator is added to the model. In other words, *c-c'* should be greater than zero.

Although this approach is the most common method used by researchers, it has important limitations. First, the method does not include estimating the magnitude and significance of the mediated effect. Also, it requires a significant total effect of X on Y (i.e., *c* path), yet it is possible that the effect of X on Y is nonsignificant but there is a mediated effect (MacKinnon, 2008). The causal steps method also requires much larger sample sizes compared to other methods to test for mediation (Fritz & Mackinnon, 2007).

The second approach is the difference in coefficients method. This method estimates the mediated effect as the difference between the total and direct effects, *c-c'*. It

is also possible to compute the standard error, confidence intervals, and significance testing for the *c*-*c*′ measure of the mediated effect. The difference in coefficients approach has an important limitation that it is cumbersome to compute individual mediated effects in a multiple mediator model (MacKinnon, 2008; MacKinnon, Kisbu-Sakarya, Gottschall, 2013).

The third approach is to calculate the point estimate of the mediated effect as the product of coefficients, *ab*; and then divide *ab* by its standard error and compare that ratio to the normal distribution to test for statistical significance. Whether the relation between X and Y is mediated through M is examined by comparing the ratio of the effect to its standard error to the normal distribution in the sample. The commonly used standard error for *ab* derived by Sobel (1982) is based on the multivariate delta method:

$$SE(ab) = \sqrt{a^2\, SE(b)^2 + b^2 SE(a)^2}$$

$$CI(ab) = ab \pm SE(ab) * z_{Type\ I\ error}$$

Confidence intervals for the product of coefficients, *ab*, method can also be computed by several other methods such as resampling methods which create an empirical sampling distribution of *ab*. Another method takes the nonnormality of the distribution of the product into account (Meeker, Cornwell, & Aroian, 1981), and constructs the confidence intervals based on the distribution of the product of two normally distributed variables (MacKinnon, Lockwood, & Williams, 2004; MacKinnon

et al., 2002; MacKinnon, Fritz, Williams, & Lockwood, 2007). Furthermore, Bayesian methods are available to compute credible intervals for the indirect effect (Yuan & MacKinnon, 2009; Pirlott et al., 2012). The product of coefficients approach is used in this dissertation.

Mediation analysis by linear regression has several assumptions. First, it is assumed that the variables in the model are reliable and valid measures of the study variables. In other words, it is assumed that there is no measurement error that may cause bias in the estimators. It is also assumed that the variables are continuous and normally distributed; however, models with non-normally distributed variables may be estimated accurately using transformations (Cohen, Cohen, West, & Aiken, 2003; MacKinnon, 2008).

There are also assumptions regarding the causal nature of the mediation model to identify the mediated effect. It is assumed that the causal paths between X, M, and Y have the correct functional form and do not have bidirectional effects. Also, it is assumed that there are no omitted variables affecting the causal paths in the mediation model. This assumption will be described in more detail in the following section (Holland, 1988; Robins & Greenland, 1992; Pearl 2001, 2012).

1.2 Estimating Causal Effects in the Presence of a Mediator

The potential outcomes approach provides a new framework to interpret mediation effects. In the case in which all assumptions are satisfied, the traditional estimator of the mediated effect, *ab*, described above is the causal indirect effect estimator. The strength of the potential outcomes approach to mediation lies in how it

clarifies underlying assumptions of traditional mediated effect estimation and how it provides a framework to estimate mediated effects in more complex models such as nonlinear models and models with confounding.

The potential outcomes approach to causal effects (Rubin, 1974, 2004, 2005; Holland, 1986, 1988; Morgan & Winship, 2007) defines the individual causal effect using the potential outcomes of the same individual. Average causal effects are then defined based on averaging effects across individuals. These average causal effects solve several problems with the estimation of causal effects for each individual. Starting with the individual level causal effect, let variable X be a treatment program with level *x* (*x*=1 for the treatment, *x*=0 for the control) and variable Y the outcome variable. An individual may be assigned to the treatment group (*x*=1) and obtain the potential outcome value Y(1). The second potential outcome for that individual is the value she would have obtained on the outcome variable if she had been assigned to the control condition (*x*=0), that is Y(0) (also referred to as the counterfactual value). The corresponding individual causal effect is then equal to the difference between the potential outcomes, Y(1) – Y(0). However, because it is often not possible to observe both outcomes for the same person (referred as the "fundamental problem of causal inference" by Holland, 1988), averages of individuals are used to compute the average causal effect, E[Y(1) – Y(0)]. The average causal effect, the difference between the means in the treatment and control groups, is a causal effect when units are randomized to conditions and the randomization has been successful.

Now suppose a potential mediating variable M with level *m* mediates the relation between X and Y for an individual. Let Y(*x*, *m*) denote the potential outcome for an individual under the treatment level *x* and mediator level *m*. X takes the value *x*=0 for the control group, and *x*=1 for the treatment group. M is a continuous variable, and if the observed value of M is *m* for an individual, then the counterfactual value of M for that individual is denoted as *m'*. Including the mediator in the model leads to the formulation of the following effects in the potential outcome framework: controlled direct effect, natural direct effect and natural indirect effect (Pearl, 2001, 2009; Robins & Greenland, 1992).

1.2.1 Natural and Controlled Effects

The controlled direct effect is the effect of X on Y at a specific value m of M. More formally, the controlled direct effect (CDE) of a treatment on the outcome is the difference between the potential outcome scores when the individuals' mediating variable score was controlled and set to a specific value (Robins & Greenland, 1992).

(4) $\quad CDE = Y(1, m) - Y(0, m).$

As opposed to the controlled direct effects that are measured at a fixed level of the mediator, the natural direct effect allows individuals to take varying values of the mediator-the value of the mediator that they would have naturally. The *average natural direct effect* (NDE) of X on Y is different from the average controlled direct effect in that M is set to the level that would have naturally occurred under one of the conditions of X

(i.e., M_x). For instance, in the case of M_0, the natural direct effect is the effect of treatment X on outcome Y when X did not influence the mediator M (or the individuals were assigned the mediator level under the control condition):

$$(5) \quad NDE = Y(1, M_x) - Y(0, M_x).$$

Following the same approach, the *average natural indirect effect* is the effect of the treatment on outcome when changing the level of the mediator but keeping X at the same value such as X in the control group:

$$(6) \quad NIE = Y(X, M_1) - Y(X, M_0).$$

In other words, the above formula indicates the effect of treatment on the outcome when the level of M is changed while X is set to a certain value (0 or 1 in this case of two conditions).

The sum of the natural direct effect and natural indirect effect equals the total effect. This additivity assumption is also sometimes referred as the no X and M interaction assumption. Yet, Pearl (2001) demonstrated that the decomposition of total effect into the natural direct and indirect effects holds even in models with interactions and importantly also in models with non-linear effects such as logistic regression.

The following two assumptions are required for the controlled direct effect to be identified (Pearl, 2001; VanderWeele, & Vansteelandt, 2009; VanderWeele, 2010, 2011):

(i) No unmeasured confounder for the relation between X and Y.

(ii) No unmeasured confounder for the relation between M and Y.

The following two assumptions are required in addition to the two assumptions above for the natural direct and indirect effects to be identified:

(iii) No unmeasured confounder for the relation between X and M

(iv) No unmeasured M to Y confounder affected by treatment.

A confounder is an extraneous variable that correlates with both the independent and the dependent variable. Omitting a confounder from a statistical model may lead to the misestimation of the statistical model. Assumptions (i) and (iii) refer to the *ignorability* of treatment assignment conditional on the observed pretreatment confounders. This assumption is usually satisfied with randomization of X. Assumption (ii) refers to the ignorability of the mediator conditional on the observed treatment and pretreatment confounders. In other words, there are no unmeasured confounders influencing the *b* path. This assumption is difficult to meet because randomization of M is usually not plausible for many studies (i.e., the mediator status is not randomly assigned, but rather self-selected by individuals). Even though we condition on observed confounders for the relation between M and Y, there can still be unobserved confounders. This assumption (ii) is strong and is usually ignored in studies, even though we cannot have a causal interpretation of a mediated effect without its existence (MacKinnon et al., 2013; MacKinnon, 2008, Chapter 13).

The linear regression approach to mediation assumes sequential ignorability, which consists of the ignorability of the treatment assignment and the ignorability of the

mediator. In other words, by successful randomization of the treatment X, we can achieve causal estimation of the *c* (i.e., the total effect of X on Y) and *a* paths (Holland, 1988). Sequential ignorability also assumes that M is randomly assigned, but this is not often possible; research participants self-select their value of the mediator. Thus further assumptions are required for causal interpretation for the *b* and *c'* paths. If individuals are randomly assigned to a treatment, there should be no confounders of the X to Y, and X to M relations. However, if individuals are not randomly assigned to values of the mediator, there can be confounders affecting M and Y, leading to inaccurate estimates of the effect of M on Y and the effect of X on Y adjusted for M. Several solutions to improve the interpretation of the *b* and *c'* coefficients as causal effects have been proposed and these methods are a focus of this dissertation as described below.

1.2.1.1 Inverse propensity weighting method

Returning to the case of one X and one Y variable, causal effects of a nonrandomized treatment on an outcome can be estimated using propensity scores that account for the effects of potential confounders of the X to Y relation. In this section, estimation using propensity scores is described first followed by the use of propensity scores in causal mediation.

Propensity scores

In the case of an effect of treatment on the outcome, the propensity score is the estimated probability of receiving the treatment given measured confounders (Rosenbaum & Rubin, 1983). Because the confounders used to estimate the propensity score are either variables that do not change such as gender, or variables measured at

baseline, the estimated propensity scores are not influenced by the treatment. Therefore, assuming all confounders are measured, comparing the treatment and control groups with similar estimated propensity scores is a causal estimator of the unconfounded effect of X on Y. In other words, propensity scores balance the distribution of confounders in the treatment and control groups so that the treatment assignment effect on the outcome is unconfounded given the propensity scores. An advantage of using propensity scores over analysis of covariance as a method to adjust for confounders is that including a large set of confounders in an analysis of covariance model is sometimes not practical whereas the propensity score is a single number summarizing all of the measured confounders. Moreover, the propensity score method estimates of treatment effect are more stable than the analysis of covariance estimates when the distributions of confounders in the treatment and control groups do not overlap adequately (i.e., number of individuals in treatment and control groups with similar confounder scores is low) (Rubin, 1997; King & Zeng, 2006). There are several propensity score methods for confounder adjustment; among them are matching (Rubin & Thomas, 1992, 1996; Rosenbaum & Rubin, 1985), stratification (Rosenbaum and Rubin, 1984), and weighting (Hirano & Imbens, 2001; Robins, Rotnitzky, & Zhao, 1995). In this dissertation, I focus on a weighting method called inverse propensity weighting to improve causal inference in the case of confounders affecting M to Y relation in the single mediator model.

Creating propensity scores and weighting in mediation context

For a nonrandomized treatment effect on an outcome, inverse propensity weighting makes the treated and control participants represent the population by

weighting each observation. The weights reflect the probability that each person would have received the treatment based on measured confounders. The weights are the inverse of the probability of being in the group that an individual actually participated, conditional on the confounders (C). In other words, individuals in the treatment group are weighted by 1 / P[X=1 | C] and individuals in the control group are weighted by 1 / (1 – P[X=1 | C]). In this framework, the causal inference challenge is viewed as a missing data problem (Robins, Rotnitzky, & Zhao, 1994), in that Y(1) is only observed for individuals under the treatment condition and is missing for the individuals in the control group. Inverse weighting allows individuals in the treatment to account for missing control participants with similar characteristics on the measured confounders (that is the counterfactual outcome scores). Additionally, it should be noted that in a non-randomized treatment context, all the confounders used for weighting are measured pre-treatment.

Even if the treatment is randomized, for the mediation model the M to Y relation is still subject to potential confounders. Propensity scores can be used to improve the causal interpretation of the indirect effects in a similar way as for the X to Y effect. If the mediator is binary with values of 0 and 1, then individuals with M=1 are given a weight of P[M=1 | X] / P[M=1 | X, C]. And individuals with M=0 are given a weight of (1 – P[M=1 | X]) / (1 – P[M=1 | X, C]). In mediation context, the confounders used for weighting are measured before the mediator. The weights reflect the additional prediction of the confounders compared to the prediction by treatment alone. The purpose of these weights is to create a new data set in which confounding by measured variables is removed so that the relation of M to Y more closely resembles a randomized relation. For

a binary mediator, the denominator model can be computed by a logistic regression of the mediator on measured confounders and the treatment condition. The predicted probabilities are the propensity score estimates (denoted as $\hat{\pi}$). If the mediator is continuous, then the denominator model can be computed by regressing the mediator on measured confounders and the treatment and then inserting the predicted values ($\hat{m}$) in a normal probability density function (Coffman & Zhong, 2012; Robins, Hernán, & Brumback, 2000) as shown below:

$$(7) \qquad \phi(M|X,C) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(m-\hat{m})^2}{2\sigma^2}}$$

where σ is the residual standard error from the regression of M on X and C.

Estimating the mediated effect

Using the potential outcomes framework, the mediation equations can be written in terms of marginal structural models (MSM) (Coffman & Zhong, 2012; VanderWeele & Vansteelandt, 2009). As can be seen in equations 8 and 9, marginal structural models are written in terms of potential outcomes rather than observed outcomes because the Expectations are written in terms of different potential values X=x and M=m. The MSM equations represent possible varying levels of the treatment and mediator which is used to define causal effects based on the potential outcomes.

$$(8) \qquad E\,[M \mid X=x] = i_{0M} + a\,x.$$

(9) $E\,[Y|\,M=m, X=x] = i_{0Y} + b\,m + c'\,x.$

The causal effect of a one unit increase in the level of a continuous M (from m to m') on Y in the control group can be defined as:

(10) $E\,[Y(0, m) - Y(0, m')] = (i_{0Y} + b\,m) - (i_{0Y} + b\,m') = b\,(m\text{-}m').$

If the treatment in the mediation model is randomized, then only equation 9 is weighted using the propensity scores. If the treatment is not randomized, then equation 8 (the effect of X on M) should also be weighted. The null hypothesis stating that the product of the *a* and *b* paths is equal to zero can be tested to assess mediation (Coffman & Zhong, 2012).

IPW with truncated weights

A possible problem in propensity weighting is the presence of extreme weights. Extreme variation in the weights can yield high variance and instability in the estimates. A solution to reduce the impact of extreme weights is weight truncation (Potter, 1993). Weight truncation is generally performed by trimming the weights that are larger or smaller than some values (e.g., cutpoints at the 1st or 99th percentile of the weight distribution). Yet, simulation studies show that even though weight trimming can improve the performance of propensity score weights in some conditions, it can also induce bias in other conditions (Lee, Lessler, & Stuart, 2011). Therefore, researchers are advised to use weight trimming with caution and focus more on improving the

specification of the propensity score model rather than relying on post-hoc methods such as trimming (Lee et al., 2011).

Assumptions

Two main identifying assumptions for causal effects under the propensity score model are (1) Stable Unit Treatment Value Assumption (SUTVA), and (2) all possible confounders for the X to Y, M to Y, and X to M relations are measured. The Stable unit treatment value assumption assumes that the potential outcomes for an individual do not depend on the treatment assignment or mediator level of other individuals (i.e., no interference between individuals). Also, using the propensity score approach, it is assumed that a set of confounders are measured such that conditional ignorability holds and the propensity score is greater than zero and less than one to support it. It should be noted that the no unmeasured confounders assumption is also made under regression adjustment methods such as analysis of covariance. Because this assumption is very difficult to satisfy, researchers should be careful when interpreting their results because the estimates could be affected by confounding. Furthermore, no X and M interaction on Y is assumed.

Augmented inverse propensity weighting estimator

Another inverse weighting estimator is the augmented inverse propensity weighting method. Similar to the inverse propensity weighting method, augmented inverse propensity weighting requires a two-step analysis: first running a regression model to create the propensity score, and then running the outcome regression models for the treatment and control groups separately. The confounders used in the propensity

approach to predict the probability of treatment assignment can also contain information about the outcome variable. The augmented inverse propensity weighting method incorporates this additional information in its formula by adjusting the inverse propensity weighting estimator by a weighted average of the two regression estimators of the outcome for treatment and control groups leading to more efficient estimates (Glynn & Quinn, 2010).

Relying on misspecified models can bias the estimated effects of the propensity score method. An advantage of augmented inverse propensity weighting is its double robustness, in that it is a consistent estimator of the treatment effect if either (i) the propensity score model is correctly specified or (ii) the outcome regression models are correctly specified (Schafstein, Rotnitzky, & Robins, 1999; see Glynn & Quinn, 2010 for a proof of double robustness). Therefore the augmented inverse propensity weighting method may lead to better results when there is uncertainty about either the propensity model or the outcome model. A Monte Carlo simulation study comparing the augmented inverse propensity weighting estimator to a regression estimator, an inverse propensity weighting estimator, and a propensity score matching estimator found that the augmented inverse propensity weighting estimator had similar or lower mean square error compared to three other estimators in the case of misspecification of one of the propensity score or outcome models (Glynn & Quinn, 2010). Additionally, different standard error estimates for the treatment effect estimated with the augmented inverse propensity weighting method are available. Many of those standard error estimates such as a sandwich estimator (Lunceford & Davidian, 2004) and bootstrap standard error estimates (Imbens,

2004) can be implemented in the R package called “Causal GAM” (Glynn & Quinn, 2009). The augmented inverse propensity weighting estimator is not yet implemented for mediator effects, to my knowledge, but may be useful for mediation analysis

1.2.1.2 Sequential g-estimation

G-computation is a method to identify the controlled direct effect in the presence of post-treatment confounders (Robins, 1986). Post-treatment confounders in a mediation model are confounders of the M to Y relation that are influenced by the treatment; they can bias the direct effect (i.e., *c'*) estimate. An example of post-treatment confounders for the M to Y relationship may be the variable socio-economic status (SES) in a mediation chain where educational attainment influences unhealthy eating behavior which then influences blood pressure. In this example, SES may be influenced by the educational attainment and also influence both eating behavior and blood pressure. Another example of post-treatment confounders in mediation may be alliance with the therapist in an intervention program to treat depression.

The g-computation method attempts to estimate all potential values in a research design by using the estimated distribution of the measured confounders given values of X. The mean outcome is computed for each combination of values of the predictors in the outcome model to find the expected outcome values within levels of X, M, and C. Estimating the conditional outcome under each possible combination implies a solution to the missing data problem in determining causal effects.

The g-computation method can be difficult to implement when estimating the joint distribution of the confounders as a function of treatment in the case of many

confounders, since the method requires estimating all predicted potential outcomes. A simpler method is sequential g-estimation that also allows the researchers to directly model the effect of treatment on the outcome (Vansteelandt, 2009; Goetgeluk, Vanteelandt, & Goetghebeur, 2008; Joffe & Green, 2009). The sequential g-estimator is equivalent to g-computation method in the case of linear models.

The sequential g-estimator is implemented in two steps in which the first step removes the effect of the mediator from the outcome variable and in the second step the direct effect is estimated. First, the outcome is regressed on the treatment, mediator, and confounders using ordinary least squares regression to find the mediator's effect on the outcome (this is referred as the mediator model). Then, the mediator's effect is removed from the outcome by using the coefficient reflecting the effect of M on Y, ($Y - \beta_m M$). Next, this residual outcome is regressed on the treatment to find the remaining direct effect of X on Y (this is referred as the outcome model):

$$(11) \quad E(Y - \beta_m M \mid X) = \alpha_0 + \psi X$$

Note that the above equation for the residual outcome can also include the baseline confounders, but not post-treatment confounders. The standard error for the sequential g-estimator, ψ, may be biased since it does not account for the uncertainty in the estimation of the mediator's effect. Therefore, bootstrapping can be used for the estimation of the standard error.

Similar to the IPW method, the sequential g-estimation also has the following assumptions: (1) Stable Unit Treatment Value Assumption (SUTVA), and (2) all possible confounders for the X to Y, M to Y, and X to M relations are measured.

Doubly robust sequential g-estimation

Since the sequential g-estimation method fits two models in its estimation (by first estimating a mediator model and then an outcome model as described above), it may be prone to be biased by misspecification in either of these two models. A *doubly robust sequential g-estimation* method is suggested in the literature in which the estimated direct effect is robust to misspecifications in either the mediator or the outcome model. The method is expected to produce bias in the direct effect estimates when both parts of its estimation process are misspecified (Schafer & Kang, 2008). Doubly robust sequential g-estimation involves the following steps: in the first step, a propensity model for the mediator is fitted as in the IPW method; then, in the second step, the outcome regression is fitted using the propensity weights. Even though this method to estimate the controlled direct effects has been recommended, its performance has never been tested in simulation studies.

Chapter 2

CURRENT STUDY

The overall purpose of the proposed study is to evaluate how confounders affect estimation of direct and indirect effects and to evaluate several methods that may improve estimation of direct and indirect effects when confounders are present. The methods are compared in a simulation study.

Recent literature suggests various methods to deal with the assumption of no unmeasured confounders for the M to Y relation in mediation analysis. The methods differ in how adjustment is made for confounders. The inverse propensity weighting (IPW) method uses a propensity model in which mediator group membership is weighted conditional on the confounders that are measured before the mediator. In a mediation context, the IPW method achieves causal estimation of direct and indirect effects by regressing the outcome on the treatment and mediator, after weighting each individual by the inverse of his/her probability of the mediator status conditional on the treatment and confounders (i.e., mimicking randomization of the mediator using measured confounders). Another method, sequential g-estimation uses an outcome model where the direct effect is estimated after removing the association between the mediator and outcome.

Simulation studies show that the inverse propensity weighting approach produces roughly unbiased estimates of the indirect effects when all confounders are measured and included in the propensity model (Coffman & Zhong, 2012). Similarly, sequential g-estimation produces unbiased estimates of the direct effect in the case of including all

post-treatment confounders in the estimation process, whereas linear regression with adjustment does not (Loeys et al., 2013). Additionally, sequential-g estimation produces roughly unbiased direct effect estimates even as the association between the post-treatment confounder and the outcome increases. In contrast, the adjusted regression and IPW estimators get increasingly biased as the association between the post-treatment confounder and the outcome increases (Vansteelandt, 2009).

The current dissertation investigates the statistical performance of methods to detect direct and indirect effects using a large simulation study. Most of the recent literature on causal inference methods for indirect and direct effects in the presence of a mediator includes small simulation studies in a limited number of conditions (Lepage et al., 2012). The studies generally ignore effect sizes, statistical power, Type I error rates, and confidence interval coverage. This study aims to compare five methods in terms of bias, relative bias, mean square error, statistical power, Type I error rates, and confidence interval coverage. I explore how robust the methods are to violation of the no omitted variables assumption in model estimation and to the size of confounder effect in the case of various sample size conditions. Another contribution of the study is that it investigates a doubly robust g-estimation method that has been briefly suggested in the literature (Vansteelandt & Keiding, 2011), but has never been, to my knowledge, described in detail nor tested.

The five methods investigated in this study are: IPW, IPW with truncated weights, sequential g-estimation, doubly robust sequential g-estimation, and linear regression with adjustment. One of the reasons I chose to focus on these specific methods is that they can

accommodate continuous mediators. The methods are used to estimate the total indirect and direct effects in a single mediator model with two post-treatment confounders of the M to Y relation.

2.1 Research Questions and Predictions

Below are the specific research questions explored followed by predictions:

1. How robust are the methods to the violation of the no unmeasured confounders assumption? Does the doubly robust sequential g-estimation outperform the other methods investigated?

   It is predicted that the doubly robust method will outperform all the other methods except when there is violation of the no omitted confounders assumption in the estimation of both the mediator and outcome models.

2. Does the IPW method lead to biased estimates when the effect of the treatment and post-treatment confounders on the mediator are extreme (i.e., the *a*, *d*, and *f* paths are either small or large in Figure 2 below)?

   In the case of the IPW method, when the effects of the treatment and the post-treatment confounders on the mediator are small, the weights used for each subject become highly influential and may lead to inaccurate results. Conversely, if M has strong predictors, then weights may be extremely large or extremely small. Because of their ability to deal with extreme weights for some individuals, it is expected that the IPW with truncated weights and the sequential g-estimation methods will lead to less biased estimates, higher statistical power, and better coverage compared to the IPW method.

3. Does the bias in estimates increase as the effect of the treatment on the confounders increases (i.e., paths *g* and *h* in Figure 2)?

   The relative bias of the direct effect in the presence of a mediator occurs when the confounders for the M to Y path are influenced by the treatment. Thus, it is expected that the relative bias will increase as the effect of the treatment on the post-treatment confounders increases. I predict that sequential g-estimation will outperform the IPW method in the case of a large effect of the treatment on the confounders, since the sequential g-estimation aims to eliminate the M to Y path when estimating the direct effect of the treatment on the outcome.

4. Does a smaller effect of the mediator on the outcome (i.e., the *b* path in Figure 2) lead to less bias in the direct effect estimates under each method?

   It is expected that the relative bias in the direct effect estimates will decrease as the *b* path decreases.

5. Does the bias in estimates for the direct effect (i.e., the *c′* path in Figure 2) decrease as the effect size for the direct effect increases?

   It is expected that the relative bias in the direct effect and the total indirect effect estimates will decrease as the *c′* path increases.

6. Does the bias of the direct and indirect effects decrease as sample size increases? Do the predicted effects specified above differ by sample size?

   The bias in the IPW estimator is expected to increase as N decreases due to weight instability.

2.2 Method

2.2.1 Simulation overview

A Monte Carlo simulation study was conducted to examine the effect of confounder effect sizes and violation of assumption of no unmeasured confounders for the M to Y relation on the performance of five analysis methods (i.e., regression with adjustment, IPW, IPW with truncated weights, sequential g-estimation, and doubly robust sequential g-estimation) in a single mediator model with post-treatment confounder variables. There are two measured post-treatment confounder variables (C_1 and C_2) that are influenced by the treatment and that influence the mediator directly, and the outcome through a spurious relation induced by an unobserved confounder U (see figure 2 below). The model was generated for different sample sizes, with different effect sizes for the paths X to M, M to Y, X to C_1, X to C_2, C_1 to M, and C_2 to M. After the generation of the data, the five methods were used to estimate the direct and indirect effect estimates in the single mediator model. To assess the effect of violation of the assumption of no unmeasured confounders, two models are estimated using the five methods: (a) a two-confounders estimation of the model by including both post-treatment confounders (C_1 and C_2) in the estimation; (b) a one-confounder estimation of the model by including only the confounder C_1 in the estimation and omitting the second confounder C_2 from the estimation. The results are then evaluated by examining bias in the parameter estimates, confidence interval coverage, statistical power and Type I error rates.

The data were generated in SAS 9.3 with a total of 1000 replications per condition. The evaluation criteria measures, including bias, relative bias, and coverage, were also computed and analyzed in SAS.

2.2.2 Data generation and simulation conditions

The following regression equations are specified in SAS in order to generate the population parameters. Figure 2 shows the simulated model. Exogenous variables are generated using the SAS RANNOR function to produce normally distributed random variables. The independent variable X is simulated to be binary to represent a treatment status (0 = control, 1 = treatment group). All other variables are simulated to be continuous with normally distributed error terms. There is an unobserved confounder U in the simulated model so that there is only one path to be traced from X to Y for ease of interpretation.

(3.1) $M = a\,X + d\,C_1 + f\,C_2 + e_1$

(3.2) $Y = c'\,X + b\,M + t\,U + e_2$

(3.3) $C_1 = gX + k\,U + e_3$

(3.4) $C_2 = h\,X + n\,U + e_4$

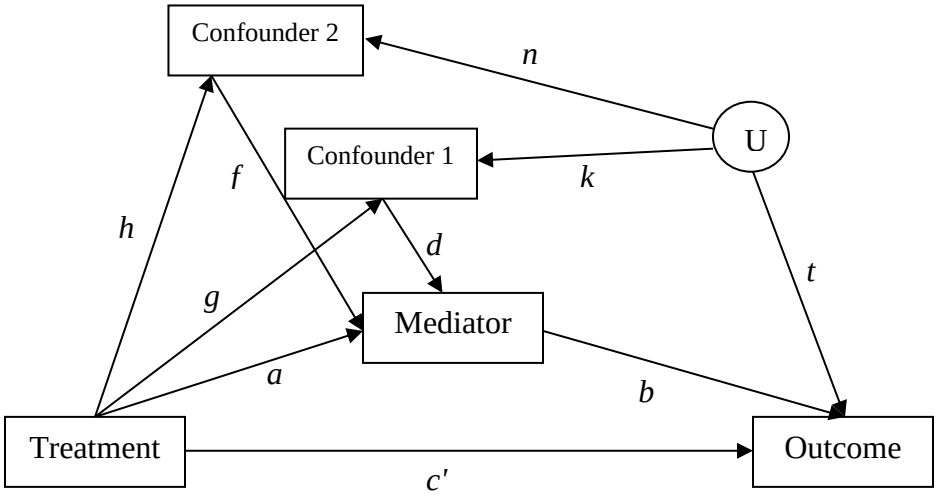


Figure 2. Generated model

The unstandardized regression parameters for the *b* and *c'* paths are varied as 0, .14, and .59. The effect of X on M (the *a* path), and the effects of C_1 on M, and C_2 on M (the paths *d* and *f*) are varied as .14, .39 and .59. The effect of X on C_1 and C_2 (the paths *g* and *h*) are varied as .14 and .59. The effects of C_1 on M and C_2 on M are set to be equal (i.e., the *d* and *f* paths), as wells as the effect of X on C_1 and C_2 (i.e., the *g* and *h* paths). The effects of the unobserved confounder U on C_1, C_2, and Y (i.e., the paths *k*, *n*, and *t*) are set equal to 1.0. Mediation effect sizes were chosen following MacKinnon et al. (2002, 2004). Example correlation matrices for the study variables for two simulation conditions are given in Appendix A.

Sample sizes were simulated to be 250, 500 or 1000 to represent the sample sizes commonly found in social sciences, and to explore the case of a larger sample size of 1000.

To summarize, a 3 (N) × 3 (X → M) × 3 (M → Y) × 3 (X → Y) × 3 (C_1 → M and C_2 → M) × 2 (X → C_1 and X → C_2) factorial design yielding a total of 486 conditions was used in the simulation study. A total of 1,000 replications of each condition were conducted so 486,000 data sets were analyzed in the simulation study.

2.2.3 Model estimation

The five methods of interest were applied to the generated data sets using two model estimation specifications: (a) The two-confounders estimation model including both of the population model confounders C_1 and C_2 in the estimation, and (b) the one-confounder estimation model including only the confounder C_1 in the estimation (i.e., omitting the second confounder C_2 from the estimation). The case of one-confounder estimation model allows a test the robustness of methods to the violation of no omitted confounders assumption. As an exception, the one-confounder model for the doubly robust sequential g-estimation had three types of estimation where C_2 was included in one part of the model but not another as will be described below.

1) *Linear regression with adjustment.* The linear outcome regression equation for the two-confounders estimation model includes X, M, and both C_1 and C_2 as predictors; and the one-confounder estimation model only includes X, M, and C_1 as predictors.
2) *Inverse propensity score weighting*. The two-confounders estimation for the propensity score model to create the weights for the mediator is specified by including X, C_1 and C_2 in estimating the denominator model. The one-confounder estimation was performed by only including X and C_1 in estimating the

denominator model. In both cases, the weighted outcome model only includes X and M as predictors.

3) *Inverse propensity score weighting with truncated weights.* The model specification is the same as the method (2) described above; yet weights are truncated at the 1st and the 99th percentile of the weight distribution as in Cole & Hérnan (2008). The truncation is conducted to avoid weighting certain observations too little or too much.

4) *Sequential g-estimation.* The first step in which the outcome is regressed on the treatment, mediator, and confounders using ordinary least squares regression (referred to as the Q-model) is specified by including X, M, C_1, and C_2 as predictors in the two-confounders estimation model. Only X, M, and C_1 are included in the one-confounder estimation model.

5) *Doubly robust sequential g-estimation*. In the first step, the propensity model for the mediator is fitted as in method (2), the IPW method. Then, in the second step, the outcome regression is fitted using the propensity weights. For the doubly robust sequential g-estimation method, three one-confounder estimation models are fitted: (a) by omitting confounder C_2 in only the mediator propensity model, (b) by omitting confounder C_2 in only the outcome model, (c) by omitting confounder C_2 in both the mediator propensity and outcome models. This allows for testing if the doubly robust method fails when both parts or one part of the estimated model omit the confounder C_2 (Schafer & Kang, 2008).

2.2.4 Data analysis, outcome measures and evaluation criteria

The *c'* path is the measure of the controlled direct effect of the treatment on the outcome. The total indirect effect is computed as subtracting the direct effect from the total effect (*c*-*c'*). For all methods, the percentile bootstrap with 1000 replications is used to calculate the 95% confidence intervals. In order to compute the percentile confidence intervals, the 1000 replicated coefficient estimates were saved and sorted in descending order to determine values at the 2.5 and 97.5 percentiles.

Bias of the *c'* path and the indirect effect *c*-*c'* are defined as:

$$\mathrm{Bias}\left(\hat{\theta}_c\right) = \mathrm{R}^{-1}\sum_{r=1}^{R}(\hat{\theta}_{rc} - \theta_c)$$

where R refers to the total number of replications, θ_c refers to the true value of the coefficients, and $\hat{\theta}_{rc}$ refers to the parameter estimate for replications r in condition c. *Relative bias* is defined as the ratio of bias to the true value. An estimator was considered as acceptable in terms of bias if the absolute value of relative bias was less than .10 (Flora & Curran, 2004).

$$\mathrm{RBias}\left(\hat{\theta}_c\right) = \mathrm{R}^{-1}\sum_{r=1}^{R}\frac{(\hat{\theta}_{rc} - \theta_c)}{\theta_c}$$

Additionally, the mean square residual is defined as follows. Smaller values of MSE indicate higher stability of the parameter estimates.

$$\mathrm{MSE} = \mathrm{R}^{-1} \sum_{\mathrm{r}=1}^{\mathrm{R}} (\hat{\theta}_{\mathrm{rc}} - \bar{\hat{\theta}})^2$$

Type I error rates and *statistical power* are calculated using 5% level of significance, as it is the most common value used in social sciences. Type I error rate indicates the error of rejecting the null hypothesis that the direct effect (or indirect effect) is equal to zero (i.e., *c'*=0) in the simulation condition where actually the direct effect is equal to zero. In other words, the proportion of replications in which a significant effect is incorrectly detected represents Type I error rates when true values are set equal to zero. The proportion of replications in which a significant effect is correctly detected represents statistical power when true values are not equal to zero. The values were evaluated against the nominal .80 criterion for statistical power (Cohen, 1988). The liberal criterion of [.025, .075] was used to evaluate Type I error rates (Bradley, 1978). The statistical power was interpreted only in the conditions in which the Type I error rates were acceptable. *Coverage* is computed as the proportion of 95% confidence intervals that contains the true value. Coverage rates greater than 90% was evaluated as satisfactory (Collins, Schafer, & Kam, 2001).

Chapter 3

RESULTS

There are two quantities that are examined in this research, the direct effect *c'* and the indirect effect *c*-*c'*. The first set of tables and figures describes results for the direct effect, and the second set of tables and figures describes results for the indirect effect. There are two types of estimated models -- a two-confounders estimation model and a one-confounder estimation model. The two-confounders estimation model has all information on all variables for the different estimation techniques. The one-confounder estimation model has all information on all variables except for one confounding variable C_2. There are three types of simulation outcome variables. One group of outcome variables investigates the characteristics of estimates in terms of bias, relative bias, and mean square error. The focus of the bias related results was the magnitude of bias rather than the sign of the bias, i.e., both larger negative bias and larger positive bias were described as increasing magnitude of bias. The second group of simulation outcome variables is based on the tests of statistical significance, Type 1 error, and statistical power. A third outcome variable is confidence interval coverage for each estimator. The results for three sample sizes are presented—N=250, 500, and 1000. Because the bias related results were comparable across sample sizes, complete tabled results for N=500 are shown and only mean square error results for all sample sizes are shown in Figures.

For the two-confounders estimation models, there are five estimation methods: (1) Ordinary Least Squares Regression, (2) Inverse probability weighting, (3) Inverse probability weighting with truncated weights, (4) Sequential g-estimation, and (5) Doubly

Robust Sequential g-estimation. For one-confounder estimation models that omit Confounder 2 from the estimation, seven estimation methods are studied: (1) Ordinary Least Squares Regression, (2) Inverse probability weighting, (3) Inverse probability weighting with truncated weights, (4) Sequential g-estimation, (5) Robust Sequential g-estimation that does not include information on Confounder 2 for the prediction of the mediator, (6) Robust Sequential g-estimation that does not include information on Confounder 2 for the prediction of the outcome variable, and (7) Robust Sequential g-estimation that does not include information on Confounder 2 for the prediction of both the mediator and outcome variables.

3.1 The direct effect, *c'*

3.1.1 Accuracy of point estimates

Tables 1 and 2 provide information about the robustness of methods to the violation of unmeasured confounders for the M to Y relation assumption. These tables address the research question, "How robust are the methods to violation of the assumption of unmeasured confounders of the M to Y relation?" As confounders of the M to Y relation are likely in most mediation studies these results are especially relevant. In Table 1, bias, relative bias, and mean square error of estimates of effects for two-confounders estimation models are presented across different effect sizes for the relation between post-treatment confounders and the mediator, i.e., paths *d* and *f*. Table 2 presents the results for one-confounder estimation models where the confounder C_2 was omitted from the analyses. This model with C_2 omitted from the analysis corresponds to the

common case in mediation studies in which no measure of a confounder is available but a confounder may affect the analysis.

For the two-confounders estimation model results in Table 1, bias, relative bias and MSE increase as the effect of the post-treatment confounder on the mediator increases for all methods except for linear regression and sequential g-estimation. IPW had acceptable relative bias when confounder effect size was .14. The truncated IPW method results showed that weight trimming did not improve the performance of the IPW method in terms of bias. This finding may be due to the trimming rule used. Although trimming may be used to optimize propensity score weights, the optimal level of trimming may be difficult to determine and may not contribute to or have adverse effects in the estimation. Sequential g-estimation was unbiased across different effect sizes of the confounder and the direct path c'. The doubly robust sequential g-estimation method had unacceptable relative bias as the confounder effect size increased. This finding may be expected considering that the doubly robust g-estimation method uses IPW in its estimation process, and IPW method did not perform well when confounder effect sizes increased. Linear regression had unacceptable relative bias across conditions, and its performance was not influenced by the size of the confounder effect size. This finding can be explained by the fact that the confounders were post-treatment. As can be seen in Table 5, the bias of the linear regression estimates increases drastically as the effect of the treatment on the confounders (paths *g* and *h*) increases.

For one-confounder estimation models in Table 2, bias, relative bias and MSE increase as the effect of the post-treatment confounder on the mediator increases for all

methods. All methods have unacceptable relative bias rates, except for one condition for the sequential g-estimation method. The sequential g-estimation method has satisfactory bias performance when the direct effect and confounder effect sizes are .14. The results also show that doubly robust sequential g-estimation performs the worst when both the mediator and the outcome estimation models violate the no omitted variables assumption. As seen in Figures 3 and 4, the pattern of results is similar across different sample sizes.

Table 1

Direct effect c' bias, relative bias, and mean square error by post-treatment confounders effect size (N=500) - two-confounders estimation model

		c' effect size								
		0			.14			.59		
		Post treatment confounder effect size (paths *d* and *f*)								
Method		.14	.39	.59	.14	.39	.59	.14	.39	.59
Regression	Bias	-.242	-.243	-.243	-.243	-.244	-.244	-.244	-.243	-.243
	Rel.Bias	na	na	na	-1.736	-1.740	-1.739	-.413	-.411	-.412
	MSE	.093	.093	.093	.093	.093	.093	.094	.093	.093
IPW	Bias	-.002	-.053	-.118	-.004	-.051	-.114	-.004	-.052	-.120
	Rel.Bias	na	na	na	-.027	-.361	-.815	-.006	-.087	-.203
	MSE	.021	.072	.126	.021	.071	.126	.021	.069	.126
IPW trunc.	Bias	-.032	-.125	-.189	-.033	-.125	-.187	-.033	-.124	.189
	Rel.Bias	na	na	na	-.39	-.892	-1.339	-.057	-.210	-.321
	MSE	.020	.042	.067	.020	.041	.067	.020	.042	.069
Seq. *g*-est.	Bias	.001	-.001	.000	-.001	-.001	.001	-.001	.002	-.001
	Rel.Bias	na	na	na	-.006	-.007	.004	-.001	.003	-.002
	MSE	.017	.018	.018	.017	.017	.018	.017	.017	.018
R. seq. *g*-est.	Bias	-.002	-.053	-.116	-.004	-.053	-.117	-.004	-.051	-.117
	Rel.Bias	na	na	na	-.025	-.375	-.834	-.006	-.086	-.198
	MSE	.017	.025	.038	.017	.024	.038	.017	.024	.038

Note: Regression: linear regression adjusting for covariates; IPW: inverse propensity weighting; IPW trunc.: IPW with truncated weights; Seq. g-est: sequential g-estimation; R. seq. g-est.: doubly robust sequential g-estimation.

Table 2

Direct effect c' bias, relative bias, and mean square error by post-treatment confounders effect size (N=500) - one-confounder estimation models

		c' effect size								
		0			.14			.59		
		Post treatment confounder effect size (paths *d* and *f*)								
Method		.14	.39	.59	.14	.39	.59	.14	.39	.59
Regression	Bias	-.209	-.253	-.275	-.210	-.253	-.276	-.210	-.252	-.276
	Rel.Bias	na	na	na	-1.498	-1.808	-1.970	-.356	-.427	-.467
	MSE	.069	.091	.105	.070	.091	.105	.070	.091	.106
IPW	Bias	-.033	-.125	-.187	-.035	-.123	-.186	-.035	-.124	-.191
	Rel.Bias	na	na	na	-.249	-.879	-1.328	-.059	-.210	-.323
	MSE	.021	.057	.094	.020	.056	.095	.021	.056	.095
IPW trunc.	Bias	-.054	-.169	-.229	-.056	-.169	-.228	-.056	-.167	-.229
	Rel.Bias	na	na	na	-.399	-.207	-1.627	-.095	-.284	-.389
	MSE	.022	.053	.082	.022	.053	.082	.022	.053	.082
Seq. *g*-est.	Bias	-.031	-.106	-.156	-.033	-.105	-.155	-.033	-.103	-.157
	Rel.Bias	na	na	na	-.236	-.753	-1.110	.056	-.175	-.266
	MSE	.018	.029	.043	.018	.028	.043	.018	.028	.044
R. seq. *g*-est. I	Bias	-.033	-.125	-.188	-.035	-.124	-.188	-.035	-.123	-.189
	Rel.Bias	na	na	na	-.246	-.889	-1.340	-.059	-.208	-.320
	MSE	.018	.036	.059	.018	.035	.059	.018	.036	.059
R. seq. *g*-est. II	Bias	-.184	-.218	-.252	-.185	-.218	-.252	-.186	-.217	-.252
	Rel.Bias	na	na	na	-1.319	-1.557	-1.801	-.315	-.368	-.427
	MSE	.059	.076	.094	.060	.076	.094	.060	.076	.094
R. seq. *g*-est. III	Bias	-.210	-.267	-.295	-.210	-.267	-.295	-.211	-.266	-.295
	Rel.Bias	na	na	na	-1.502	-1.904	-2.104	-.358	-.450	-.501
	MSE	.070	.101	.120	.070	.100	.119	.071	.101	.119

Note: Regression: linear regression adjusting for covariates; IPW: inverse propensity weighting; IPW trunc.: IPW with truncated weights; Seq. g-est: sequential g-estimation; R. seq. g-est. I: doubly robust sequential g-estimation with one-confounder estimation mediator model; R. seq. g-est. II: doubly robust sequential g-estimation with one-confounder estimation outcome model; R. seq. g-est. III: doubly robust sequential g-estimation with one-confounder estimation mediator and outcome models.

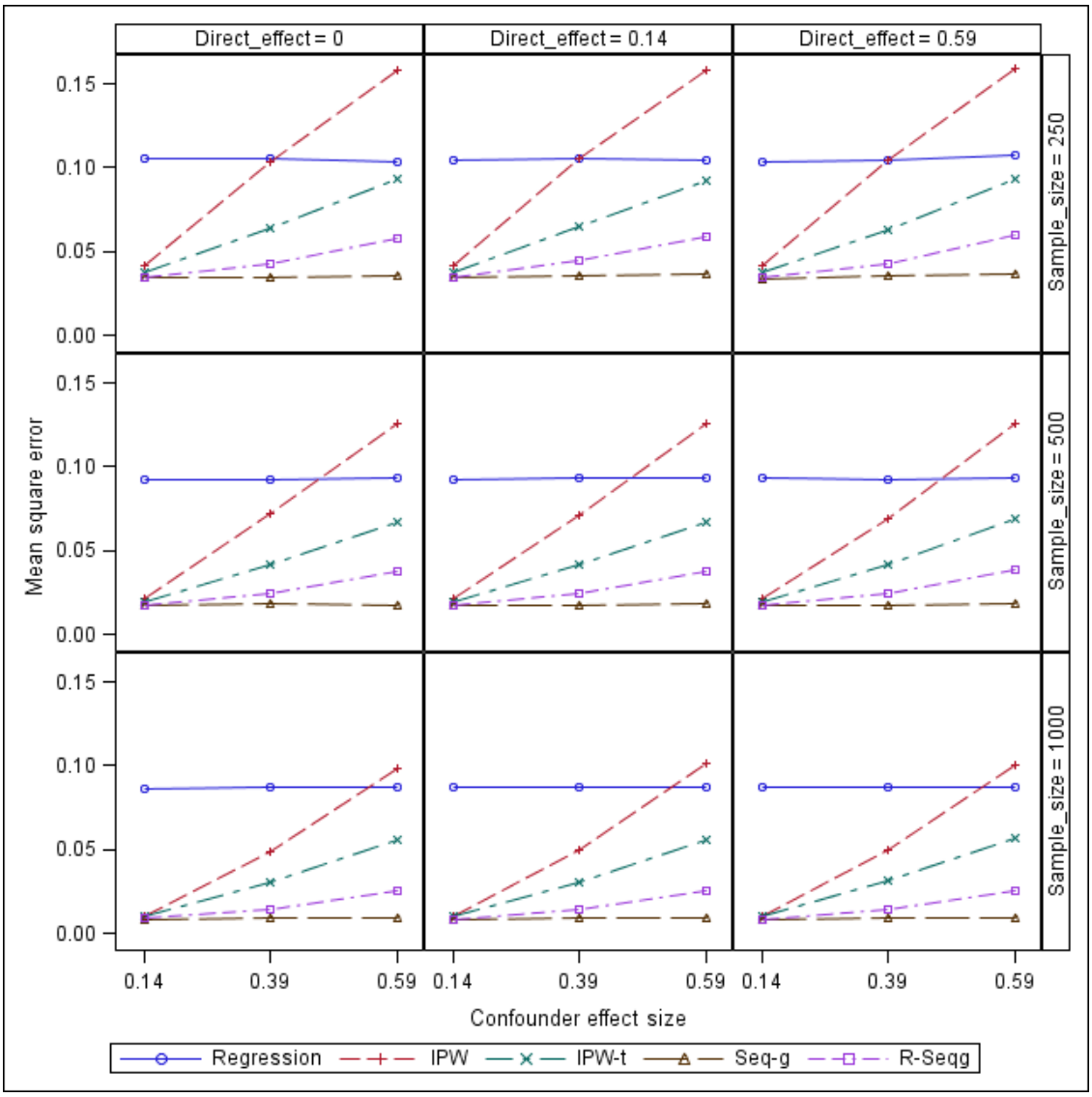


Figure 3. Direct effect c' mean square error by post-treatment confounders effect size and sample size - two-confounders estimation model

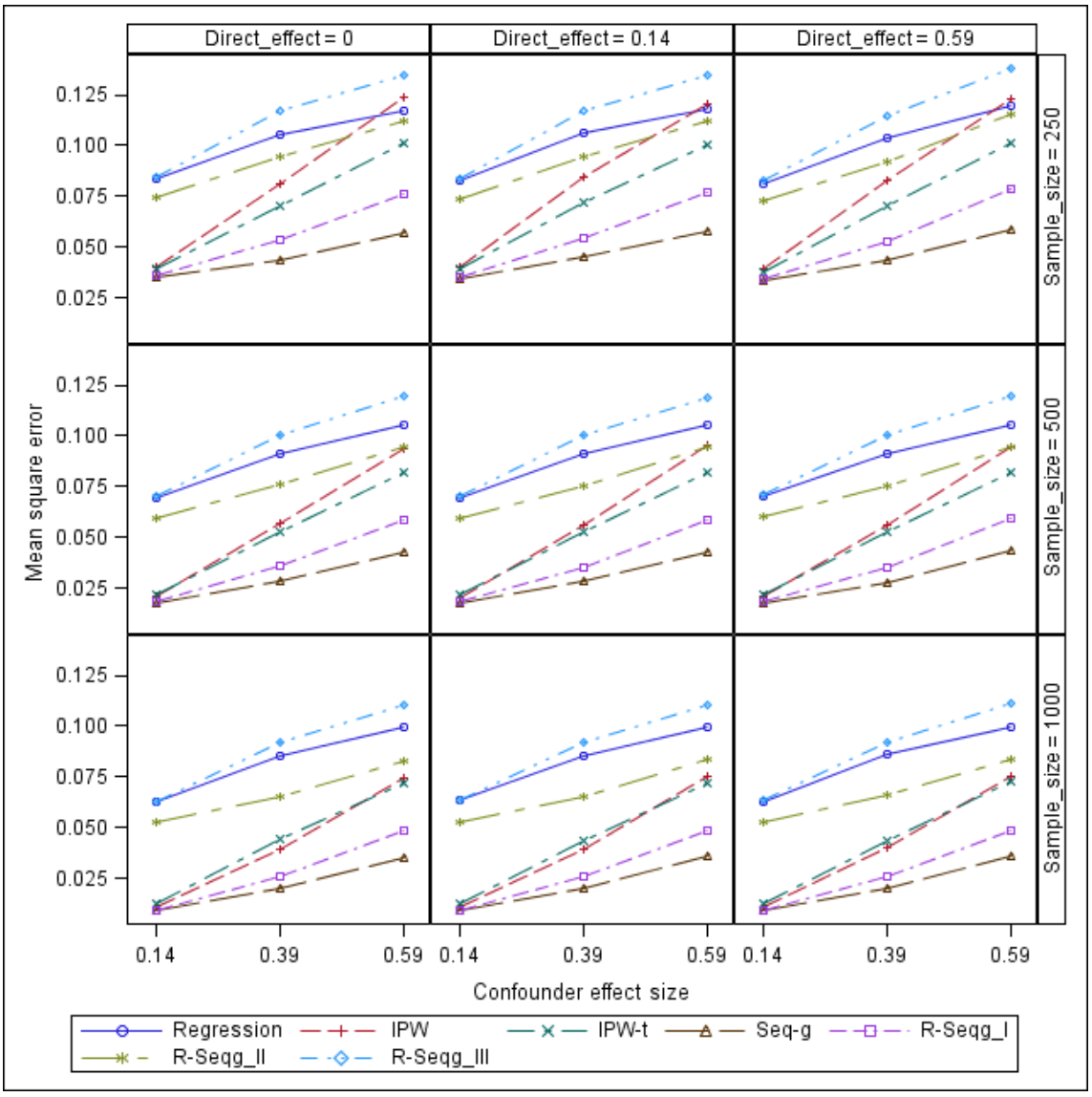


Figure 4. Direct effect c' mean square error by post-treatment confounders effect size and sample size - one-confounder estimation model

Tables 3 and 4 investigate the research question, "Does the IPW method lead to biased estimates when the effect of the treatment (the *a* path) and post-treatment covariates (the *d* and *f* paths) on the mediator are extreme?" In the case of the IPW method, when the effects of the treatment and the post-treatment confounders on the mediator are too small, the weights used for each subject become highly influential and may lead to inaccurate results. Conversely, if M has strong predictors, then weights may be extremely large or extremely small. Because of its ability to deal with extreme weights for some individuals, IPW with truncated weights method was expected to lead to less biased estimates compared to the IPW method.

Tables 3 and 4 show that the bias in both the IPW and IPW with truncated weights estimates of the *c'* path increases as the *a, d,* and , *f* paths increase in both two-confounders and one-confounder estimation models. For the one-confounder estimation models reported in Table 4, relative bias of both IPW and IPW-truncated methods are unacceptable in nearly all conditions. For the two-confounders estimation models reported in Table 3, IPW has unacceptable relative bias when the *a* path and confounders effect sizes are larger than .14. Additionally, the truncated IPW method does not perform better than the conventional IPW method, contrary to my prediction. This can be explained by the weight trimming method used in this study (weights were trimmed at the 1st and 99th percentile of the weight distribution). Other simulation studies showing that the truncated IPW method did not perform better than the conventional IPW method in many situations suggested that since the optimal trimming is difficult to determine, researchers should be focusing on the proper specification of the propensity scores rather

than weight trimming as a post-hoc method (Lee, Lessler, & Stuart, 2011). Figures 5 and 6 show that the pattern of results does not differ substantially across different sample sizes.

Table 3

Direct effect c' bias, relative bias, and mean square error by a, d, and f paths effect size (N=500) - two-confounders estimation models

		c' effect size								
		0			.14			.59		
		a=d=f paths effect size								
Method		.14	.39	.59	.14	.39	.59	.14	.39	.59
Regression	Bias	-.242	-.242	-.241	-.241	-.246	-.244	-.243	-.241	-.244
	Rel.Bias	na	na	na	-1.723	-1.755	-1.744	-.411	-.409	-.413
	MSE	.092	.093	.093	.091	.094	.094	.093	.092	.094
IPW	Bias	.000	-.051	-.146	-.000	-.052	-.140	-.001	-.050	-.152
	Rel.Bias	na	na	na	-.002	-.371	-1	-.001	-.085	-.258
	MSE	.020	.074	.137	.020	.069	.138	.020	.068	..136
IPW trunc.	Bias	-.015	-.128	-.236	-.015	-.130	-.237	-.015	-.126	-.242
	Rel.Bias	na	na	na	-.104	-.931	-1.690	-.025	-.213	-.410
	MSE	.018	.042	.086	.018	.041	.087	.018	.041	.091
Seq. *g*-est.	Bias	.002	.000	.004	.002	-.004	.001	.002	.003	-.003
	Rel.Bias	na	na	na	.014	-.028	-.008	.003	.005	-..005
	MSE	.016	.018	.018	.016	.017	.019	.016	.017	.020
R. seq. *g*-est.	Bias	.000	-.053	-.143	.000	-.056	-.145	.000	-.050	-.150
	Rel.Bias	na	na	na	.003	-.402	-1.036	.001	-.086	-.254
	MSE	.016	.025	.048	.016	.024	.050	.016	.027	.050

Note: Regression: linear regression adjusting for covariates; IPW: inverse propensity weighting; IPW trunc.: IPW with truncated weights; Seq. g-est: sequential g-estimation; R. seq. g-est.: doubly robust sequential g-estimation.

Table 4

Direct effect c' bias, relative bias, and mean square error by a, d, and f paths effect size (N=500) - one-confounder estimation models

		c' effect size								
		0			.14			.59		
		a=d=f paths effect size								
Method		.14	.39	.59	.14	.39	.59	.14	.39	.59
Regression	Bias	-.192	-.254	-.315	-.191	-.258	-.318	-.193	-.253	-.319
	Rel.Bias	na	na	na	-1.366	-1.841	-2.271	-.327	-.429	-.540
	MSE	.062	.092	.128	.062	.093	.130	.062	.091	.130
IPW	Bias	-.015	-.125	-.239	-.016	-.127	-.234	-.016	-.124	-.243
	Rel.Bias	na	na	na	-.117	-.907	-1.668	-.028	-.211	-.411
	MSE	.018	.057	.115	.018	.054	.118	.018	.055	.117
IPW trunc.	Bias	-.027	-.172	-.286	-.026	-.174	-.289	-.027	-.170	-.291
	Rel.Bias	na	na	na	-0.187	-1.246	-2.064	-.045	-.289	-.493
	MSE	.018	.052	.109	.018	.052	.112	.018	.051	.113
Seq. *g*-est.	Bias	-.015	-.106	-.195	-.014	-.111	-.197	-.015	-.104	-.200
	Rel.Bias	na	na	na	-.102	-.790	-1.407	-.025	-.177	-.340
	MSE	.016	.028	.056	.016	.028	.057	.016	.027	.059
R. seq. *g*-est. I	Bias	-.016	-.126	-.236	-.015	-.130	-.236	-.015	-.125	-.241
	Rel.Bias	na	na	na	-.109	-.926	-1.687	-.026	-.211	-.409
	MSE	.016	.035	.079	.016	.035	.081	.016	.035	.082
R. seq. *g*-est. II	Bias	-.182	-.218	-.280	-.182	-.221	-.283	-.183	-.217	-.284
	Rel.Bias	na	na	na	-1.297	-1.581	-2.019	-.310	-.368	-.482
	MSE	.058	.076	.111	.058	.076	.113	.059	.075	.113
R. seq. *g*-est. III	Bias	-.192	-.268	-.344	-.192	-.272	-.345	.193	-.267	-.347
	Rel.Bias	na	na	na	-1.371	-1.940	-2.461	-.327	-.453	-.588
	MSE	.062	.100	.150	.062	.101	.152	.063	.100	.153

Note: Regression: linear regression adjusting for covariates; IPW: inverse propensity weighting; IPW trunc.: IPW with truncated weights; Seq. g-est: sequential g-estimation; R. seq. g-est. I:

doubly robust sequential g-estimation with one-confounder estimation mediator model; R. seq. g-est. II: doubly robust sequential g-estimation with one-confounder estimation outcome model; R. seq. g-est. III: doubly robust sequential g-estimation with one-confounder estimation mediator and outcome models.

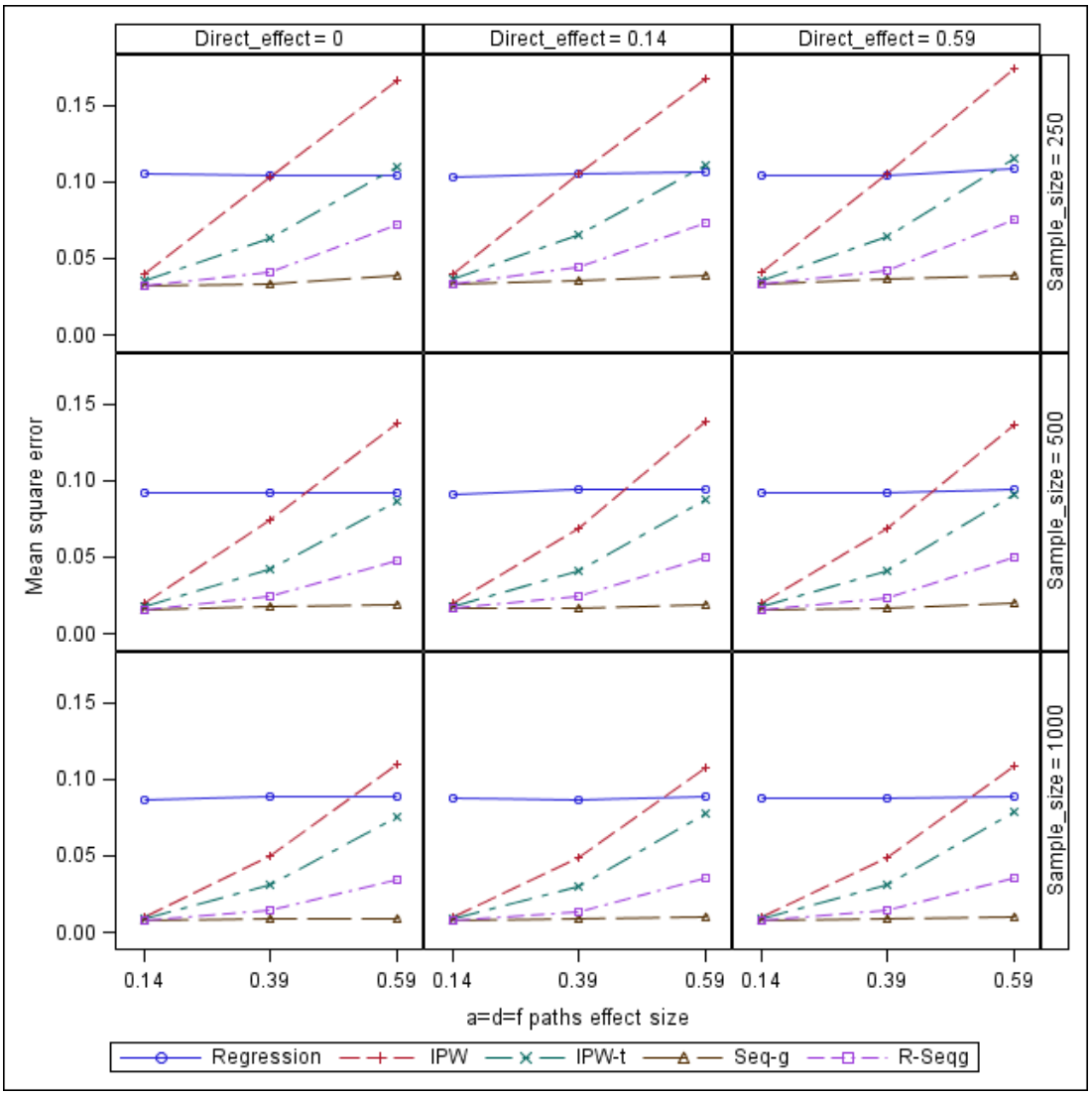


Figure 5. Direct effect c' and mean square error by a, d, and f paths effect size and sample size - two-confounders estimation model

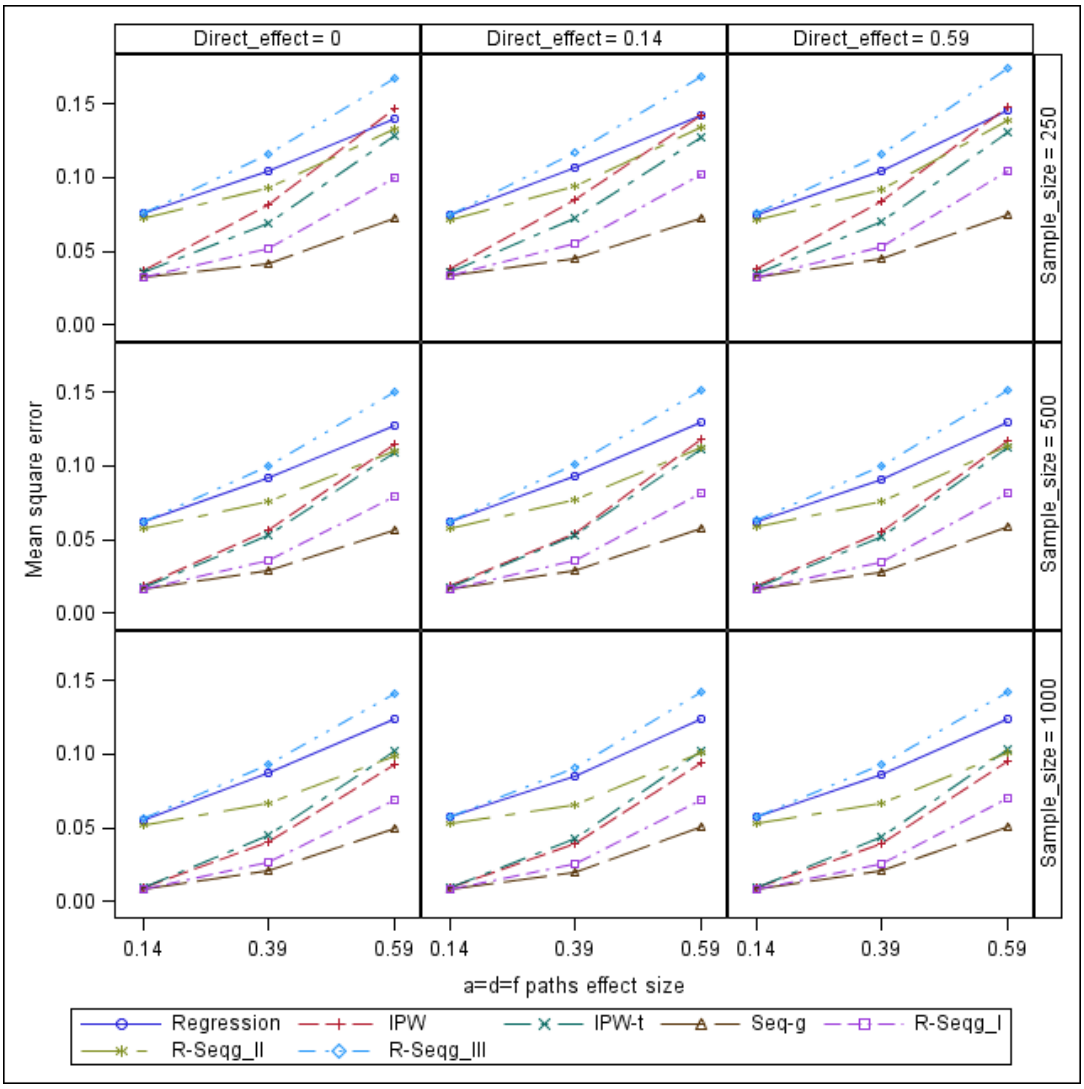


Figure 6. Direct effect c' mean square error by a, d, and f paths effect size and sample size - one-confounder estimation model

Tables 5 and 6 investigate the research question, "Does the bias in estimates for the methods increase as the effect of the treatment on the confounders increases, i.e., as paths *g* and *h* increase?" For the two-confounders estimation models in Table 5, results show that the bias, relative bias and MSE increase as the effect of the treatment on confounders increases for all methods except for sequential g-estimation. The sequential g-estimation is unbiased across conditions. Linear regression has unacceptable relative bias across conditions, and its bias increases drastically as the association between the treatment and post-treatment confounders increase. IPW has acceptable relative bias only when the direct effect is .59 and the relation between the treatment and confounders is .14. This indicates that neither IPW is a good choice of method of analysis in the case of post-treatment confounders. Additionally, weight trimming does not improve the performance of the IPW method. The doubly robust sequential g-estimation performs similar to the IPW method since it is using IPW in its estimation process.

In the case of one-confounder estimation model, as can be seen in Table 6, none of the methods has acceptable relative bias across conditions. As expected, the doubly robust g-estimation method performs the worst when the confounder C_2 is omitted from both the mediator and outcome models. Figures 7 and 8 show that the pattern of results does not differ importantly across different sample sizes.

Table 5

Direct effect c' bias, relative bias, and mean square error by g and h paths effect size (N=500) - two-confounders estimation models

		c' effect size					
		0		.14		.59	
		g=h paths effect size					
Method		.14	.59	.14	.59	.14	.59
Regression	Bias	-.093	-.392	-.094	-.393	-.094	-.393
	Rel.Bias	na	na	-.673	-2.804	-.159	-.666
	MSE	.020	.166	.020	.166	.020	.166
IPW	Bias	-.041	-.074	-.040	-.072	-.040	-.076
	Rel.Bias	na	na	-.289	-.513	-.068	-.129
	MSE	.067	.079	.066	.079	.066	.078
IPW trunc.	Bias	-.083	-.148	-.083	-.148	-.082	-.150
	Rel.Bias	na	na	-.590	-1.057	-.139	-.254
	MSE	.032	.055	.031	.054	.032	.055
Seq. *g*-est.	Bias	-.001	-.001	-.001	.000	.000	-.000
	Rel.Bias	na	na	-.009	.003	.000	-.000
	MSE	.017	.018	.017	.018	.017	.018
R. seq. *g*-est.	Bias	-.040	-.073	-.041	-.074	-.040	-.074
	Rel.Bias	na	na	-.291	-.532	-.067	-.126
	MSE	.021	.033	.020	.032	.021	.032

Note: Regression: linear regression adjusting for covariates; IPW: inverse propensity weighting; IPW trunc.: IPW with truncated weights; Seq. g-est: sequential g-estimation; R. seq. g-est.: doubly robust sequential g-estimation.

Table 6

Direct effect c' bias, relative bias, and mean square error by g and h paths effect size (N=500) - one-confounder estimation models

		c' effect size					
		0		.14		.59	
		g=h paths effect size					
Method		.14	.59	.14	.59	.14	.59
Regression	Bias	-.127	-.364	-.127	-.365	-.127	-.365
	Rel.Bias	na	na	-.911	-2.607	-.215	-.619
	MSE	.030	.148	.030	.148	.030	.148
IPW	Bias	-.082	-.148	-.083	-.147	-.083	-.150
	Rel.Bias	na	na	-.590	-1.048	-.140	-.254
	MSE	.044	.070	.045	.069	.045	.069
IPW trunc.	Bias	-.109	-.193	-.109	-.192	-.108	-.194
	Rel.Bias	na	na	-.781	-1.375	-.183	-.328
	MSE	.035	.070	.035	.069	.035	.070
Seq. *g*-est.	Bias	-.070	-.125	-.071	-.125	-.070	-.126
	Rel.Bias	na	na	-.506	-.893	-.118	-.213
	MSE	.022	.038	.022	.037	.022	.038
R. seq. *g*-est. I	Bias	-.083	-.148	-.083	-.148	-.082	-.149
	Rel.Bias	na	na	-.591	-1.059	-.139	-.252
	MSE	.026	.050	.025	.050	.025	.050
R. seq. *g*-est. II	Bias	-.101	-.334	-.102	-.334	-.102	-.335
	Rel.Bias	na	na	-.730	-2.388	-.172	-.567
	MSE	.025	.128	.025	.128	.025	.128
R. seq. *g*-est. III	Bias	-.137	-.378	-.137	-.377	-.137	-.378
	Rel.Bias	na	na	-.978	-2.695	-.232	-.641
	MSE	.034	.160	.034	.159	.034	.160

Note: Regression: linear regression adjusting for covariates; IPW: inverse propensity weighting; IPW trunc.: IPW with truncated weights; Seq. g-est: sequential g-estimation; R. seq. g-est. I: doubly robust sequential g-estimation with one-confounder estimation mediator model; R. seq. g-est. II: doubly robust sequential g-estimation with one-confounder estimation outcome model; R. seq. g-est. III: doubly robust sequential g-estimation with one-confounder estimation mediator and outcome models.

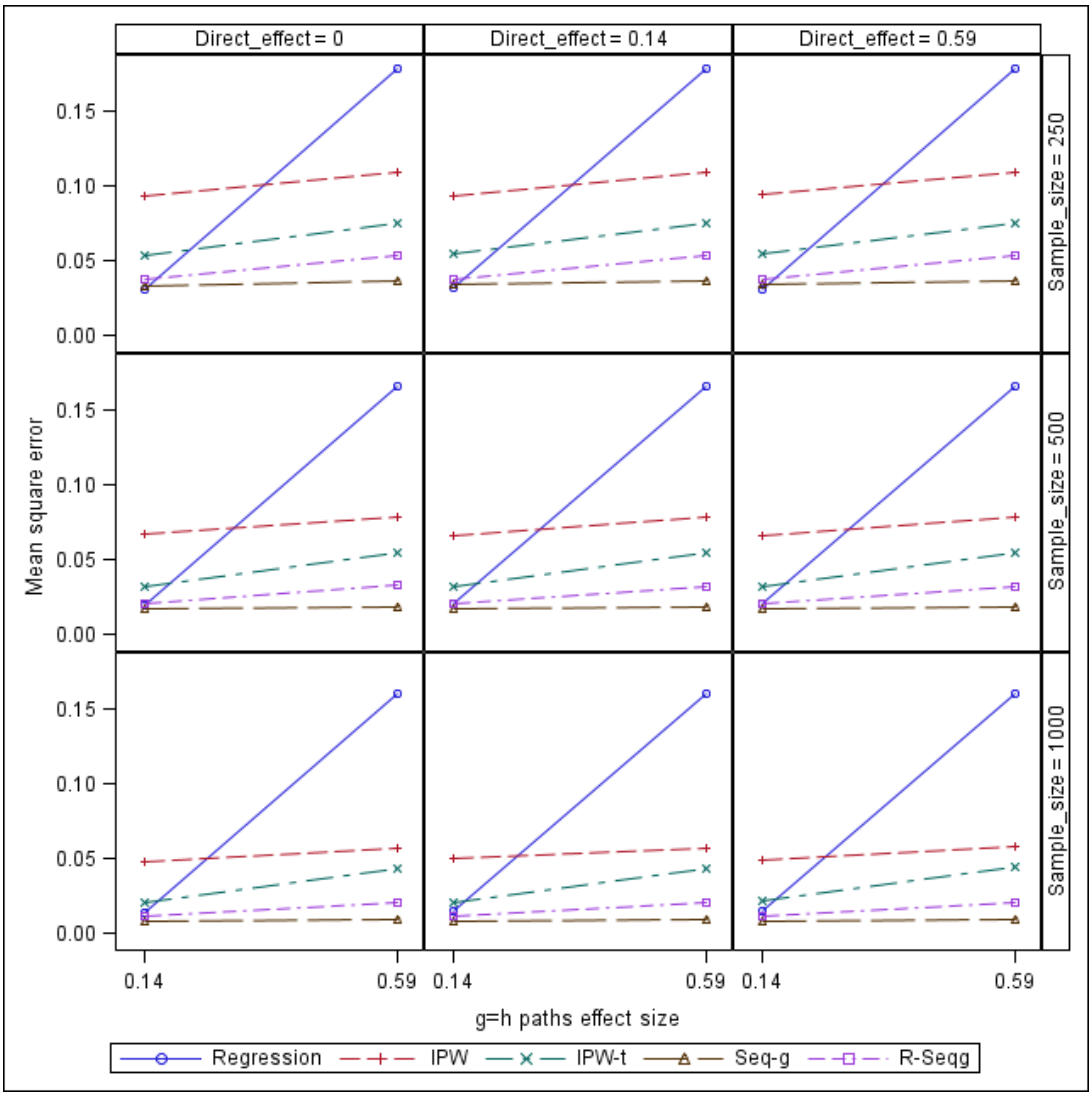


Figure 7. Direct effect *c'* mean square error by *g* and *h* paths effect size and sample size - two-confounders estimation model

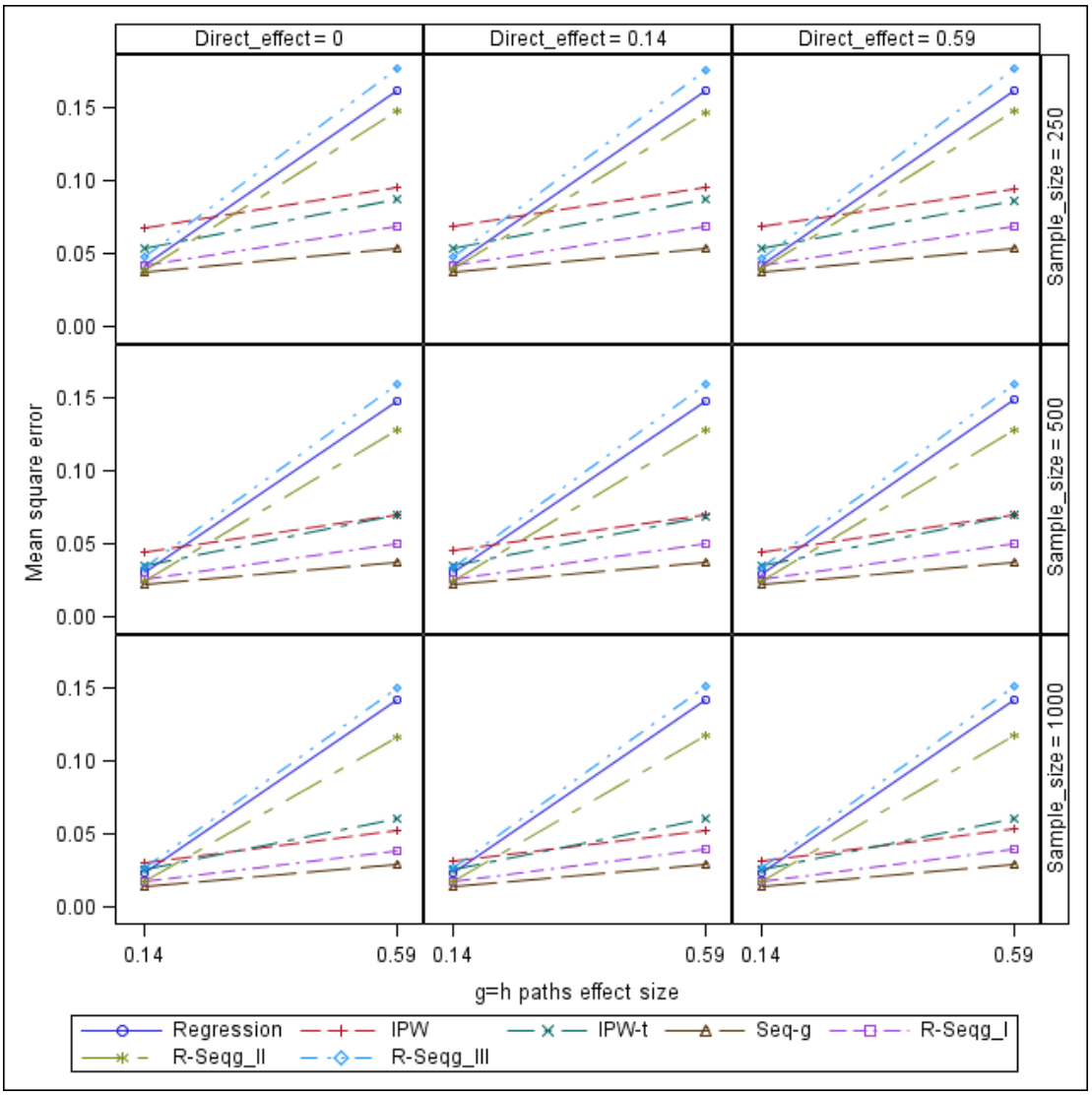


Figure 8. Direct effect *c'* mean square error by *g* and *h* paths effect size and sample size - one-confounder estimation model

Tables 7 and 8 investigate the research question, “Does a small effect of the mediator on the outcome (i.e., the *b* path) lead to less bias in the controlled direct estimates under each method?” Since the *b* path is biased by the post-treatment confounders that influence both the mediator and the outcome, it is possible that the bias in mediation parameter estimates may be influenced by the effect size of the *b* path.

Tables 7 and 8 show that for both two-confounders and one-confounder estimation models, bias in the c' parameter estimate was not influenced by the effect size of the *b* path. In the case of two-confounders estimation model, linear regression did not have acceptable relative bias across conditions and had the highest bias compared to other methods. IPW had acceptable relative bias only when the direct effect was .59. IPW truncated did not have acceptable relative bias for neither condition. Sequential g-estimation was unbiased across conditions. Yet, the doubly robust g-estimation was only unbiased when c' was equal to .59. This finding again indicates that the doubly robust sequential g-estimation followed the pattern of IPW method since its estimation partly uses IPW. In the case of one-confounder estimation models, none of the methods had acceptable relative bias. Figures 9 and 10 show that results were similar across sample size conditions.

Table 7

Direct effect c' bias, relative bias, and mean square error by b path effect size (N=500) - two-confounders estimation models

		c' effect size								
		0			.14			.59		
		b path effect size								
Method		0	.14	.59	0	.14	.59	0	.14	.59
Regression	Bias	-.242	-.243	-.244	-.244	-.243	-.243	-.243	-.243	-.244
	Rel.Bias	na	na	na	-1.743	-1.738	-1.735	-.412	-.411	-.413
	MSE	.092	.093	.094	.093	.093	.093	.094	.093	.093
IPW	Bias	-.053	-.059	-.060	-.057	-.057	-.054	-.060	-.058	-.058
	Rel.Bias	na	na	na	-.410	-.400	-.386	-.101	-.097	-.098
	MSE	.072	.074	.073	.073	.074	.071	.072	.073	.071
IPW trunc.	Bias	-.114	-.115	-.118	-.116	-.116	-.114	-.116	-.114	-.116
	Rel.Bias	na	na	na	-.831	-.826	-.813	-.197	-.194	.197
	MSE	.042	.043	.044	.043	.043	.042	.044	.043	.043
Seq. *g*-est.	Bias	.002	.000	-.002	-.002	.000	.000	-.000	.000	-.000
	Rel.Bias	na	na		-.012	.000	.003	-.000	.001	-.000
	MSE	.017	.018	.017	.017	.017	.017	.018	.017	.017
R. seq. *g*-est.	Bias	-.055	-.056	-.059	-.059	-.057	-.056	-.057	-.056	-.057
	Rel.Bias	na	na	na	-.424	-.410	-.400	-.097	-.095	-.097
	MSE	.026	.027	.027	.027	.027	.026	.027	.026	.026

Note: Regression: linear regression adjusting for covariates; IPW: inverse propensity weighting; IPW trunc.: IPW with truncated weights; Seq. g-est: sequential g-estimation; R. seq. g-est.: doubly robust sequential g-estimation.

Table 8

Direct effect c' bias, relative bias, and mean square error by b path effect size (N=500) - one-confounder estimation models

		c' effect size								
		0			.14			.59		
		b path effect size								
Method		0	.14	.59	0	.14	.59	0	.14	.59
Regression	Bias	-.244	-.245	-.247	-.247	-.246	-.246	-.246	-.246	-.247
	Rel.Bias	na	na	na	- 1.764	- 1.758	- 1.754	-.417	-.416	-.418
	MSE	.088	.089	.090	.089	.089	.089	.089	.089	.089
IPW	Bias	-.114	-.113	-.118	-.117	-.116	-.111	-.116	-.117	-.117
	Rel.Bias	na	na	na	-.838	-.825	-.794	-.196	-.198	-.198
	MSE	.056	.057	.058	.058	.057	.056	.057	.059	.057
IPW trunc.	Bias	-.149	-.150	-.153	-.152	-.152	-.149	-.151	-.150	-.151
	Rel.Bias	na	na	na	- 1.087	- 1.084	- 1.063	-.255	-.255	-.257
	MSE	.052	.052	.053	.052	.052	.051	.053	.052	.052
Seq. *g*-est.	Bias	-.096	-.097	-.100	-.099	-.098	-.097	-.097	-.097	-.098
	Rel.Bias	na	na	na	-.709	-.697	-.692	-.165	-.165	-.166
	MSE	.029	.030	.030	.030	.030	.029	.030	.030	.030
R. seq. *g*-est. I	Bias	-.114	-.114	-.118	-.117	-.115	-.114	-.115	-.115	-.116
	Rel.Bias	na	na	na	-.837	-.824	-.813	-.195	.195	-.197
	MSE	.037	.038	.038	.038	.037	.037	.038	.037	.038
R. seq. *g*-est. II	Bias	-.217	-.217	-.220	-.220	-.218	-.217	-.219	-.217	-.219
	Rel.Bias	na	na	na	- 1.568	- 1.559	- 1.550	-.370	-.368	-.371
	MSE	.076	.076	.077	.077	.077	.076	.077	.076	.077
R. seq. *g*-est. III	Bias	-.256	-.257	-.259	-.258	-.257	-.256	-.257	-.257	-.258
	Rel.Bias	na	na	na	- 1.845	- 1.837	- 1.828	-.436	-.435	-.438
	MSE	.096	.096	.098	.097	.097	.096	.097	.097	.097

Note: Regression: linear regression adjusting for covariates; IPW: inverse propensity weighting; IPW trunc.: IPW with truncated weights; Seq. g-est: sequential g-estimation; R. seq. g-est. I: doubly robust sequential g-estimation with one-confounder estimation mediator model; R. seq. g-est. II: doubly robust sequential g-estimation with one-confounder estimation outcome model; R. seq. g-est. III: doubly robust sequential g-estimation with one-confounder estimation mediator and outcome models.

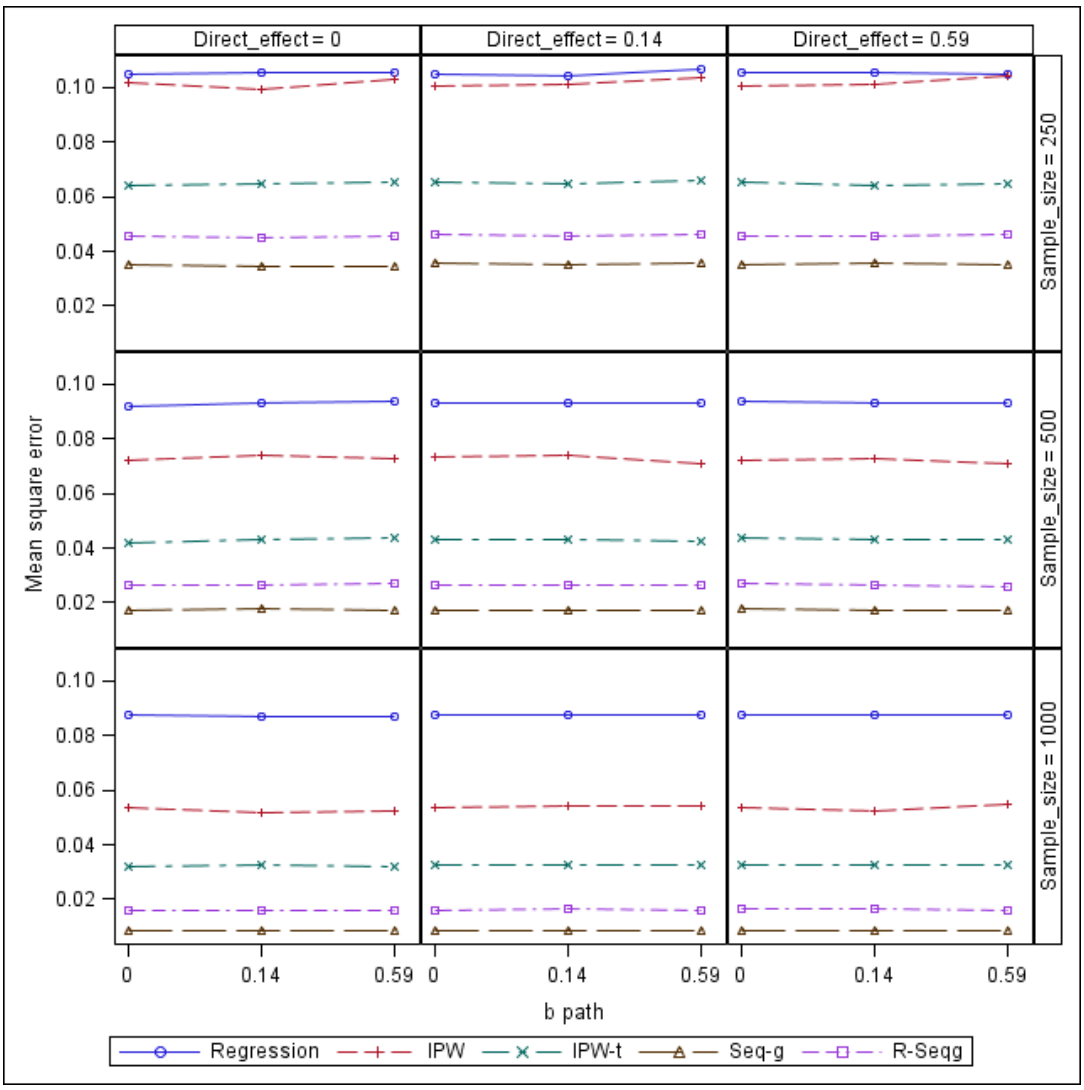


Figure 9. Direct effect *c'* mean square error by *b* path effect size and sample size - two-confounders estimation model

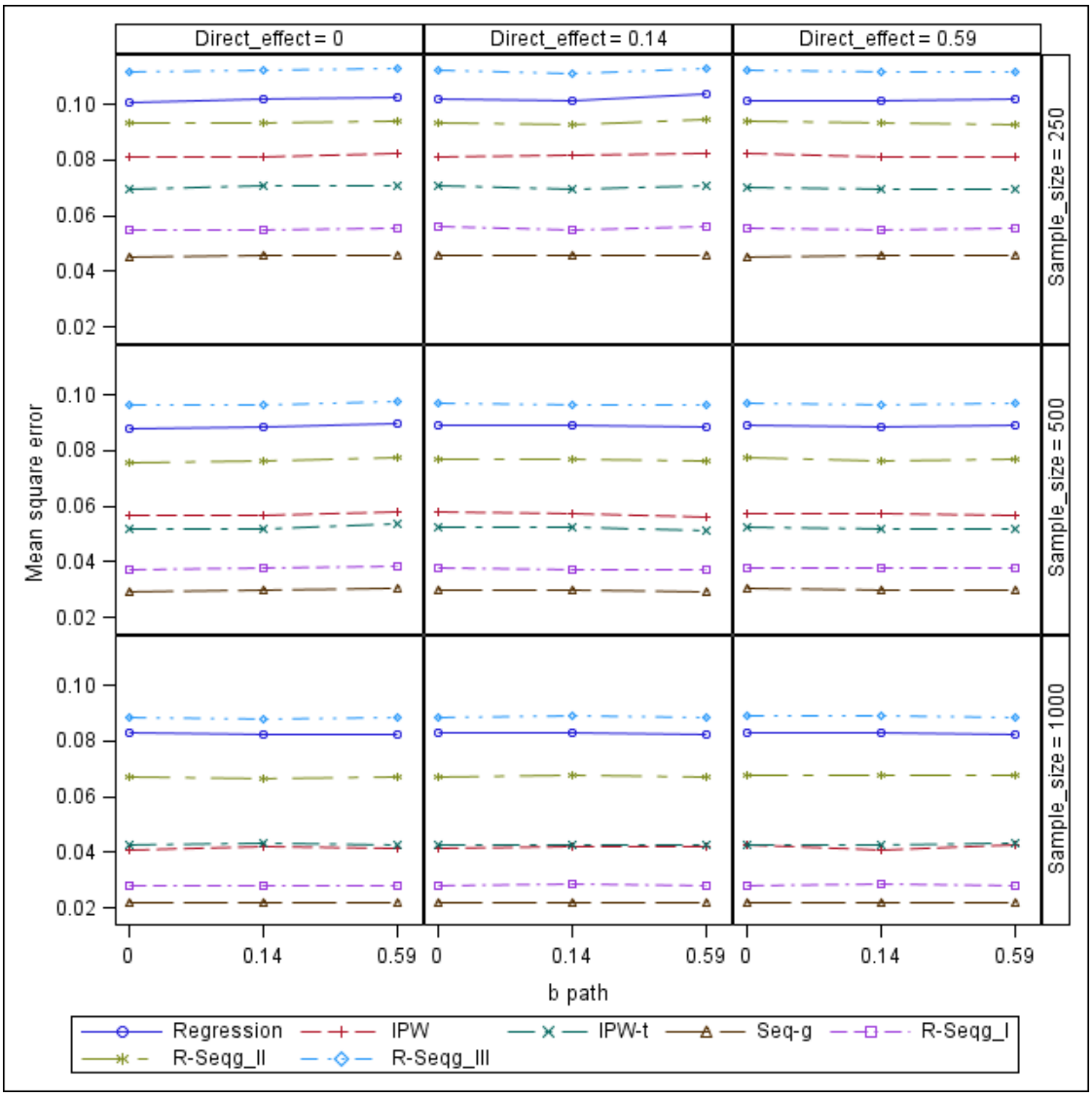


Figure 10. Direct effect *c'* mean square error by *b* path effect size and sample size - one-confounder estimation model

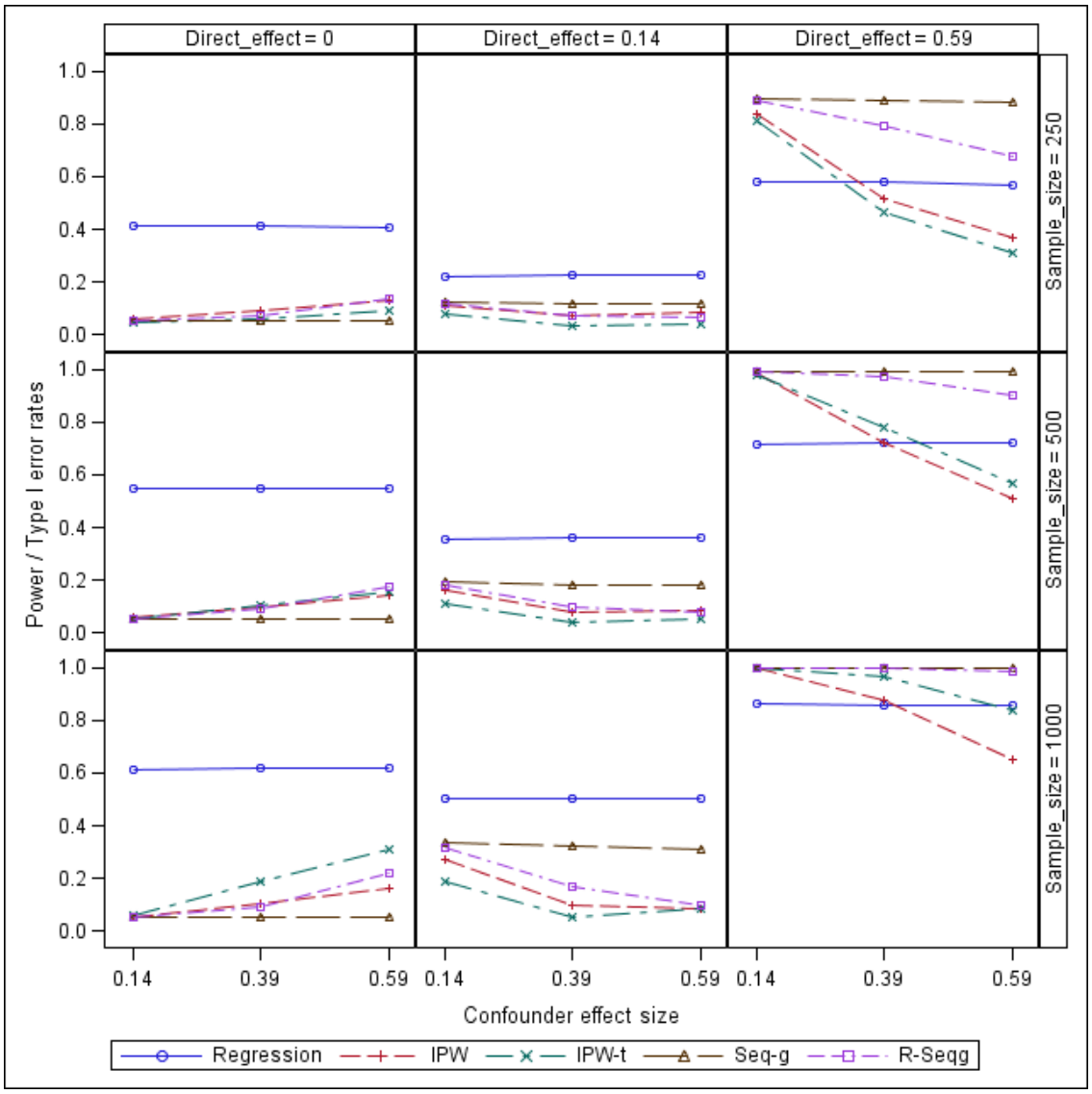
3.1.2 Statistical power and Type I error rates

The liberal criterion of [.025, .075] was used to evaluate Type I error rates (Bradley, 1978). Statistical power values were evaluated against the nominal .80 criterion (Cohen, 1988). The statistical power was interpreted only for the conditions in which the Type I error rates were acceptable.

In the case of two-confounders estimation models in Figure 11, Type I error rates for the linear regression were out of bounds across confounder effect size conditions and thus the statistical power results for that method were not interpreted. For the IPW methods, as the confounder effect size increased, Type I error rates increased and power decreased. Yet, the IPW methods had acceptable Type I error rates only when the confounder effect size was .14, and had statistical power greater than .80 for that condition only when the direct effect was equal to .59. The sequential g-estimation had good Type I error rates across conditions with a nominal value around .05 and good statistical power when the direct effect was .59. Yet, the doubly robust sequential g-estimation only had acceptable Type I error rates when the confounder effect size was .14 and good statistical power for that condition when the direct effect was equal to .59, again following the pattern of the IPW method.

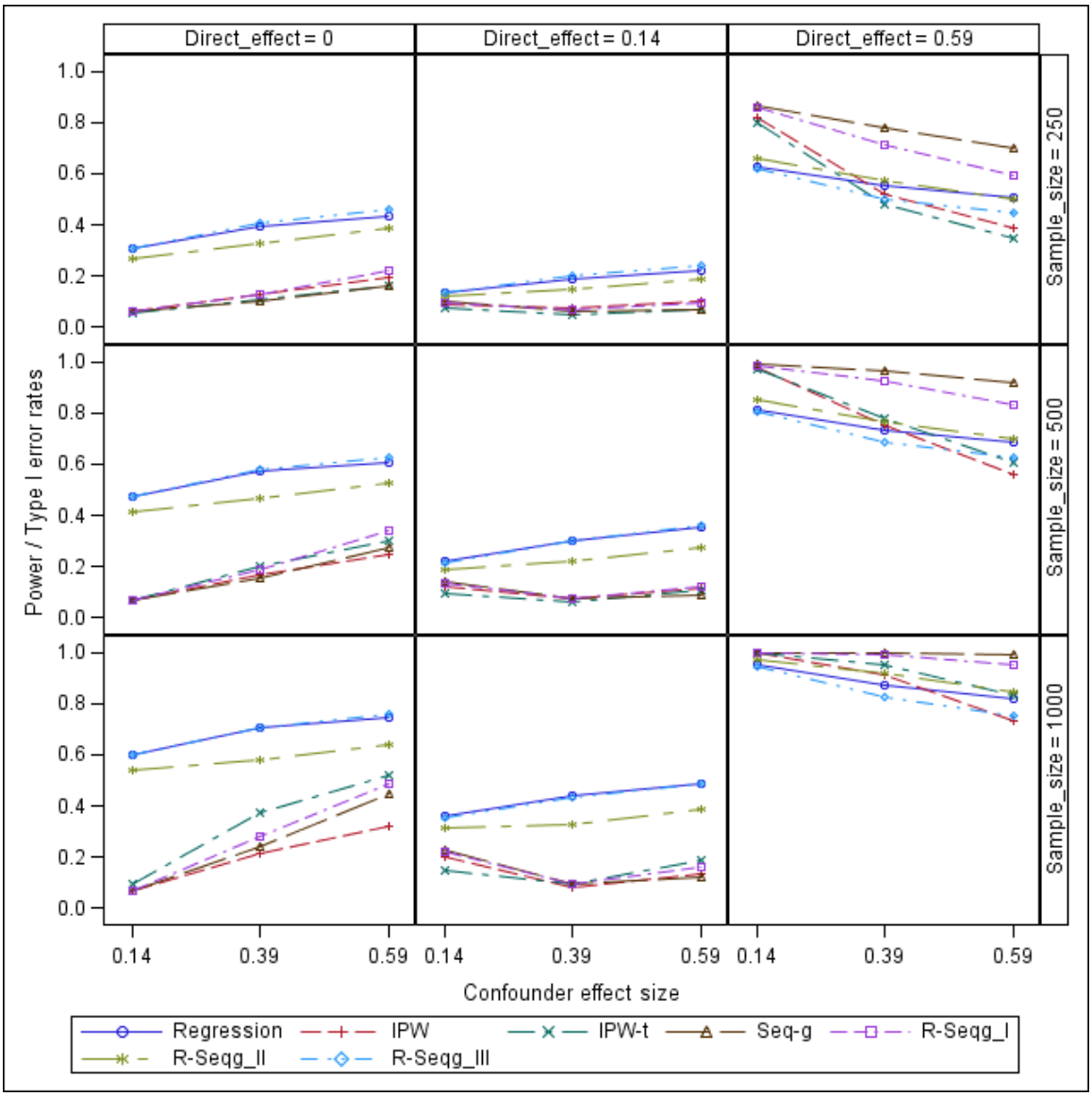
In the case of one-confounder estimation models in Figure 12, Type I error rates were out of bounds for the linear regression, doubly robust g-estimation with the omitted C_2 in the outcome model, and doubly robust g-estimation with the omitted C_2 in both the mediator and outcome models; thus the power results were not interpreted for these methods. Type I error rates were acceptable for the IPW methods, and the sequential g-

estimation, and the doubly robust sequential g-estimation when the confounder effect size was .14 and for that condition they reached power greater than .80 only when the direct effect was equal to .59.



Note1: Values represent Type I error rates when *c'*=0.
Note2: Regression: linear regression adjusting for covariates; IPW: inverse propensity weighting; IPW-t.: IPW with truncated weights; Seq-g: sequential g-estimation; R-Seqg: robust sequential g-estimation.

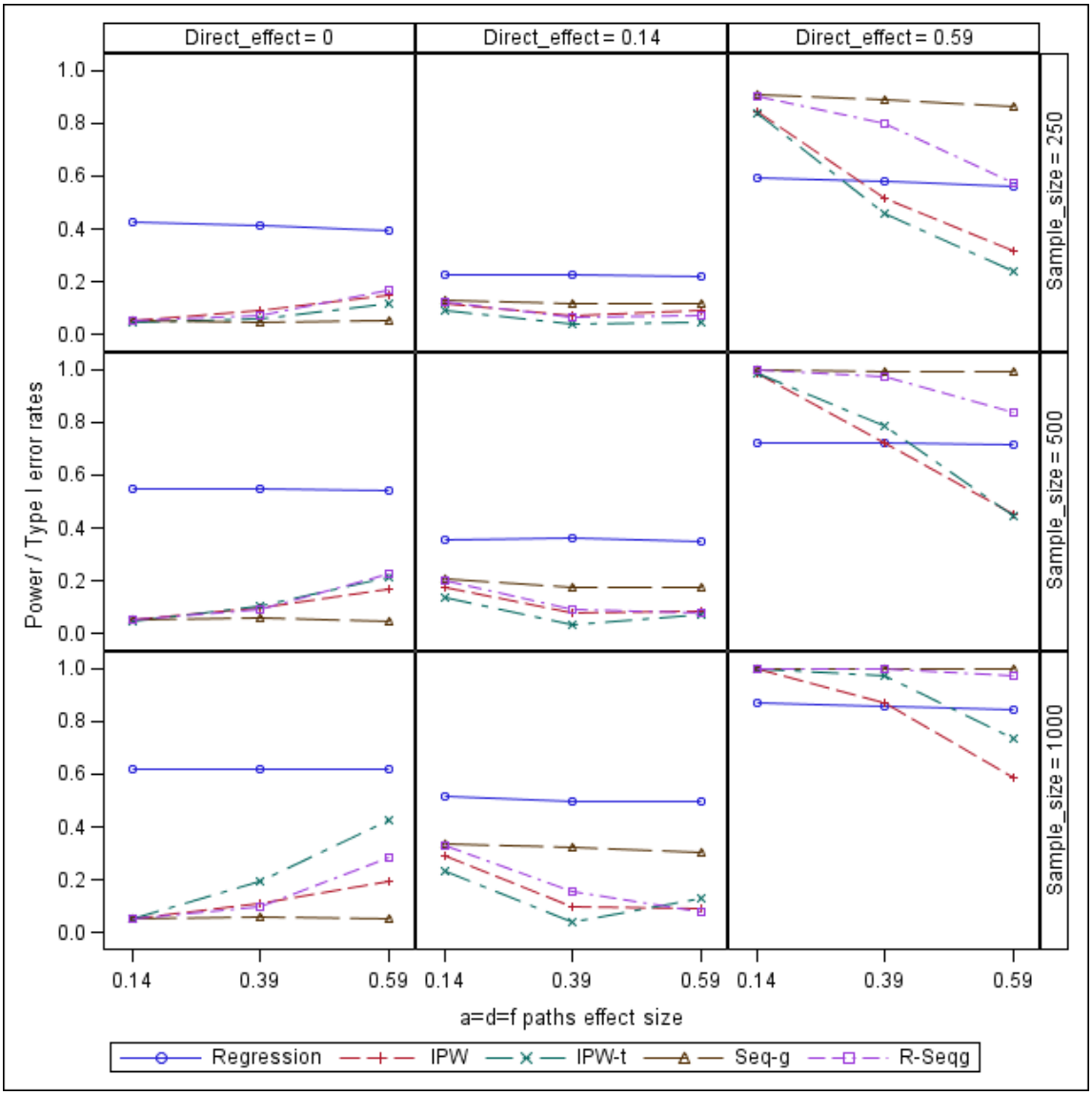
Figure 11. Direct effect *c'* power and Type I error rates by post-treatment confounders effect size and sample size - two-confounders estimation model



Note1: Values represent Type I error rates when *c'*=0.
Note2: Regression: linear regression adjusting for covariates; IPW: inverse propensity weighting; IPW-t.: IPW with truncated weights; Seq-g: sequential g-estimation; R-Seqg_I: doubly robust sequential g-estimation with one-confounder estimation mediator model; R-Seqg_II: doubly robust sequential g-estimation with one-confounder estimation outcome model; R-Seqg_III: doubly robust sequential g-estimation with one-confounder estimation mediator and outcome models.

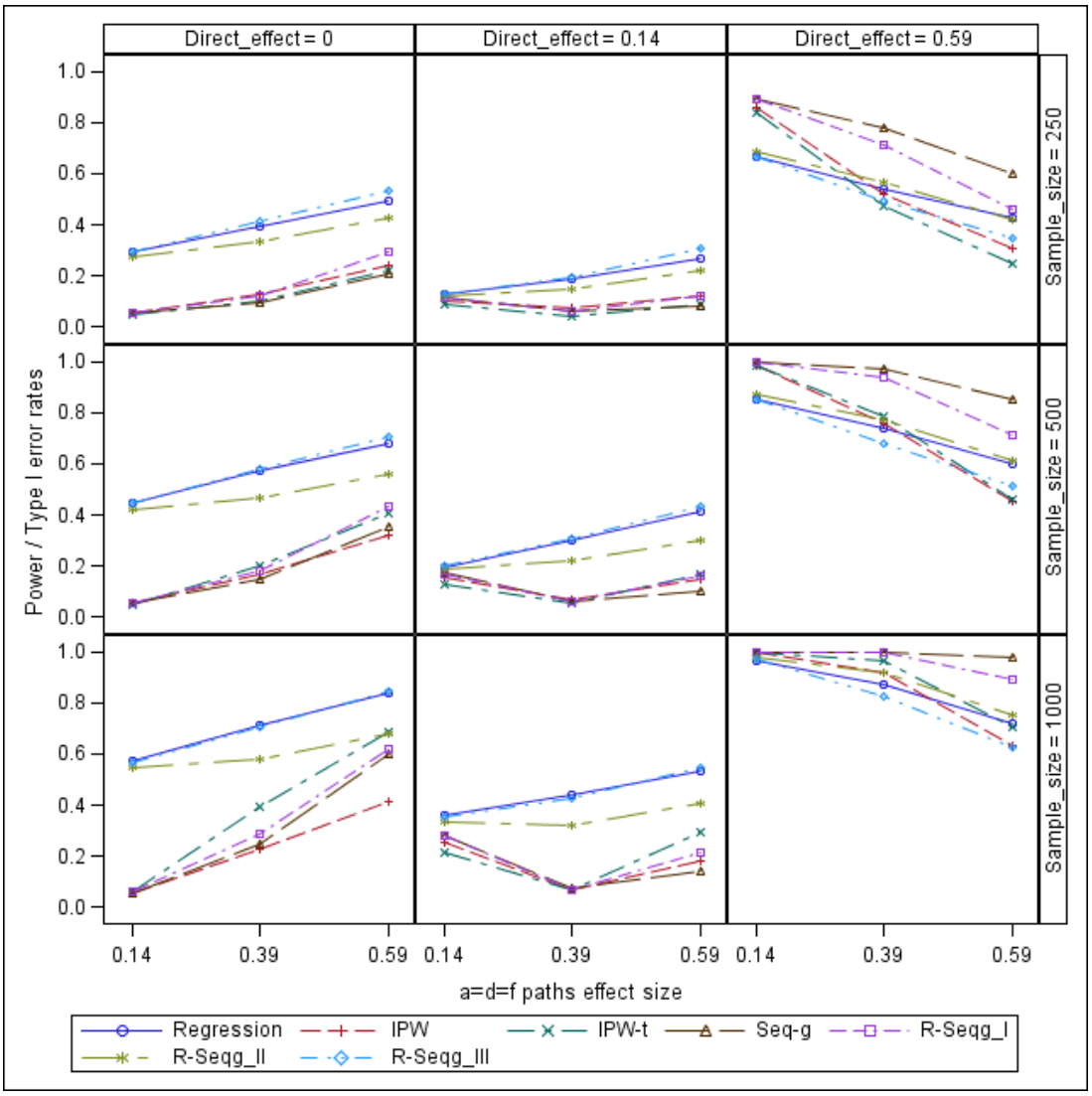
Figure 12. Direct effect *c'* power and Type I error rates by post-treatment confounders effect size and sample size - one-confounder estimation model

Figures 13 and 14 investigate whether the IPW method performs better when the *a* path and confounder effect sizes are not extreme and if the truncated IPW method has a better performance than the conventional IPW method. In the case of both two-confounders and one-confounder estimation models, the IPW method had acceptable Type I error rates when the *a* path and confounder effect sizes were .14. When the *a* path and confounder effect sizes were equal to .14, IPW had power greater than .80 only when the direct effect was equal to .59. Additionally, weight trimming did not contribute to the Type I error rate and power performance of the IPW method.



Note1: Values represent Type I error rates when *c'*=0.
Note2: Regression: linear regression adjusting for covariates; IPW: inverse propensity weighting; IPW-t.: IPW with truncated weights; Seq-g: sequential g-estimation; R-Seqg: robust sequential g-estimation.

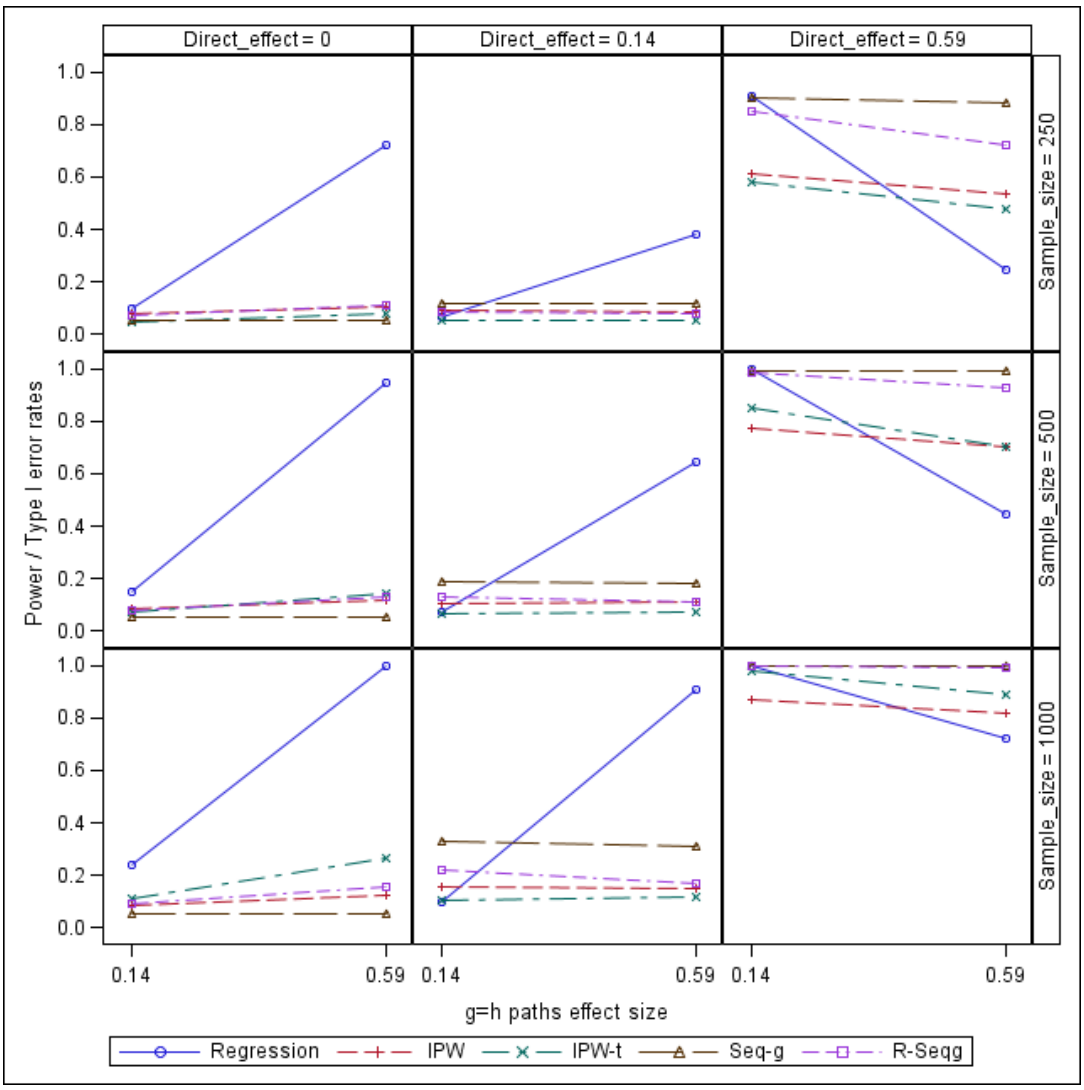
Figure 13. Direct effect *c'* power and Type I error rates by *a*, *d*, and *f* paths effect size and sample size - two-confounders estimation model



Note1: Values represent Type I error rates when *c'*=0.
Note2: Regression: linear regression adjusting for covariates; IPW: inverse propensity weighting; IPW-t.: IPW with truncated weights; Seq-g: sequential g-estimation; R-Seqg_I: doubly robust sequential g-estimation with one-confounder estimation mediator model; R-Seqg_II: doubly robust sequential g-estimation with one-confounder estimation outcome model; R-Seqg_III: doubly robust sequential g-estimation with one-confounder estimation mediator and outcome models.

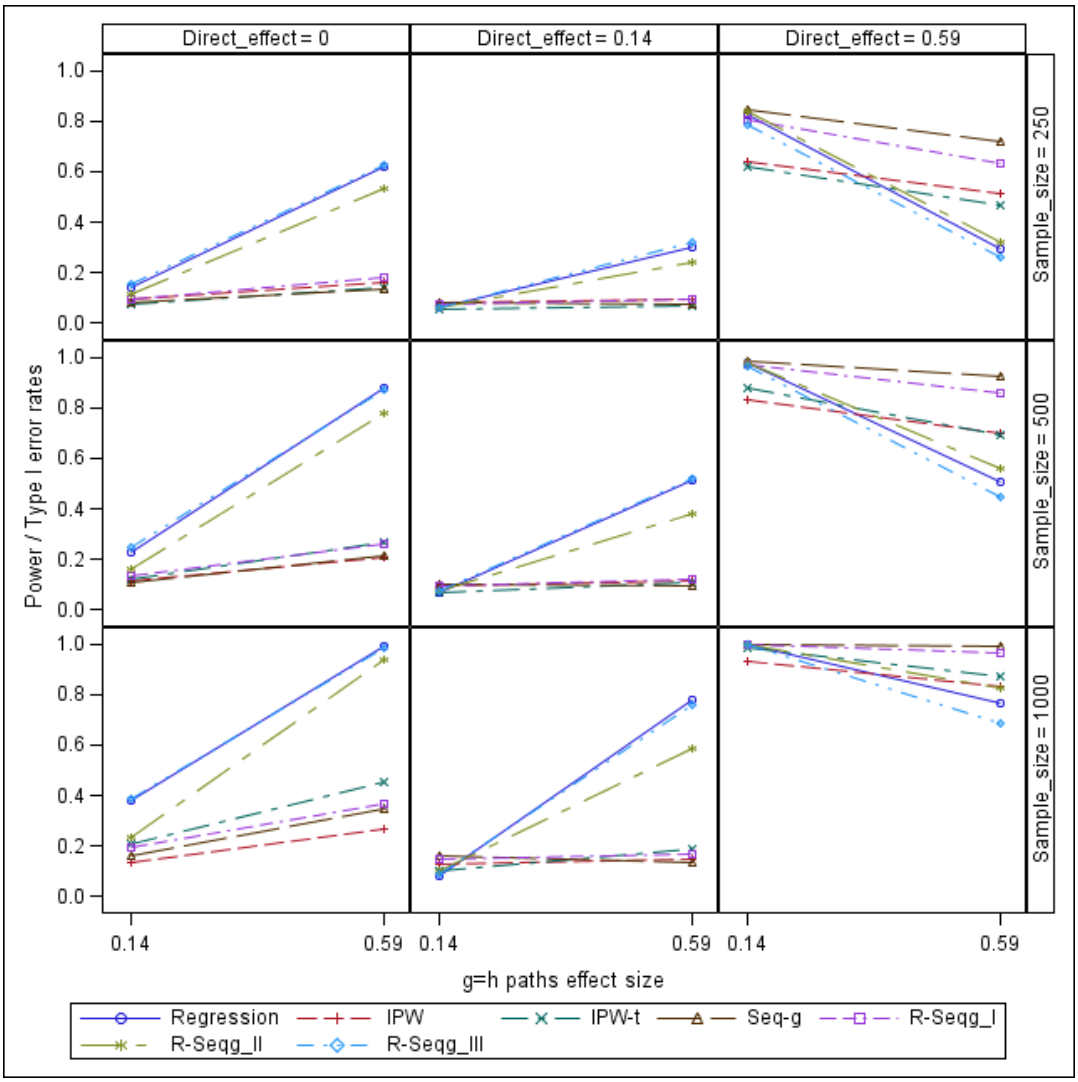
Figure 14. Direct effect *c'* power and Type I error rates by *a*, *d*, and *f* paths effect size and sample size - one-confounder estimation model

Figures 15 and 16 investigate the Type I error rate and statistical power performance of the direct effect for different values of the treatment and confounder relation. Results indicate that linear regression had increasing Type I error rates as the effect of the treatment on the confounder increased and its Type I error rates were out of bounds across conditions for both the two-confounders and one-confounder estimation models. For the two-confounders estimation models in Figure 15, IPW methods had acceptable Type I error rates only when the effect of the treatment on the confounder was equal to .14; and for that condition, the IPW methods had power greater than .80 when the direct effect was .59 for sample sizes of 500 and 1000. The sequential g-estimation had Type I error rates around .05 across conditions, and its power was greater than .80 across all sample sizes when c' was .59. In the case of one-confounder estimation models in Figure 16, none of the methods had acceptable Type I error rates across conditions.



Note1: Values represent Type I error rates when *c'*=0.
Note2: Regression: linear regression adjusting for covariates; IPW: inverse propensity weighting; IPW-t.: IPW with truncated weights; Seq-g: sequential g-estimation; R-Seqg: robust sequential g-estimation.

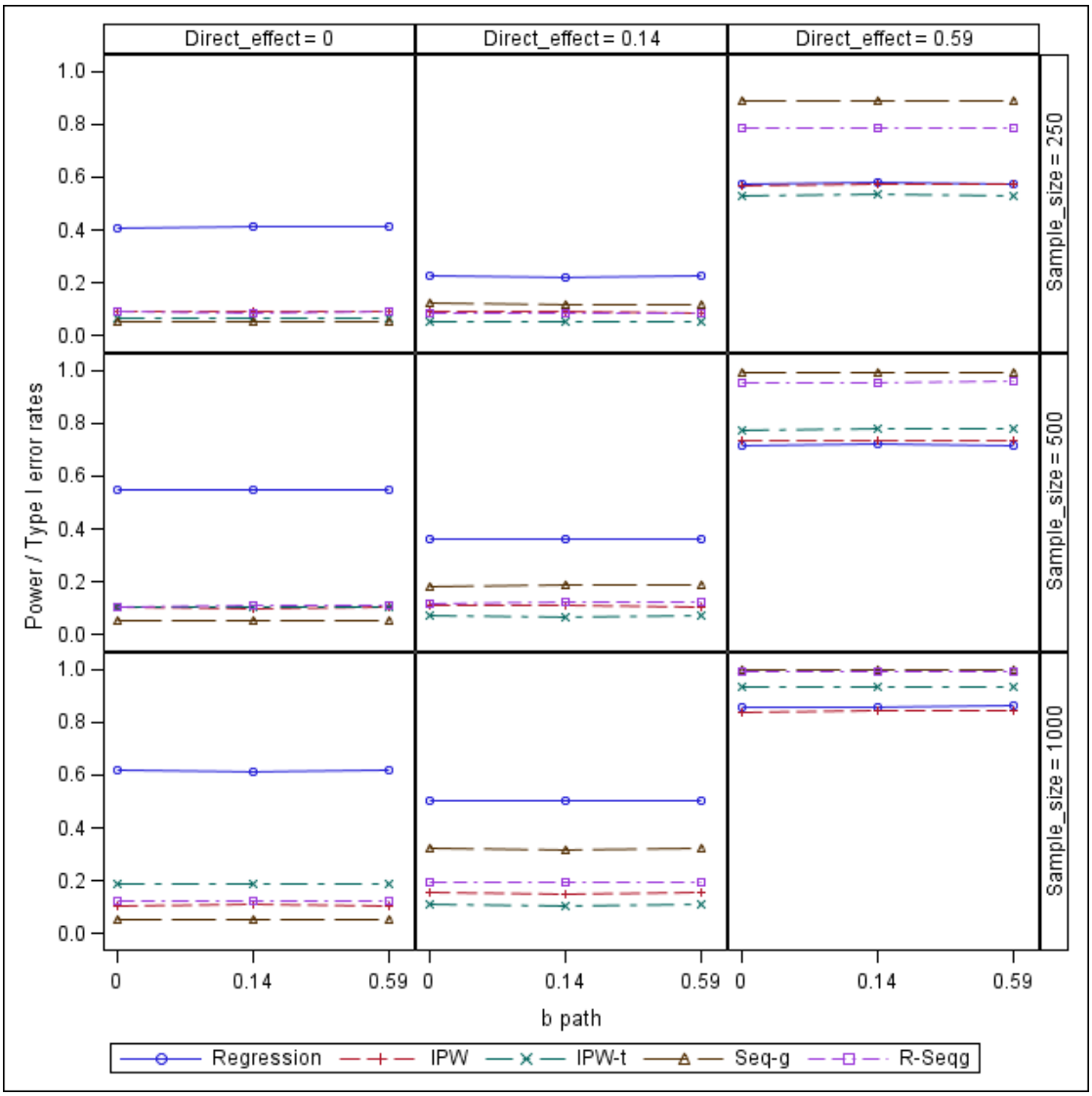
Figure 15. Direct effect *c'* power and Type I error rates by *g* and *h* paths effect size and sample size - two-confounders estimation model



Note1: Values represent Type I error rates when *c'*=0.
Note2: Regression: linear regression adjusting for covariates; IPW: inverse propensity weighting; IPW-t.: IPW with truncated weights; Seq-g: sequential g-estimation; R-Seqg_I: doubly robust sequential g-estimation with one-confounder estimation mediator model; R-Seqg_II: doubly robust sequential g-estimation with one-confounder estimation outcome model; R-Seqg_III: doubly robust sequential g-estimation with one-confounder estimation mediator and outcome models.

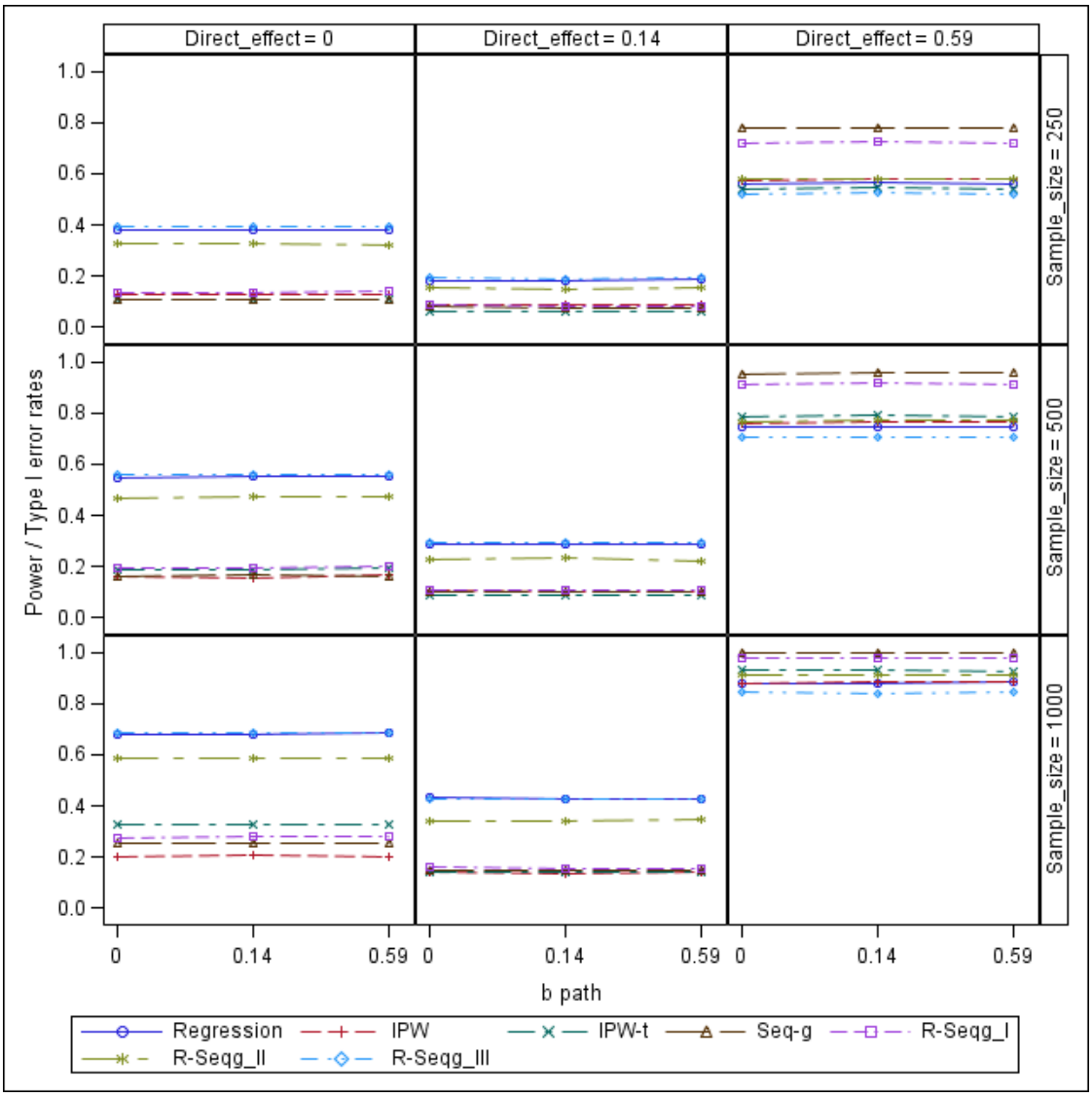
Figure 16. Direct effect *c'* power and Type I error rates by *g* and *h* paths effect size and sample size - one-confounder estimation model

Figures 17-18 show that for both two-confounders estimation and one-confounder estimation models Type I error rates and statistical power were not influenced by the effect size of the *b* path. For the two-confounders estimation models shown in Figure 17, only sequential g-estimation had acceptable Type I error rates across conditions, and its statistical power was greater than .80 when c' was equal to .59. In the case of one-confounder estimation models shown in Figure 18, none of the methods had acceptable Type I error rates across conditions; thus power results were not interpreted.



Note1: Values represent Type I error rates when *c'*=0.
Note2: Regression: linear regression adjusting for covariates; IPW: inverse propensity weighting; IPW-t.: IPW with truncated weights; Seq-g: sequential g-estimation; R-Seqg: robust sequential g-estimation.

Figure 17. Direct effect *c'* power and Type I error rates by *b* path effect size and sample size - two-confounders estimation model

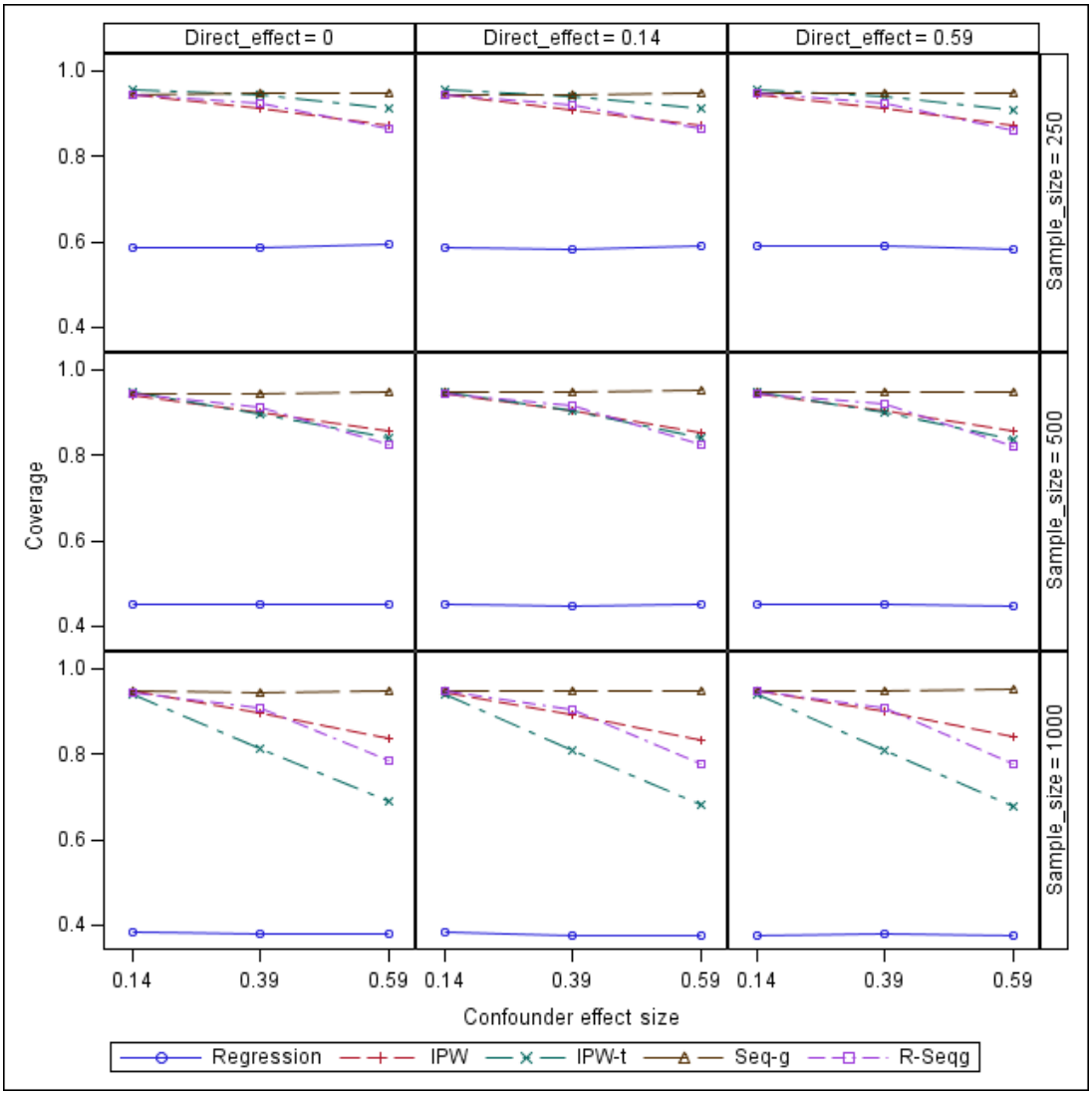


Note1: Values represent Type I error rates when *c'*=0.
Note2: Regression: linear regression adjusting for covariates; IPW: inverse propensity weighting; IPW-t.: IPW with truncated weights; Seq-g: sequential g-estimation; R-Seqg_I: doubly robust sequential g-estimation with one-confounder estimation mediator model; R-Seqg_II: doubly robust sequential g-estimation with one-confounder estimation outcome model; R-Seqg_III: doubly robust sequential g-estimation with one-confounder estimation mediator and outcome models.

Figure 18. Direct effect *c'* power and Type I error rates by *b* path effect size and sample size - one-confounder estimation model

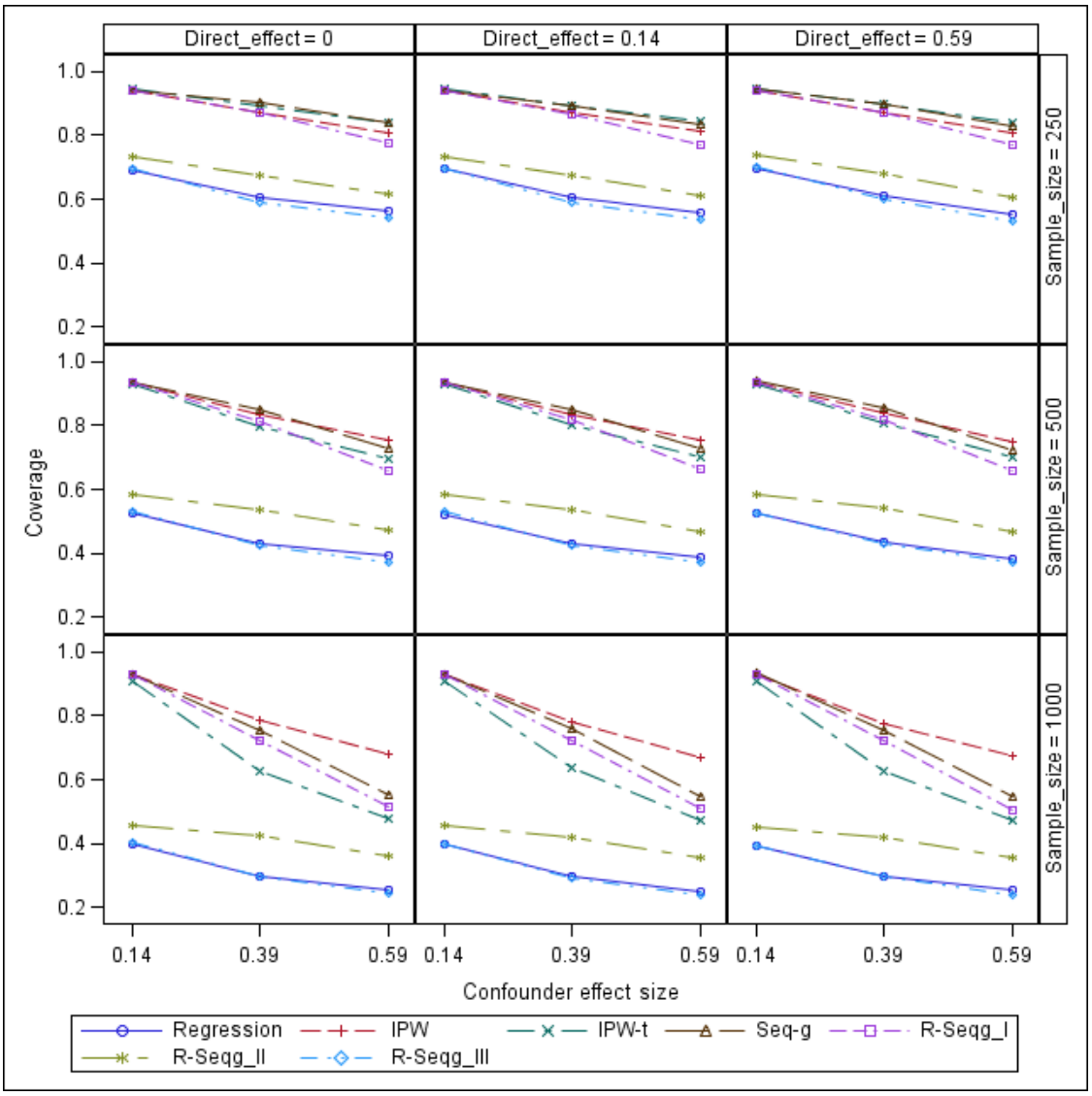
3.1.3 Confidence interval coverage

Figures 19-20 show that confidence interval coverage for the direct effect c' decreases as the effect of the confounders on the mediator decrease, except for the sequential g-estimation and linear regression methods. Sequential g-estimation has good coverage across different confounder effect sizes when the models are correctly specified. The linear regression method has the lowest coverage compared to other methods for both two-confounders estimation and one-confounder estimation models. Also, for the one-confounder estimation case, the coverage of the doubly robust sequential g-estimation with the one-confounder estimation outcome model and doubly robust sequential g-estimation with the one-confounder estimation outcome and mediator models is as low as the coverage of the adjusted linear regression method. The pattern of results is similar across different sample sizes.



Note: Regression: linear regression adjusting for covariates; IPW: inverse propensity weighting; IPW-t.: IPW with truncated weights; Seq-g: sequential g-estimation; R-Seqg: robust sequential g-estimation.

Figure 19. Direct effect *c'* confidence interval coverage by post-treatment confounders effect size and sample size - two-confounders estimation model

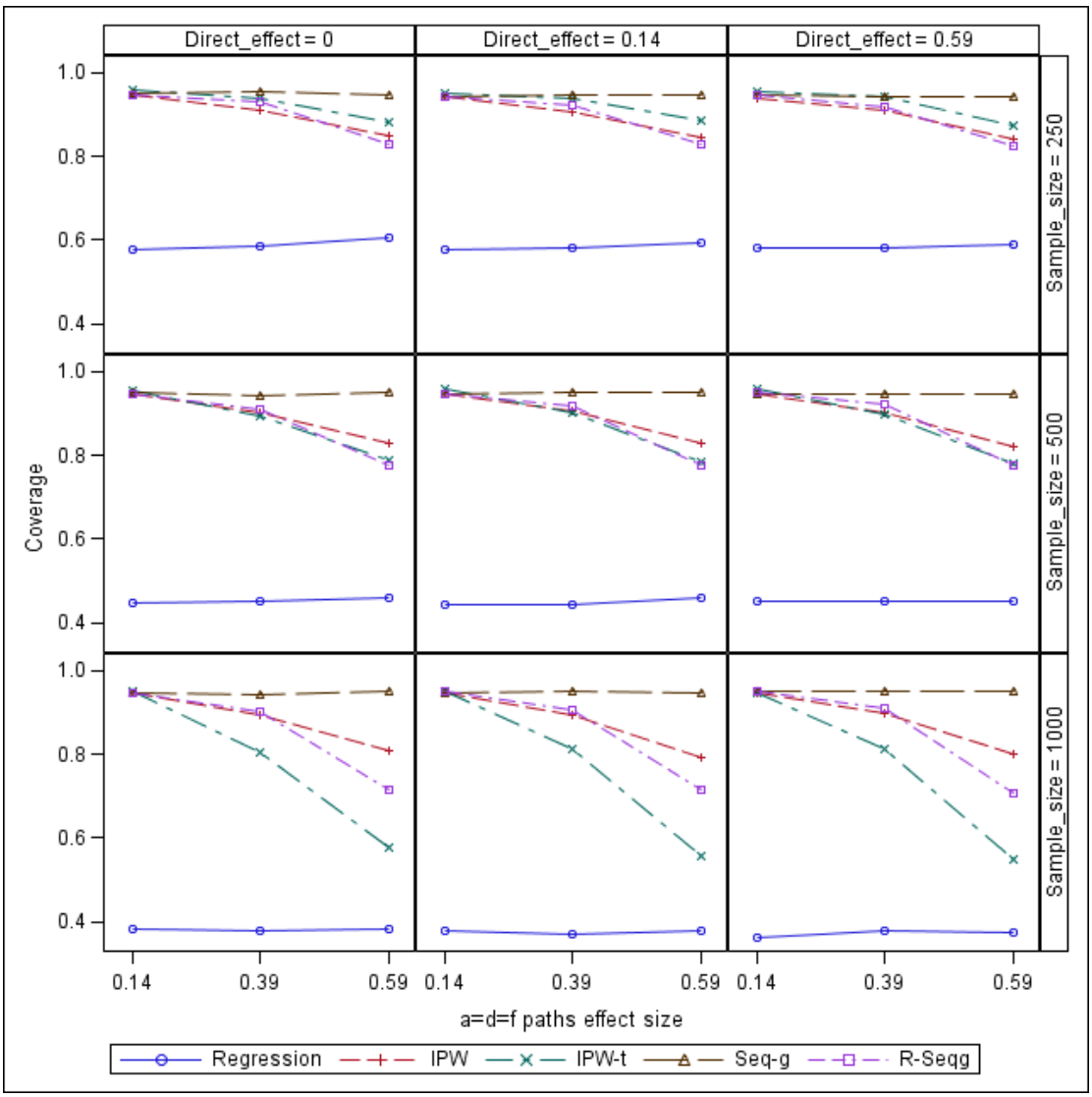


Note: Regression: linear regression adjusting for covariates; IPW: inverse propensity weighting; IPW-t.: IPW with truncated weights; Seq-g: sequential g-estimation; R-Seqg_I: doubly robust sequential g-estimation with one-confounder estimation mediator model; R-Seqg_II: doubly robust sequential g-estimation with one-confounder estimation outcome model; R-Seqg_III: doubly robust sequential g-estimation with one-confounder estimation mediator and outcome models.

Figure 20. Direct effect *c'* confidence interval coverage by post-treatment confounders effect size and sample size - one-confounder estimation model

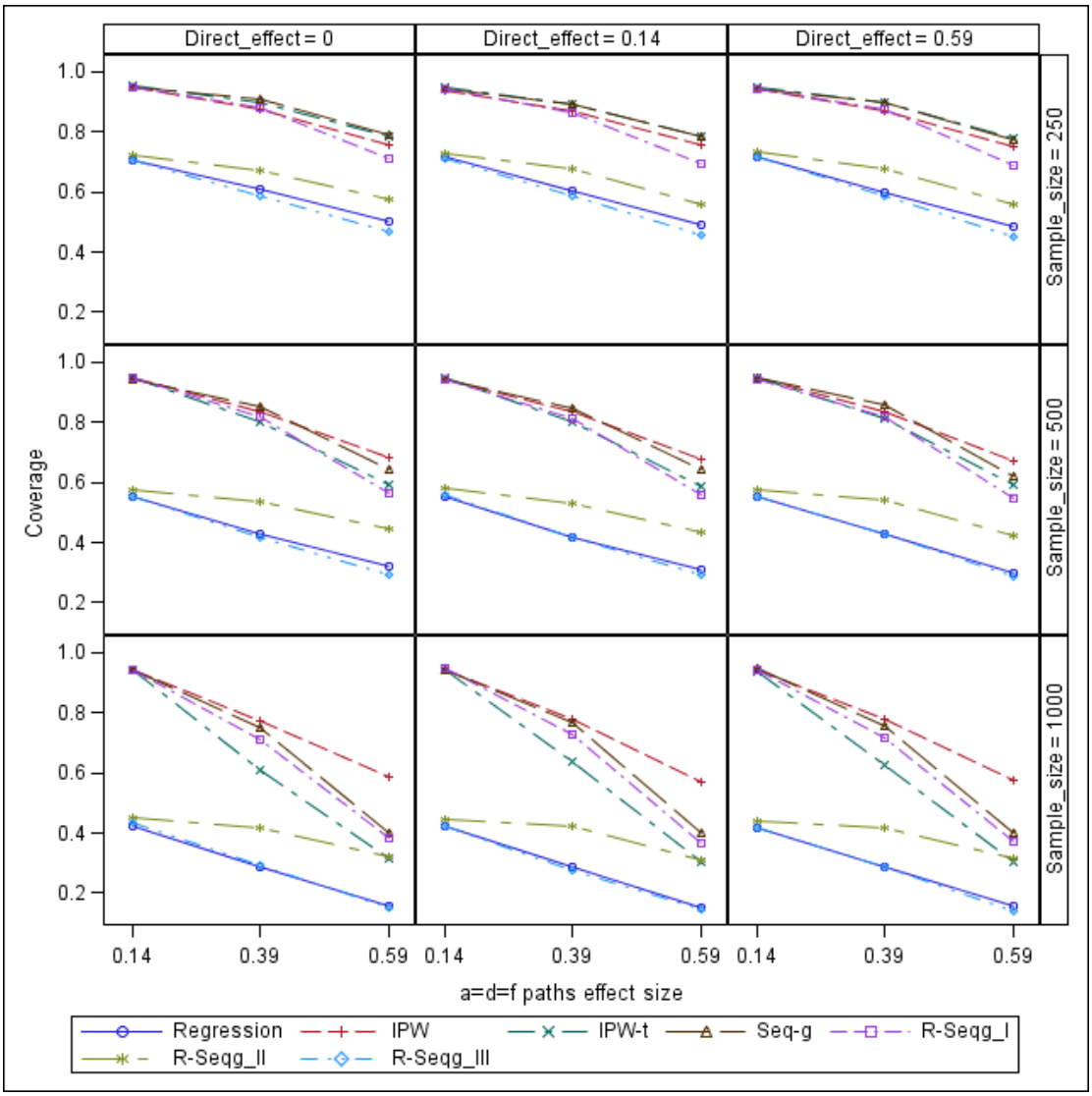
Figures 21-22 show that confidence interval coverage decreases as the effect of the treatment and confounders on the mediator decrease, except for the sequential g-estimation and linear regression methods. Sequential g-estimation has good coverage across different confounder effect sizes when the models are correctly specified. For the two-confounders estimation models, the coverage of the IPW-truncated method gets much lower as sample size increases compared to the conventional IPW method.

The linear regression method has the lowest coverage compared to other methods for both two-confounders estimation and one-confounder estimation models. Also, for the one-confounder estimation case, the coverage of the doubly robust sequential g-estimation with the one-confounder estimation outcome model and doubly robust sequential g-estimation with the one-confounder estimation outcome and mediator models is as low as the coverage of the adjusted linear regression method. The pattern of results is similar across different sample sizes.



Note: Regression: linear regression adjusting for covariates; IPW: inverse propensity weighting; IPW-t.: IPW with truncated weights; Seq-g: sequential g-estimation; R-Seqg: robust sequential g-estimation.

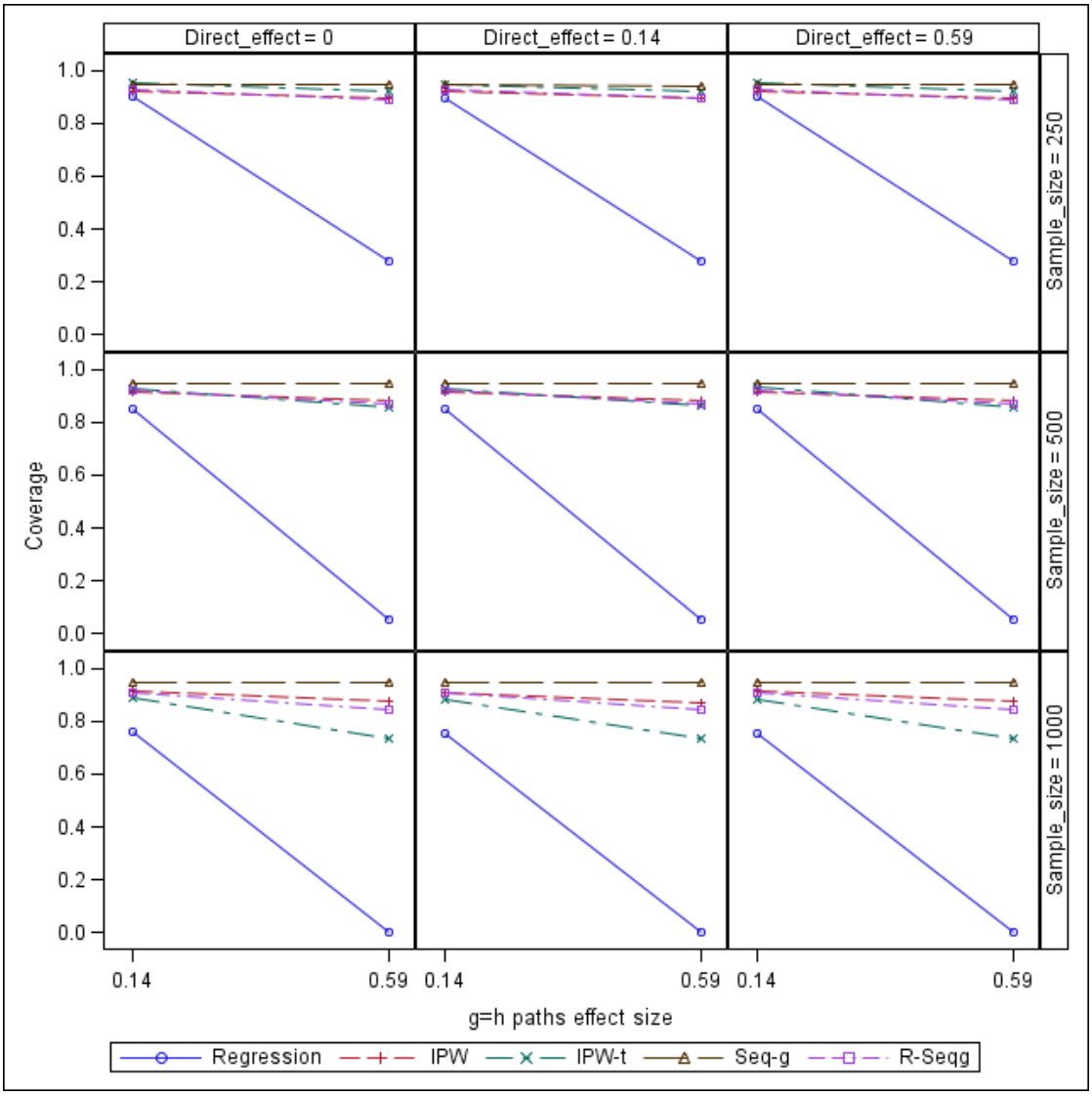
Figure 21. Direct effect *c'* confidence interval coverage by *a*, *d*, and *f* paths effect size and sample size - two-confounders estimation model



Note: Regression: linear regression adjusting for covariates; IPW: inverse propensity weighting; IPW-t.: IPW with truncated weights; Seq-g: sequential g-estimation; R-Seqg_I: doubly robust sequential g-estimation with one-confounder estimation mediator model; R-Seqg_II: doubly robust sequential g-estimation with one-confounder estimation outcome model; R-Seqg_III: doubly robust sequential g-estimation with one-confounder estimation mediator and outcome models.

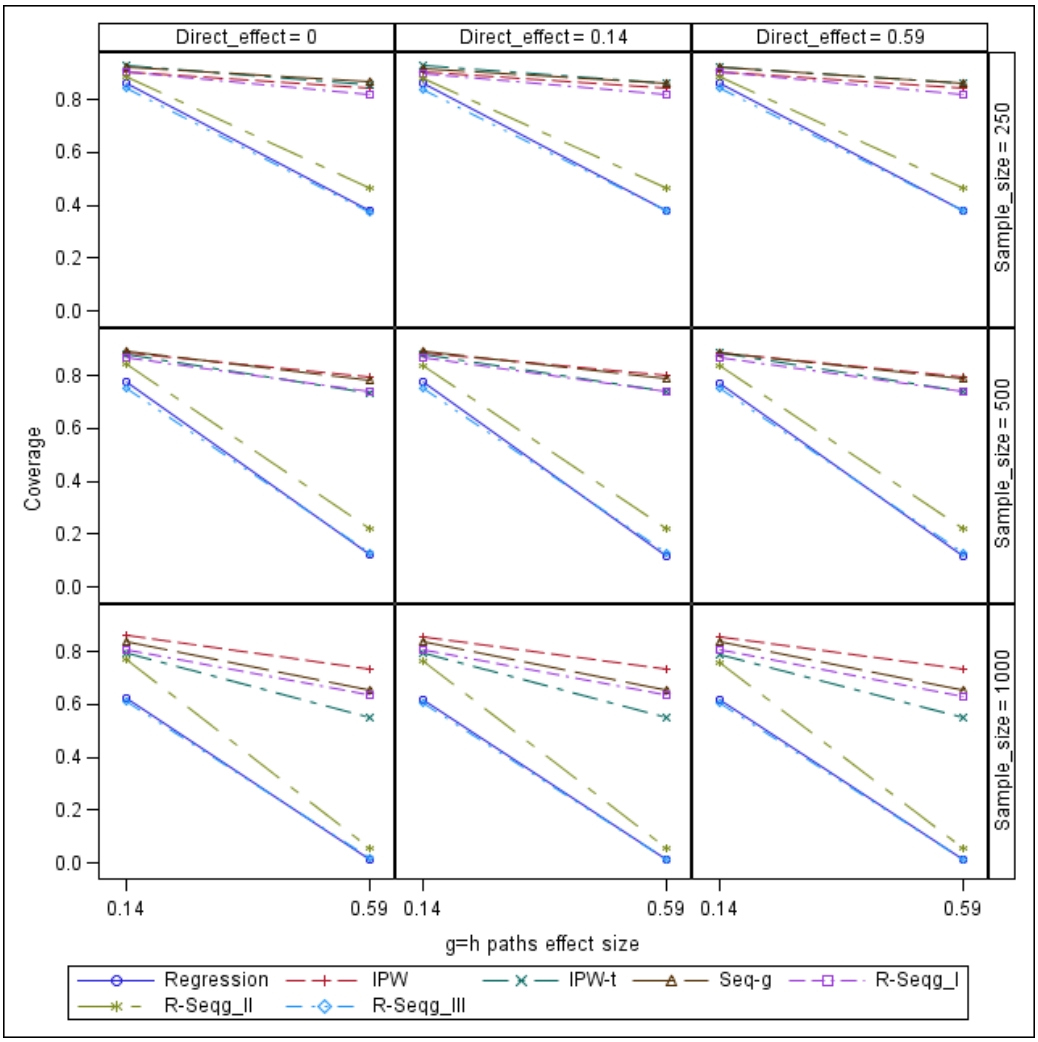
Figure 22. Direct effect *c'* confidence interval coverage by *a*, *d*, and *f* paths effect size and sample size - one-confounder estimation model

Figure 23 show that for the two-confounders estimation models, the coverage of the linear regression method decreases as the effect of the treatment on the confounders increases. For the one-confounder estimation model as depicted in Figure 24, the coverage of the linear regression, doubly robust sequential g-estimation with the one-confounder estimation outcome model and doubly robust sequential g-estimation with the one-confounder estimation outcome and mediator models decreases steeply as the effect of the treatment on the confounders increases.



Note: Regression: linear regression adjusting for covariates; IPW: inverse propensity weighting; IPW-t.: IPW with truncated weights; Seq-g: sequential g-estimation; R-Seqg: robust sequential g-estimation.

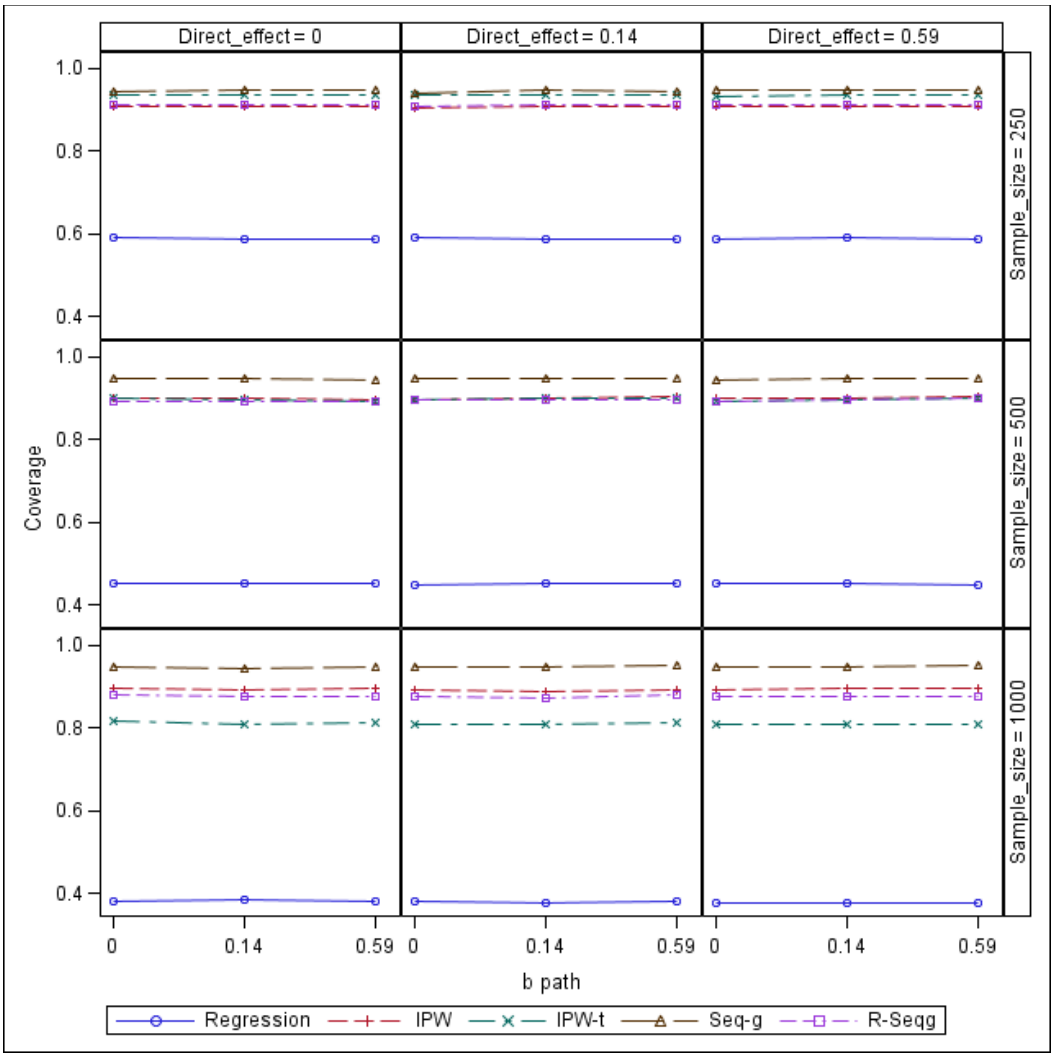
Figure 23. Direct effect *c'* confidence interval coverage by *g* and *h* paths effect size and sample size - two-confounders estimation model



Note: Regression: linear regression adjusting for covariates; IPW: inverse propensity weighting; IPW-t.: IPW with truncated weights; Seq-g: sequential g-estimation; R-Seqg_I: doubly robust sequential g-estimation with one-confounder estimation mediator model; R-Seqg_II: doubly robust sequential g-estimation with one-confounder estimation outcome model; R-Seqg_III: doubly robust sequential g-estimation with one-confounder estimation mediator and outcome models.

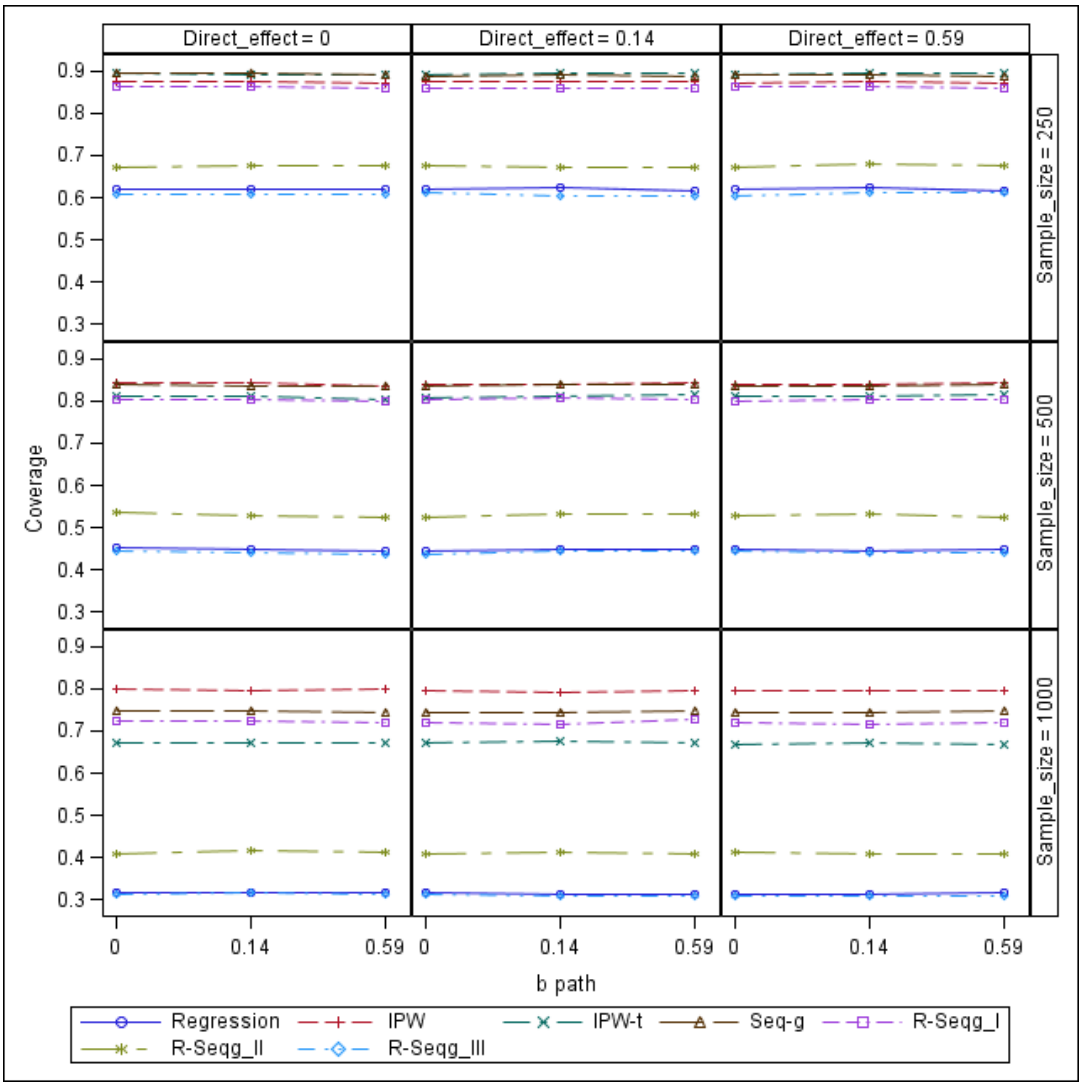
Figure 24. Direct effect *c'* confidence interval coverage by *g* and *h* paths effect size and sample size - one-confounder estimation model

Figures 25-26 show that the confidence interval coverage of the direct effect is not influenced by the size of the *b* path for any of the methods.



Note: Regression: linear regression adjusting for covariates; IPW: inverse propensity weighting; IPW-t.: IPW with truncated weights; Seq-g: sequential g-estimation; R-Seqg: robust sequential g-estimation.

Figure 25. Direct effect *c'* confidence interval coverage by *b* path effect size and sample size - two-confounders estimation model



Note2: Regression: linear regression adjusting for covariates; IPW: inverse propensity weighting; IPW-t.: IPW with truncated weights; Seq-g: sequential g-estimation; R-Seqg_I: doubly robust sequential g-estimation with one-confounder estimation mediator model; R-Seqg_II: doubly robust sequential g-estimation with one-confounder estimation outcome model; R-Seqg_III: doubly robust sequential g-estimation with one-confounder estimation mediator and outcome models.

Figure 26. Direct effect *c'* confidence interval coverage by *b* path effect size and sample size - one-confounder estimation model

3.2 The indirect effect, *c*-*c'*

3.2.1 Accuracy of point estimates

Tables 9 and 10 address the research question, "How robust are the methods to the violation of the assumption of unmeasured confounders of the M to Y relation?" for the indirect effect. In Table 9, bias, relative bias, and mean square error of estimates of effects for two-confounders estimation models are presented across different effect sizes for the relation between post-treatment confounders and the mediator, i.e., paths *d* and *f*. Table 10 presents the results for one-confounder estimation models where the confounder C_2 was omitted from the analyses.

For the two-confounders estimation model results in Table 9, bias, relative bias and MSE for the indirect effect increase as the effect of the post-treatment confounder on the mediator increases for all methods except for linear regression and sequential g-estimation. IPW had acceptable relative bias when only confounder effect size was .14. The truncated IPW method did not have acceptable relative bias in any of the simulation conditions. Again, this finding may be due to the trimming rule used. Determining which level of trimming to be used may be difficult and so weight trimming may have adverse effects in the estimation. Sequential g-estimation was unbiased across different effect sizes of the confounder and the indirect effect. The doubly robust sequential g-estimation had unacceptable relative bias as the confounder effect size increased. This finding may be consistent with the fact that the doubly robust g-estimation method uses IPW in its estimation process and so follow the bias pattern of the IPW method. Linear regression had unacceptable relative bias across conditions and its performance was not influenced

by the size of the confounder effect size. This finding again can be explained by the fact that the confounders were post-treatment.

In the case of one-confounder estimation models in Table 10, all methods have unacceptable relative bias rates. Bias, relative bias and MSE for the indirect effect estimates increase as the effect of the post-treatment confounder on the mediator increases for all methods.. The results again also show that doubly robust sequential g-estimation performs the worst when both the mediator and the outcome estimation models violate the no omitted variables assumption. Figures 27 and 28 show that the pattern of results is similar across different sample sizes.

Table 9

Indirect effect c-c' bias, relative bias, and mean square error by post-treatment confounders effect size (N=500) - two-confounders estimation models

		c-c' effect size								
		0			0 < ... <.39			>.39		
		Post treatment confounder effect size (paths *d* and *f*)								
Method		.14	.39	.59	.14	.39	.59	.14	.39	.59
Regression	Bias	.243	.242	.244	.229	.227	.206	.391	.293	.319
	Rel.Bias	na	na	na	2.649	1.963	1.655	.878	.547	.536
	MSE	.088	.087	.088	.084	.085	.077	.166	.125	.145
IPW	Bias	.003	.051	.117	.003	.047	.107	.006	.066	.140
	Rel.Bias	na	na	na	.033	.419	.880	.013	.130	.236
	MSE	.005	.056	.111	.006	.056	.115	.009	.065	.134
IPW trunc.	Bias	.032	.125	.189	.031	.111	.167	.053	.163	.234
	Rel.Bias	na	na	na	.322	.997	1.369	.120	.319	.398
	MSE	.004	.028	.056	.006	.028	.054	.011	.047	.088
Seq. *g*-est.	Bias	-.000	-.000	.000	.000	.000	.000	.000	-.001	-.001
	Rel.Bias	na	na	na	.000	-.002	.004	.000	-.001	.002
	MSE	.001	.001	.002	.002	.003	.004	.005	.007	.012
R. seq. *g*-est.	Bias	.003	.052	.118	.003	.046	.103	.003	.068	.144
	Rel.Bias	na	na	na	.028	.414	.843	.008	.133	.245
	MSE	.001	.009	.025	.003	.010	.025	.006	.021	.047

Note: Regression: linear regression adjusting for covariates; IPW: inverse propensity weighting; IPW trunc.: IPW with truncated weights; Seq. g-est: sequential g-estimation; R. seq. g-est.: doubly robust sequential g-estimation.

Table 10

Indirect effect c-c' bias, relative bias, and mean square error by post-treatment confounders effect size (N=500) - one-confounder estimation models

		c-c' effect size								
		0			0 < ... <.39			>.39		
		Post treatment confounder effect size (paths *d* and *f*)								
Method		.14	.39	.59	.14	.39	.59	.14	.39	.59
Regression	Bias	.209	.252	.276	.198	.231	.239	.336	.316	.350
	Rel.Bias	na	na	na	2.252	2.041	1.937	.755	.602	.594
	MSE	.062	.084	.098	.060	.079	.087	.125	.129	.157
IPW	Bias	.034	.124	.190	.032	.112	.166	.055	.160	.231
	Rel.Bias	na	na	na	.335	1.012	1.353	.124	.313	.394
	MSE	.005	.042	.083	.006	.042	.080	.011	.061	.115
IPW trunc.	Bias	.055	.169	.229	.052	.150	.203	.088	.221	.282
	Rel.Bias	na	na	na	.542	1.359	1.658	.198	.432	.481
	MSE	.006	.040	.071	.008	.038	.066	.016	.068	.111
Seq. *g*-est.	Bias	.032	.104	.156	.031	.094	.138	.052	.138	.194
	Rel.Bias	na	na	na	.317	.845	1.125	.116	.269	.331
	MSE	.002	.014	.031	.004	.015	.030	.008	.030	.057
R. seq. *g*-est. I	Bias	.034	.124	.189	.032	.111	.166	.053	.162	.233
	Rel.Bias	na	na	na	.332	1.000	1.358	.120	.380	.399
	MSE	.002	.022	.048	.004	.022	.044	.009	.042	.080
R. seq. *g*-est. II	Bias	.185	.217	.253	.174	.201	.217	.297	.268	.321
	Rel.Bias	na	na	na	2.004	1.754	1.763	.666	.507	.544
	MSE	.052	.068	.087	.051	.065	.077	.099	.104	.141
R. seq. *g*-est. III	Bias	.211	.266	.296	.198	.244	.257	.338	.334	.373
	Rel.Bias	na	na	na	2.251	2.150	2.084	.759	.609	.633
	MSE	.063	.093	.112	.060	.086	.099	.126	.144	.177

Note: Regression: linear regression adjusting for covariates; IPW: inverse propensity weighting; IPW trunc.: IPW with truncated weights; Seq. g-est: sequential g-estimation; R. seq. g-est. I: doubly robust sequential g-estimation with one-confounder estimation mediator model; R. seq. g-est. II: doubly robust sequential g-estimation with one-confounder estimation outcome model; R. seq. g-est. III: doubly robust sequential g-estimation with one-confounder estimation mediator and outcome models.

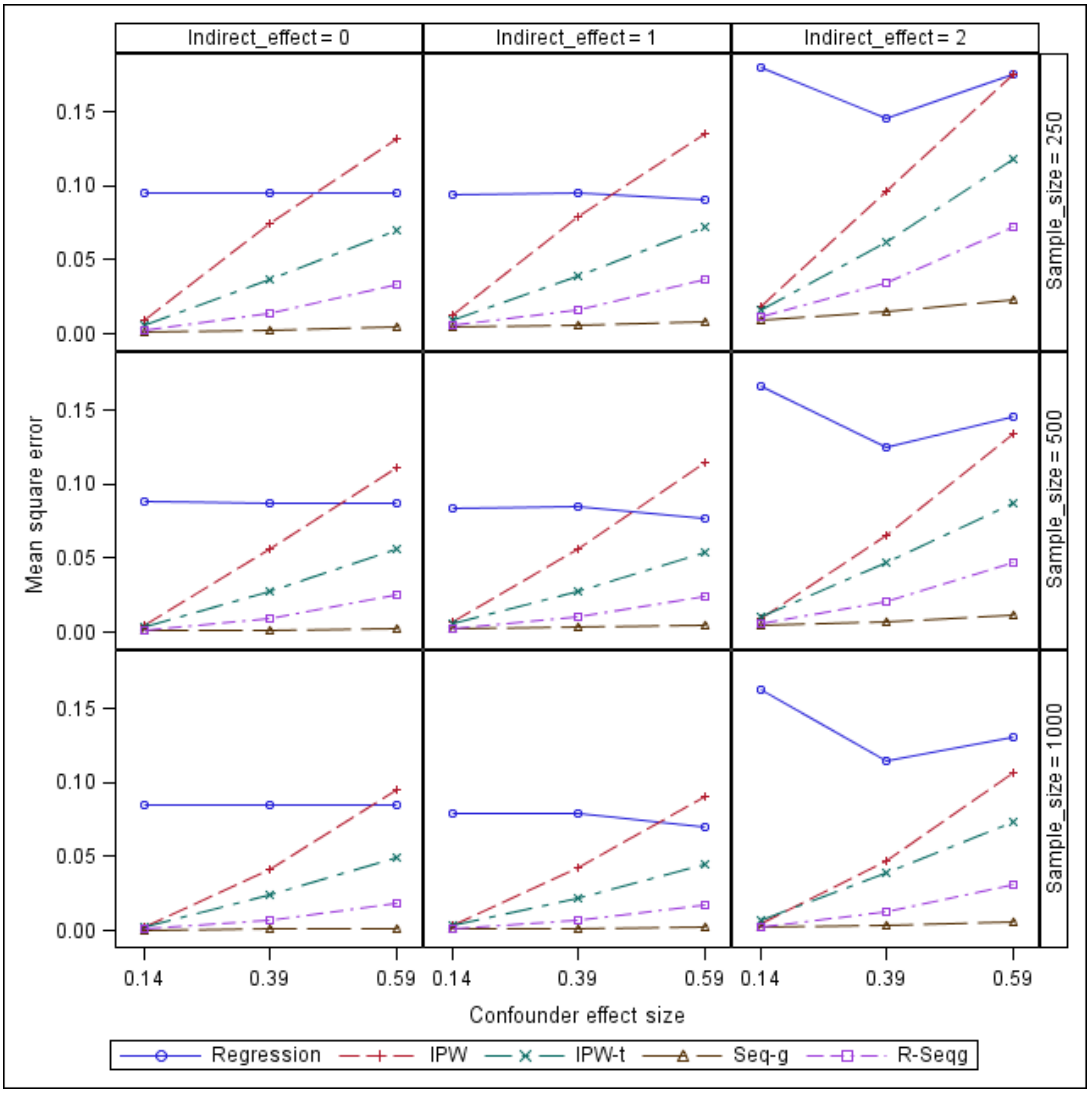


Figure 27. Indirect effect *c-c'* mean square error by post-treatment confounders effect size and sample size - two-confounders estimation model

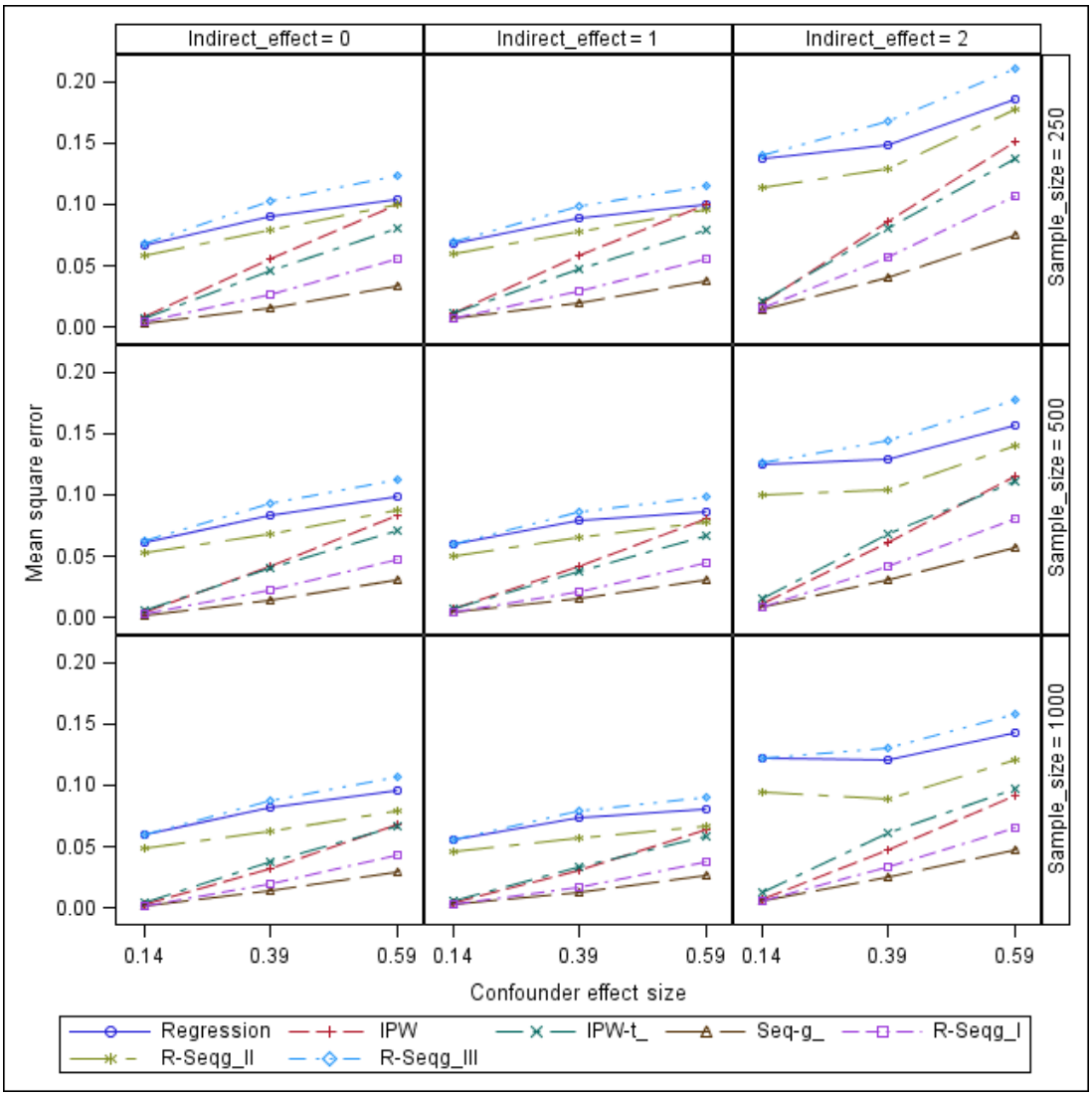


Figure 28. Indirect effect *c-c'* mean square error by post-treatment confounders effect size and sample size - one-confounder estimation model

Tables 11 and 12 especially investigate the performance of the IPW indirect effect estimates depending on the effect of the treatment (the *a* path) and post-treatment covariates (the *d* and *f* paths). Results show that the bias in both the IPW and IPW-truncated estimates of the *c*-*c'* estimate get larger as the *a, d,* and, *f* paths increase in both two-confounders estimation and one-confounder estimation models. The IPW method had only acceptable relative bias less than .10 in one condition where the confounder effect size was .14 and the indirect effect was between 0 and .39 for the two-confounders estimation model in Table 11. For the one-confounder estimation model, neither the IPW nor the IPW-truncated had acceptable relative bias in any of the simulation conditions. Figures 29 and 30 showing the MSE results for the indirect effect indicate that the pattern of the results were similar across sample sizes.

Please note that some tables and figures for the indirect effect results have empty cells for some conditions (e.g., tables 11 and 12 have the empty cells for the case of $c\text{-}c' \geq .39$ and $a=d=f=.14$). These are the conditions that were not available in the simulation. The study was designed to simulate specific effect sizes for the direct effect *c'*, and not for the indirect effect *c*-*c'*. Thus, the true indirect effect estimates were realized as a result of the simulated values for the direct effect which led to some of these missing conditions.

Table 11

Indirect effect c-c' bias, relative bias, and mean square error by a, d, and f paths effect size (N=500) - two-confounders estimation models

		c-c' effect size								
		0			0 < ... <.39			>.39		
		a=d=f paths effect size								
Method		.14	.39	.59	.14	.39	.59	.14	.39	.59
Regression	Bias	.245	.241	.244	.243	.193	.247	na	.391	.243
	Rel.Bias	na	na	na	3.980	1.638	1.559	na	.780	.362
	MSE	.088	.087	.088	.090	.069	.093	na	.173	.109
IPW	Bias	.002	.052	.146	.001	.047	.157	na	.059	.140
	Rel.Bias	na	na	na	.041	.396	1.106	na	.118	.228
	MSE	.004	.056	.123	.006	.056	.129	na	.064	.135
IPW trunc.	Bias	.018	.128	.239	.016	.116	.243	na	.160	.237
	Rel.Bias	na	na	na	.302	1.003	1.709	na	.318	.390
	MSE	.002	.027	.076	.004	.028	.080	na	.045	.092
Seq. *g*-est.	Bias	.000	-.001	.000	-.000	-.000	.002	na	-.001	-.000
	Rel.Bias	na	na	na	-.003	-.005	.014	na	-.003	-.001
	MSE	.000	.001	.003	.002	.003	.004	na	.007	.012
R. seq. *g*-est.	Bias	.002	.052	.148	.001	.048	.149	na	.066	.145
	Rel.Bias	na	na	na	.023	.409	1.042	na	.131	.241
	MSE	.000	.009	.036	.002	.010	.038	na	.020	.049

Note1: Regression: linear regression adjusting for covariates; IPW: inverse propensity weighting; IPW trunc.: IPW with truncated weights; Seq. g-est: sequential g-estimation; R. seq. g-est.: doubly robust sequential g-estimation.
Note2: na: not applicable.

Table 12

Indirect effect c-c' bias, relative bias, and mean square error by a, d, and f paths effect size (N=500) - one-confounder estimation models

		c-c' effect size								
		0			0 < ... <.39			>.39		
		a=d=f paths effect size								
Method		.14	.39	.59	.14	.39	.59	.14	.39	.59
Regression	Bias	.194	.254	.318	.193	.215	.321	na	.373	.317
	Rel.Bias	na	na	na	3.206	1.838	2.164	na	.743	.506
	MSE	.055	.084	.122	.058	.069	.127	na	.157	.143
IPW	Bias	.018	.125	.238	.017	.116	.243	na	.152	.238
	Rel.Bias	na	na	na	.325	1.001	1.706	na	.302	.393
	MSE	.003	.042	.106	.005	.042	.109	na	.058	.120
IPW trunc.	Bias	.029	.172	.289	.028	.156	.293	na	.215	.288
	Rel.Bias	na	na	na	.513	1.349	2.058	na	.428	.475
	MSE	.003	.039	.101	.005	.037	.105	na	.065	.118
Seq. *g*-est.	Bias	.017	.106	.199	.016	.097	.200	na	.134	.198
	Rel.Bias	na	na	na	.296	.836	1.402	na	.267	.328
	MSE	.001	.014	.045	.003	.015	.048	na	.028	.060
R. seq. *g*-est. I	Bias	.017	.126	.239	.017	.115	.241	na	.159	.238
	Rel.Bias	na	na	na	.311	.988	1.692	na	.316	.393
	MSE	.001	.021	.070	.003	.021	.072	na	.040	.085
R. seq. *g*-est. II	Bias	.186	.217	.284	.183	.180	.285	na	.334	.282
	Rel.Bias	na	na	na	2.994	1.531	1.906	na	.666	.445
	MSE	.052	.068	.105	.053	.056	.109	na	.132	.125
R. seq. *g*-est. III	Bias	.196	.268	.346	.193	.228	.349	na	.389	.345
	Rel.Bias	na	na	na	3.196	1.952	2.372	na	.775	.553
	MSE	.056	.093	.144	.058	.077	.149	na	.172	.165

Note: Regression: linear regression adjusting for covariates; IPW: inverse propensity weighting; IPW trunc.: IPW with truncated weights; Seq. g-est: sequential g-estimation; R. seq. g-est. I: doubly robust sequential g-estimation with one-confounder estimation mediator model; R. seq. g-est. II: doubly robust sequential g-estimation with one-confounder estimation outcome model; R. seq. g-est. III: doubly robust sequential g-estimation with one-confounder estimation mediator and outcome models.

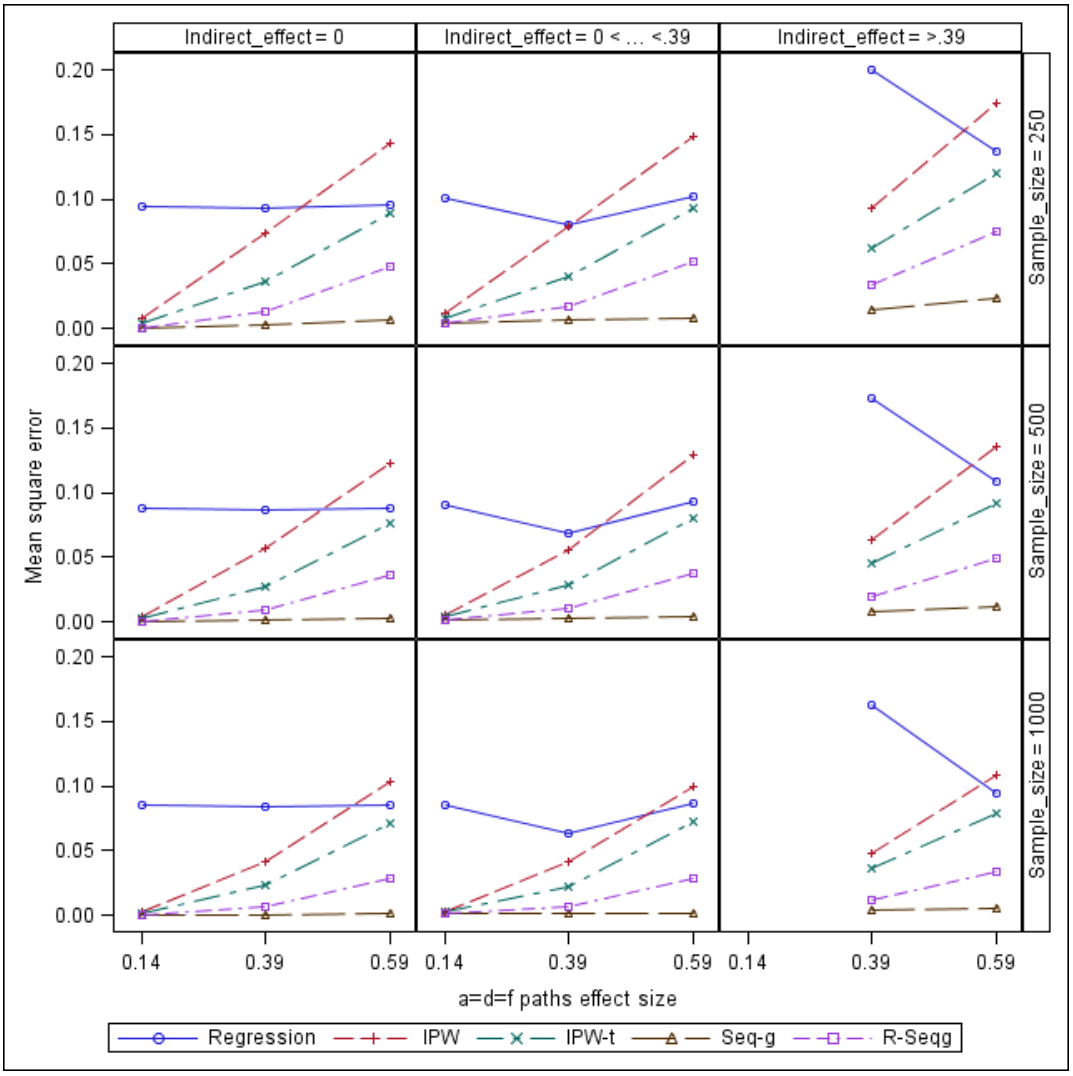


Figure 29. Indirect effect *c-c'* mean square error by *a*, *d*, and *f* paths effect size and sample size - two-confounders estimation model

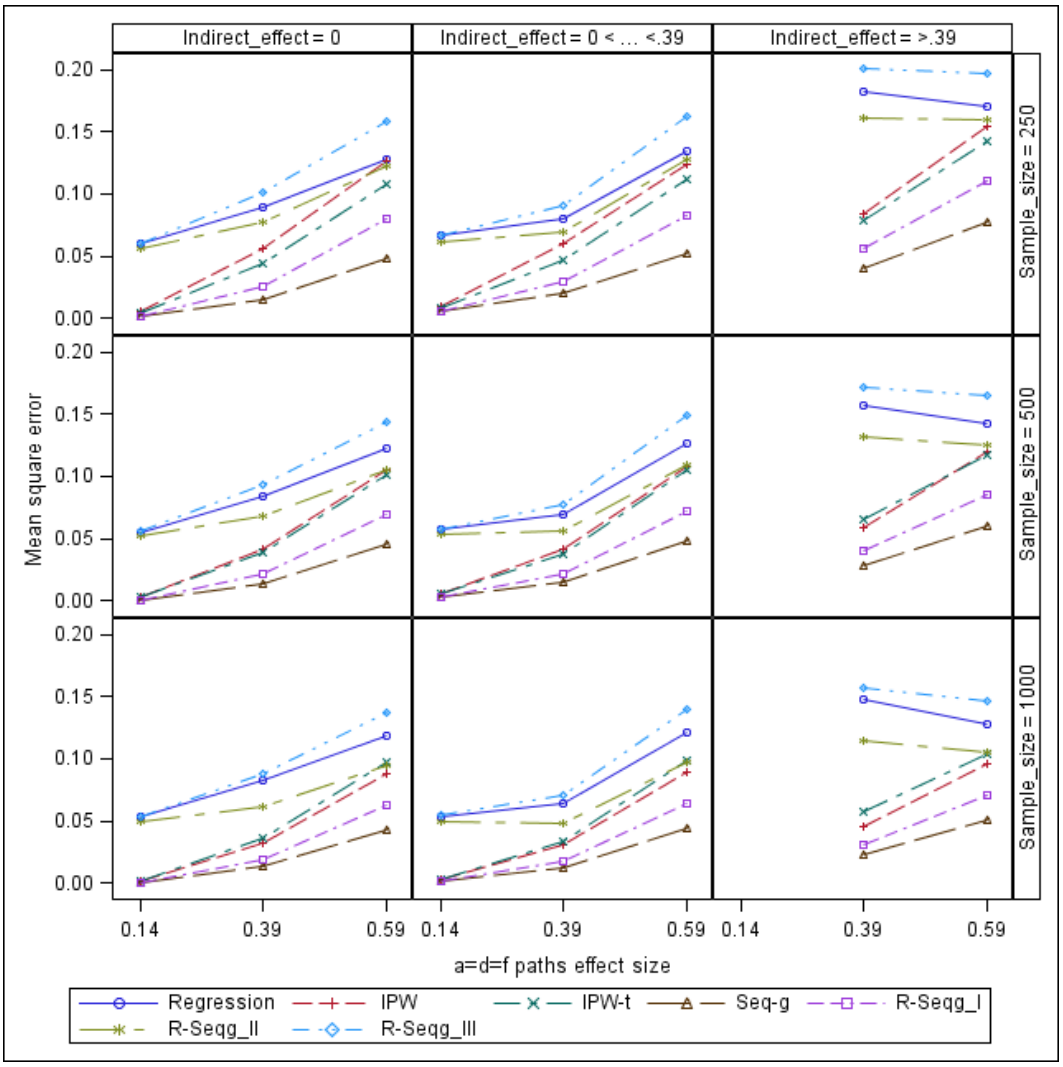


Figure 30. Indirect effect *c-c'* mean square error by *a*, *d*, and *f* paths effect size and sample size - one-confounder estimation model

Tables 13 and 14 investigate whether the bias in the indirect effect estimates increases as the effect of the treatment on the confounders increases, i.e., as paths *g* and *h* increase. For the two-confounders estimation models in Table 13, the bias for the linear regression estimate of the indirect effect significantly increases as the relation between the treatment and confounders increases. Linear regression had no acceptable relative bias in any of the conditions. Similar to the direct effect results, the indirect effect results also indicate that linear regression produces biased effect estimates when confounders are post-treatment. Sequential g-estimation had unbiased estimates across conditions. The IPW methods and the robust g-estimation had unacceptable relative bias across conditions. In the case of one-confounder estimation model, none of the methods had acceptable relative bias in any of the conditions. Figures 31 and 32 indicate that results were similar across sample size conditions.

Table 13

Indirect effect c-c' bias, relative bias, and mean square error by g and h paths effect size (N=500) - two-confounders estimation models

		c-c' effect size					
		0		0 < ... <.39		>.39	
		g=h paths effect size					
Method		.14	.59	.14	.59	.14	.59
Regression	Bias	.093	.393	.093	.393	.092	.394
	Rel.Bias	na	na	1.173	3.440	.214	.706
	MSE	.015	.161	.021	.165	.031	.177
IPW	Bias	.040	.074	.036	.061	.076	.102
	Rel.Bias	na	na	.377	.428	.175	.166
	MSE	.051	.064	.052	.055	.089	.094
IPW trunc.	Bias	.082	.149	.073	.126	.150	.196
	Rel.Bias	na	na	.773	.925	.348	.329
	MSE	.018	.041	.020	.035	.047	.068
Seq. *g*-est.	Bias	.000	.000	.000	-.000	-.001	.001
	Rel.Bias	na	na	.001	-.000	-.002	.001
	MSE	.001	.002	.003	.003	.008	.009
R. seq. *g*-est.	Bias	.040	.075	.034	.060	.081	.103
	Rel.Bias	na	na	.357	.422	.188	.169
	MSE	.006	.018	.008	.016	.024	.034

Note: Regression: linear regression adjusting for covariates; IPW: inverse propensity weighting; IPW trunc.: IPW with truncated weights; Seq. g-est: sequential g-estimation; R. seq. g-est.: doubly robust sequential g-estimation.

Table 14

Indirect effect c-c' bias, relative bias, and mean square error by g and h paths effect size (N=500) - one-confounder estimation models

		c-c' effect size					
		0		0 < ... <.39		>.39	
		g=h paths effect size					
Method		.14	.59	.14	.59	.14	.59
Regression	Bias	.127	.365	.120	.354	.178	.388
	Rel.Bias	na	na	1.400	3.020	.415	.684
	MSE	.022	.140	.026	.136	.053	.173
IPW	Bias	.082	.149	.073	.126	.151	.193
	Rel.Bias	na	na	.778	.932	.351	.323
	MSE	.031	.056	.032	.048	.066	.087
IPW trunc.	Bias	.108	.193	.097	.166	.193	.249
	Rel.Bias	na	na	1.036	1.241	.449	.420
	MSE	.022	.056	.023	.048	.060	.091
Seq. *g*-est.	Bias	.070	.125	.063	.107	.128	.164
	Rel.Bias	na	na	.663	.791	.296	.275
	MSE	.008	.024	.011	.021	.030	.044
R. seq. *g*-est. I	Bias	.082	.149	.073	.126	.152	.195
	Rel.Bias	na	na	.775	.926	.352	.327
	MSE	.012	.036	.014	.031	.040	.063
R. seq. *g*-est. II	Bias	.102	.335	.096	.327	.139	.351
	Rel.Bias	na	na	1.145	2.801	.324	.619
	MSE	.017	.121	.022	.118	.044	.147
R. seq. *g*-est. III	Bias	.137	.378	.128	.365	.200	.405
	Rel.Bias	na	na	1.480	3.093	.466	.713
	MSE	.026	.153	.030	.146	.064	.190

Note: Regression: linear regression adjusting for covariates; IPW: inverse propensity weighting; IPW trunc.: IPW with truncated weights; Seq. g-est: sequential g-estimation; R. seq. g-est. I: doubly robust sequential g-estimation with one-confounder estimation mediator model; R. seq. g-

est. II: doubly robust sequential g-estimation with one-confounder estimation outcome model; R. seq. g-est. III: doubly robust sequential g-estimation with one-confounder estimation mediator and outcome models.

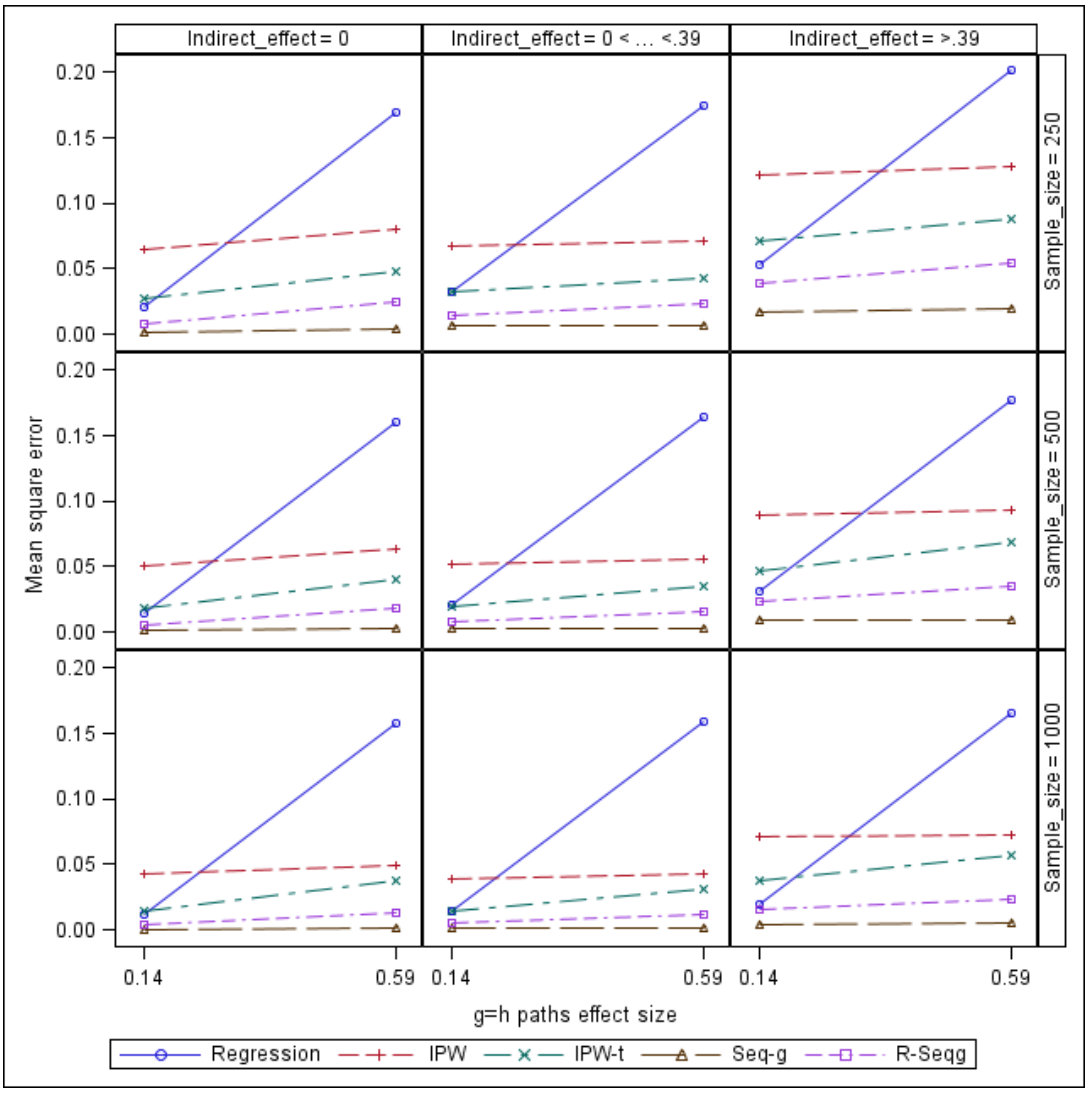


Figure 31. Indirect effect *c-c'* mean square error by *g* and *h* paths effect size and sample size - two-confounders estimation model

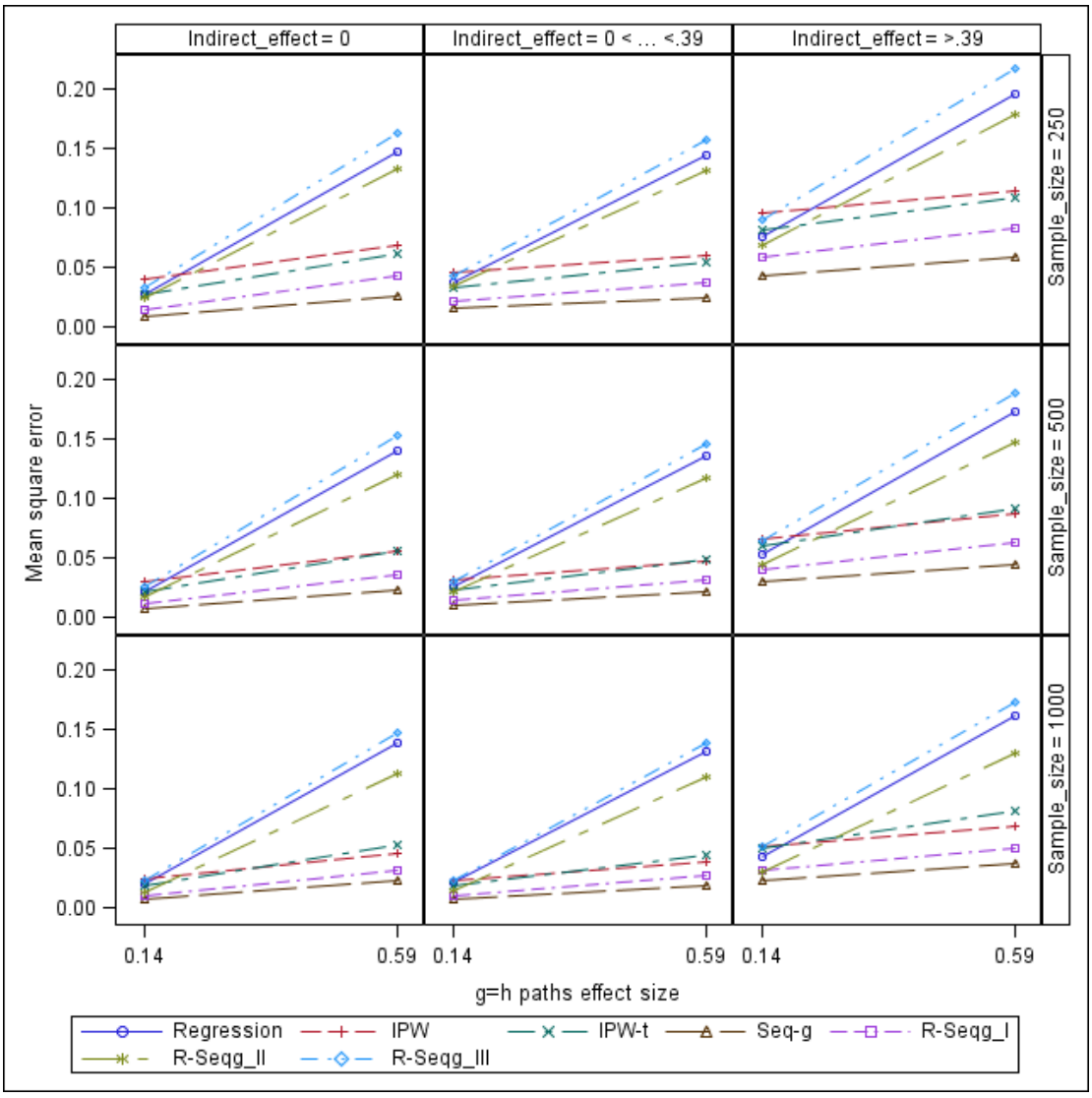


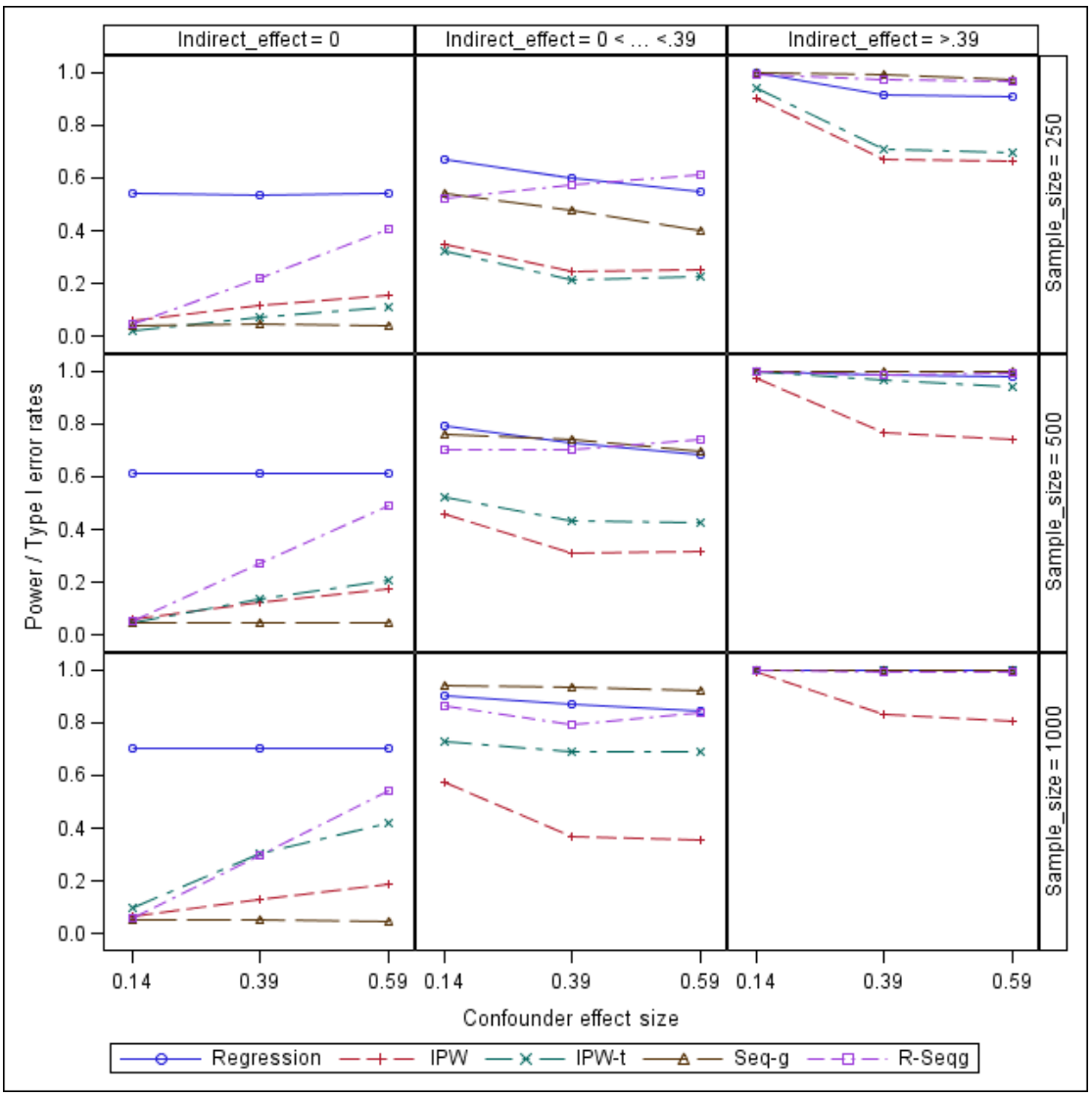
Figure 32. Indirect effect *c-c'* mean square error by *g* and *h* paths effect size and sample size - one-confounder estimation model

3.2.2 Statistical power and Type I error rates

The liberal criterion of [.025, .075] was used to evaluate Type I error rates for the indirect effect (Bradley, 1978). Statistical power values were evaluated against the nominal .80 criterion (Cohen, 1988). Again, the statistical power was interpreted only for the conditions in which the Type I error rates were acceptable.

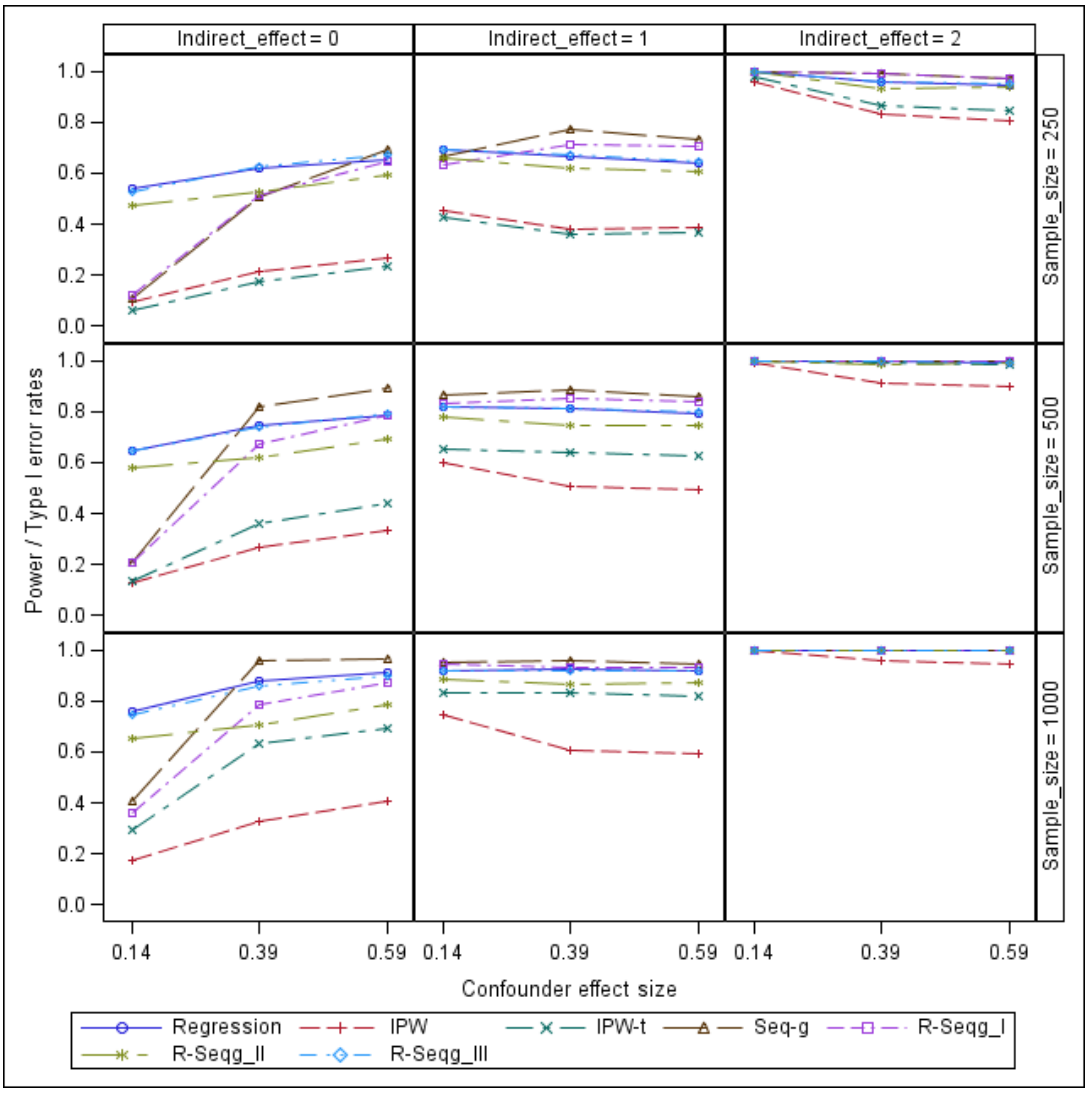
Figures 33 and 34 investigate the relation between confounder effect size and the statistical power and Type I error rates. In the case of two-confounders estimation models, Type I error rates for the linear regression were out of bounds across confounder effect size conditions and thus the statistical power results for that method were not interpreted. Sequential g-estimation had good Type I error rates around the nominal value .05 and good statistical power across conditions. Both IPW methods had acceptable Type I error rates only when the confounder effect size was .14. When the confounder effect size was .14, IPW methods had power greater than .80 when the indirect effect estimate was greater than .39. The doubly robust g-estimation again followed the same pattern as the IPW method and had good Type I error rates only when the confounder effect size was .14 and good statistical power for that condition when the indirect effect was greater than .39.

In the case of one-confounder estimation models in Figure 34, Type I error rates for the indirect effect were greater than .075 for all methods across conditions; thus the power results were not interpreted.



Note1: Values represent Type I error rates when *c-c'* =0.
Note2: Regression: linear regression adjusting for covariates; IPW: inverse propensity weighting; IPW-t.: IPW with truncated weights; Seq-g: sequential g-estimation; R-Seqg: robust sequential g-estimation.

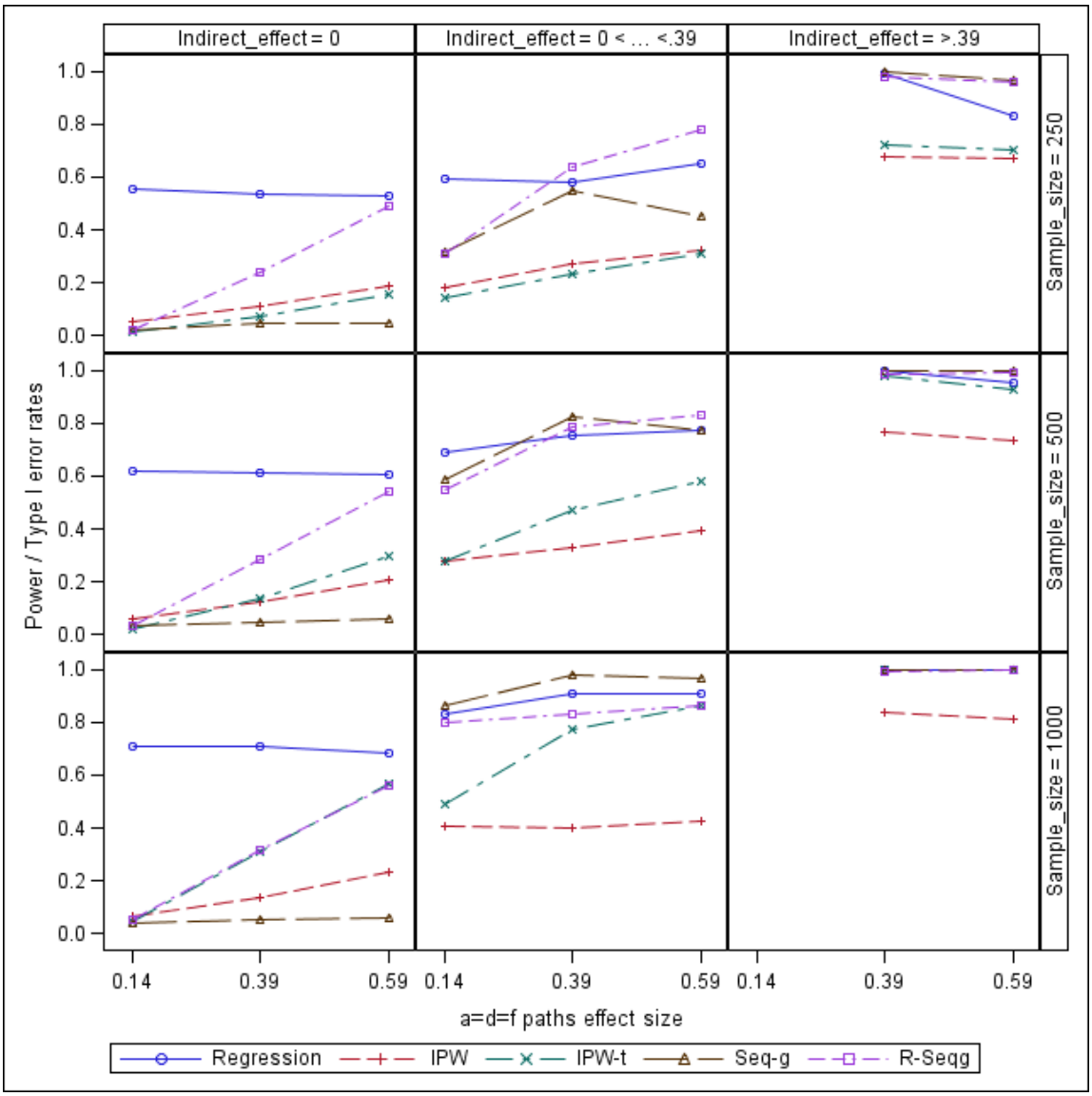
Figure 33. Indirect effect *c-c'* power and Type I error rates by post-treatment confounders effect size and sample size - two-confounders estimation model



Note1: Values represent Type I error rates when *c*-*c′* =0.
Note2: Regression: linear regression adjusting for covariates; IPW: inverse propensity weighting; IPW-t.: IPW with truncated weights; Seq-g: sequential g-estimation; R-Seqg_I: doubly robust sequential g-estimation with one-confounder estimation mediator model; R-Seqg_II: doubly robust sequential g-estimation with one-confounder estimation outcome model; R-Seqg_III: doubly robust sequential g-estimation with one-confounder estimation mediator and outcome models.

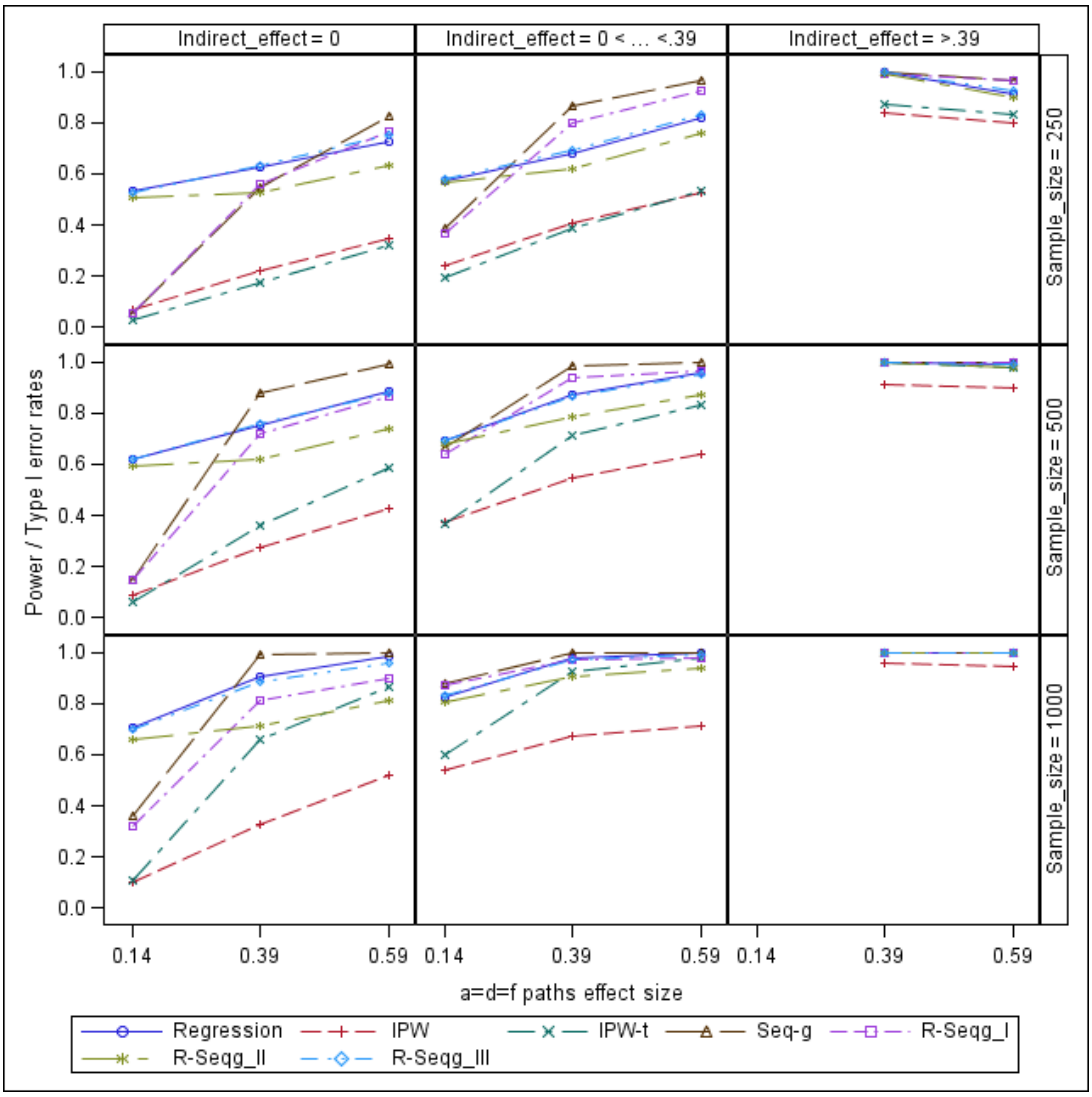
Figure 34. Indirect effect *c*-*c′* power and Type I error rates by post-treatment confounders effect size and sample size - one-confounder estimation model

Figures 35 and 36 investigate if the IPW method performs better when the *a* path and confounder effect sizes are not extreme and if the truncated IPW method has a better performance than the conventional IPW method. In the case of both two-confounders and one-confounder estimation models, both IPW and IPW-truncated methods had acceptable Type I error rates when the *a* path and confounder effect sizes were equal to .14. It would have been expected that when the *a* path and confounder effect sizes were .14, both IPW methods would have power greater than .80 when the indirect effect was greater than .39; however I did not have that simulation condition available to interpret.



Note1: Values represent Type I error rates when *c-c'* =0.
Note2: Regression: linear regression adjusting for covariates; IPW: inverse propensity weighting; IPW-t.: IPW with truncated weights; Seq-g: sequential g-estimation; R-Seqg: robust sequential g-estimation.

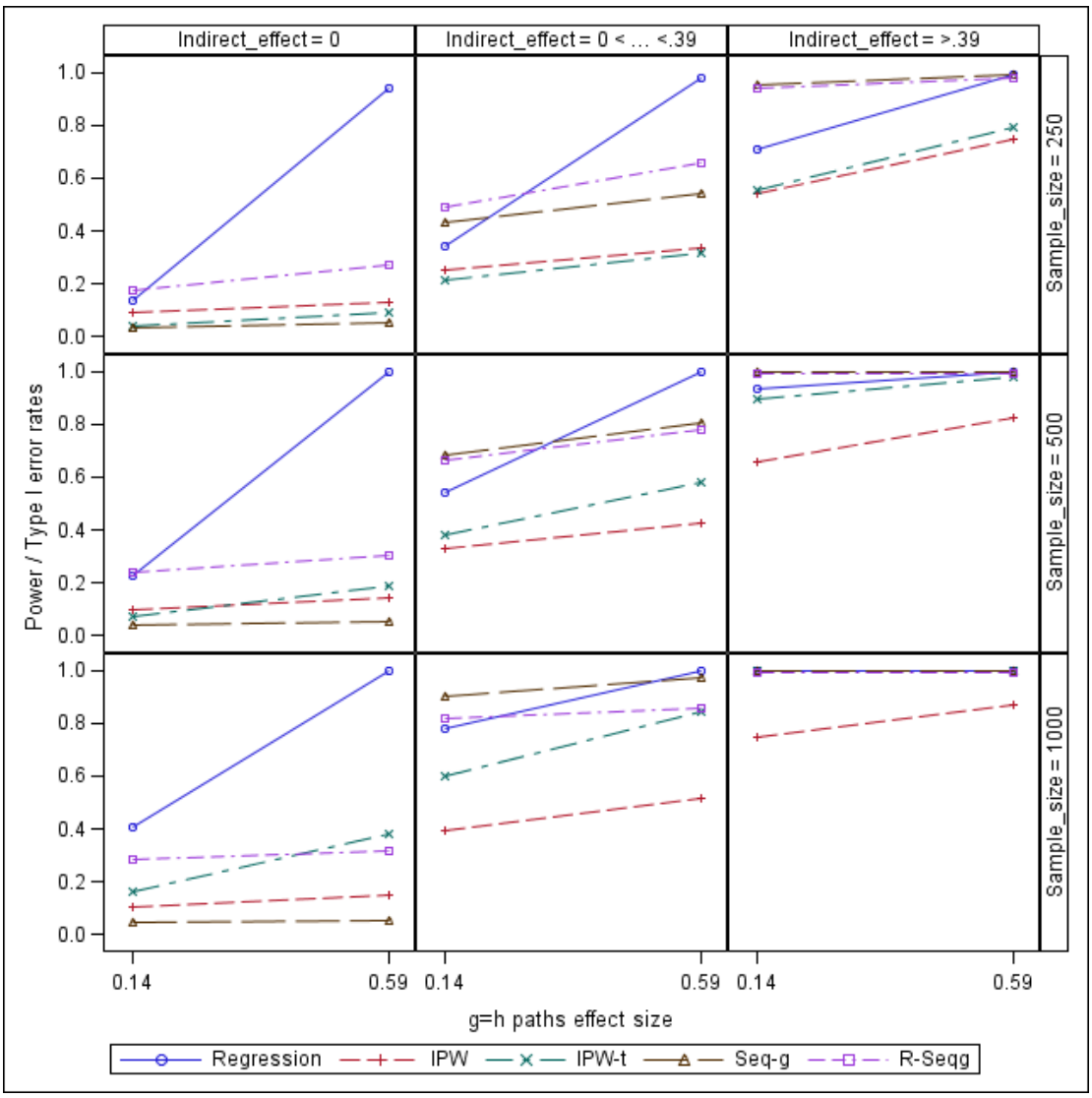
Figure 35. Indirect effect *c-c'* power and Type I error rates by *a*, *d*, and *f* paths effect size and sample size - two-confounders estimation model



Note1: Values represent Type I error rates when *c*-*c'* =0.
Note2: Regression: linear regression adjusting for covariates; IPW: inverse propensity weighting; IPW-t.: IPW with truncated weights; Seq-g: sequential g-estimation; R-Seqg_I: doubly robust sequential g-estimation with one-confounder estimation mediator model; R-Seqg_II: doubly robust sequential g-estimation with one-confounder estimation outcome model; R-Seqg_III: doubly robust sequential g-estimation with one-confounder estimation mediator and outcome models.

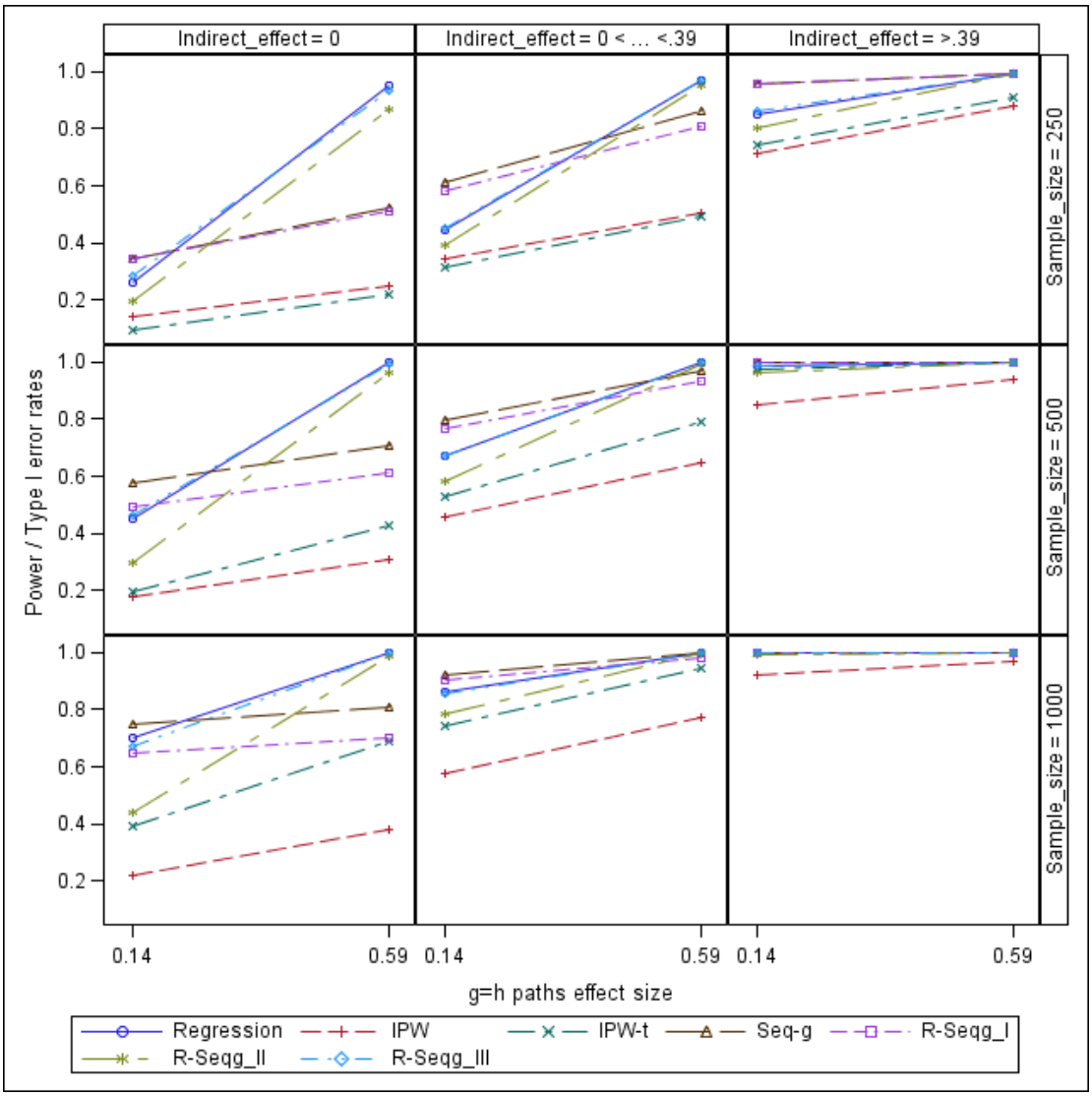
Figure 36. Indirect effect *c*-*c'* power and Type I error rates by *a*, *d*, and *f* paths effect size and sample size - one-confounder estimation model

Figures 37-38 explore whether the Type I error rates and statistical power of the indirect effect are influenced by the effect of the treatment on confounder (*g* and *h* paths). In the case of two-confounders estimation models, Type I error rates of linear regression increased as the effect size of the *g* and *h* paths increased; for all other methods, Type I error rates were not influenced by the size of the *g* and *h* paths. Linear regression had unacceptable Type I error rates and thus its statistical power was not interpreted. IPW methods had acceptable Type I error rates when the confounder effect size was equal to .14 for the sample sizes 250 and 500. Yet, only the truncated-IPW method had power greater than .80 for those conditions. In the case of one-confounder model shown in Figure 38, none of the methods had acceptable Type I error rates across conditions.



Note1: Values represent Type I error rates when *c-c'* =0.
Note2: Regression: linear regression adjusting for covariates; IPW: inverse propensity weighting; IPW-t.: IPW with truncated weights; Seq-g: sequential g-estimation; R-Seqg: robust sequential g-estimation.

Figure 37. Indirect effect *c-c'* power and Type I error rates by *g* and *h* paths effect size and sample size - two-confounders estimation model



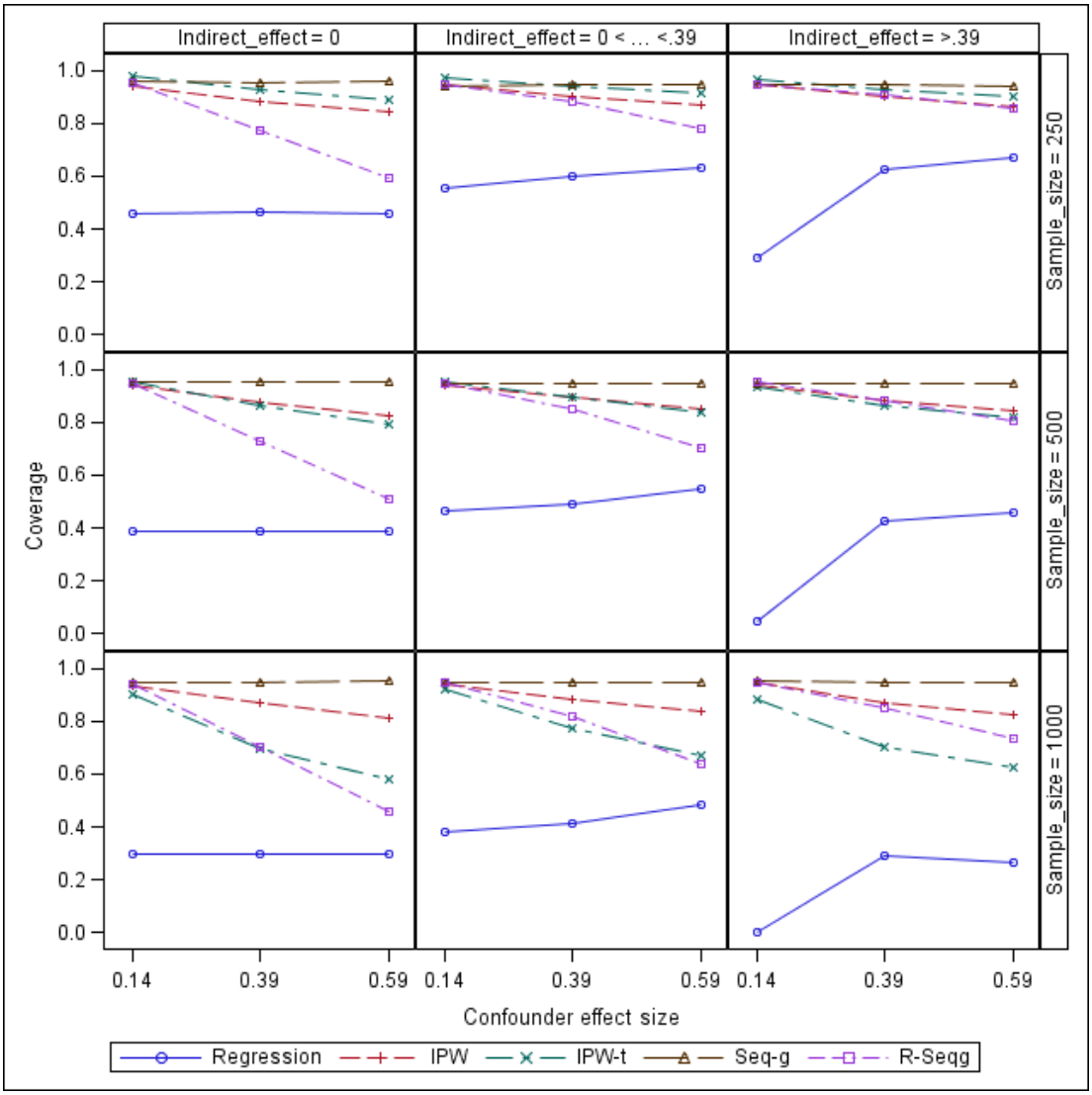
Note1: Values represent Type I error rates when c-c' =0.
Note2: Regression: linear regression adjusting for covariates; IPW: inverse propensity weighting; IPW-t.: IPW with truncated weights; Seq-g: sequential g-estimation; R-Seqg_I: doubly robust sequential g-estimation with one-confounder estimation mediator model; R-Seqg_II: doubly robust sequential g-estimation with one-confounder estimation outcome model; R-Seqg_III: doubly robust sequential g-estimation with one-confounder estimation mediator and outcome models.

Figure 38. Indirect effect c-c' power and Type I error rates by g and h paths effect size and sample size - one-confounder estimation model

3.2.3 Confidence interval coverage

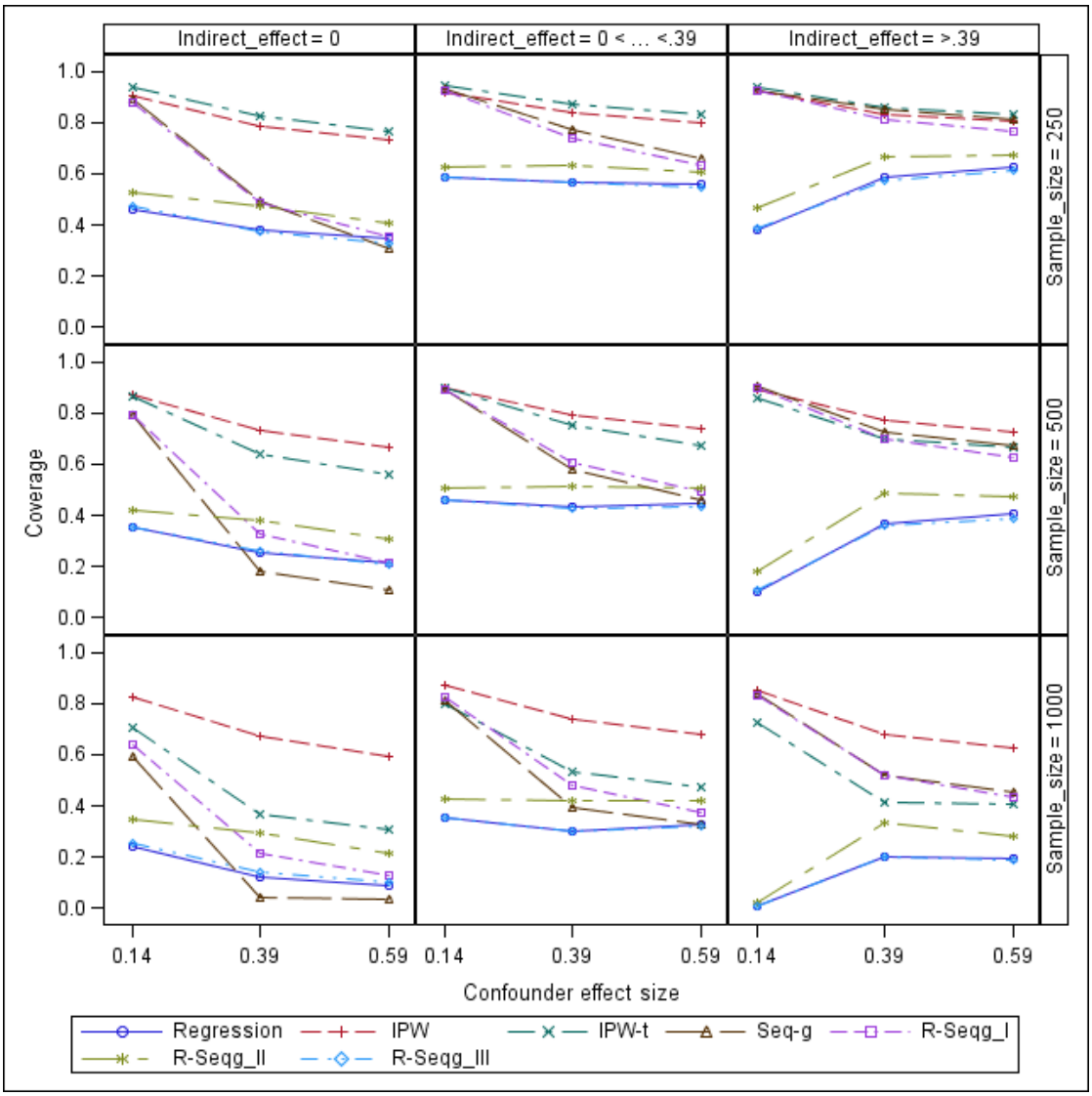
Figures 39-40 investigate if the confidence interval coverage of the indirect effect estimates is influenced by the size of the confounder effect (the *d* and *f* paths). For the two-confounders estimation models, linear regression did not have coverage greater .90 in any of the simulation conditions. For all other methods except for sequential g-estimation, coverage decreases as the size of the *d* and *f* paths increase. For the IPW, IPW-truncated, and the doubly robust g-estimation methods, coverage was greater than .90 only when the confounder effect size was equal to .14. Coverage of the sequential g-estimation was robust to the size of the confounder effect and greater than .90 across conditions.

For the one-confounder estimation models in Figure 40, linear regression again did not have coverage greater than .90 in any of the conditions. IPW methods, sequential g-estimation and doubly robust sequential g-estimation with one-confounder estimation mediator model had coverage greater than .90 only when the confounder effect size was equal to .14 for the sample sizes 250 and 500. The doubly robust sequential g-estimation methods with one-confounder estimation outcome model and one-confounder estimation mediator and outcome model did not reach coverage greater than .90 in any conditions.



Note: Regression: linear regression adjusting for covariates; IPW: inverse propensity weighting; IPW-t.: IPW with truncated weights; Seq-g: sequential g-estimation; R-Seqg: robust sequential g-estimation.

Figure 39. Indirect effect *c-c'* confidence interval coverage by post-treatment confounders effect size and sample size - two-confounders estimation model

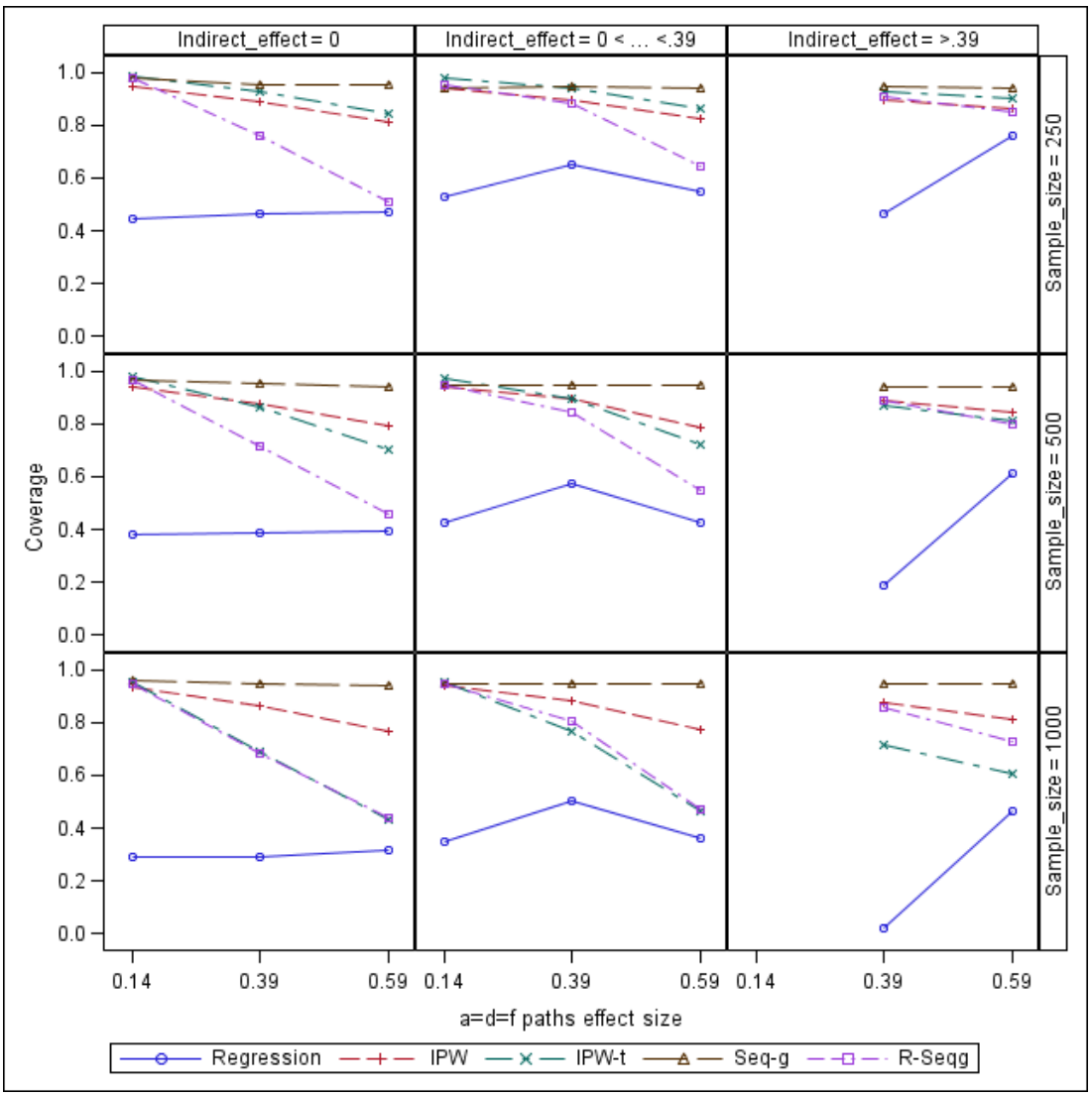


Note: Regression: linear regression adjusting for covariates; IPW: inverse propensity weighting; IPW-t.: IPW with truncated weights; Seq-g: sequential g-estimation; R-Seqg_I: doubly robust sequential g-estimation with one-confounder estimation mediator model; R-Seqg_II: doubly robust sequential g-estimation with one-confounder estimation outcome model; R-Seqg_III: doubly robust sequential g-estimation with one-confounder estimation mediator and outcome models.

Figure 40. Indirect effect *c-c'* confidence interval coverage by post-treatment confounders effect size and sample size - one-confounder estimation model

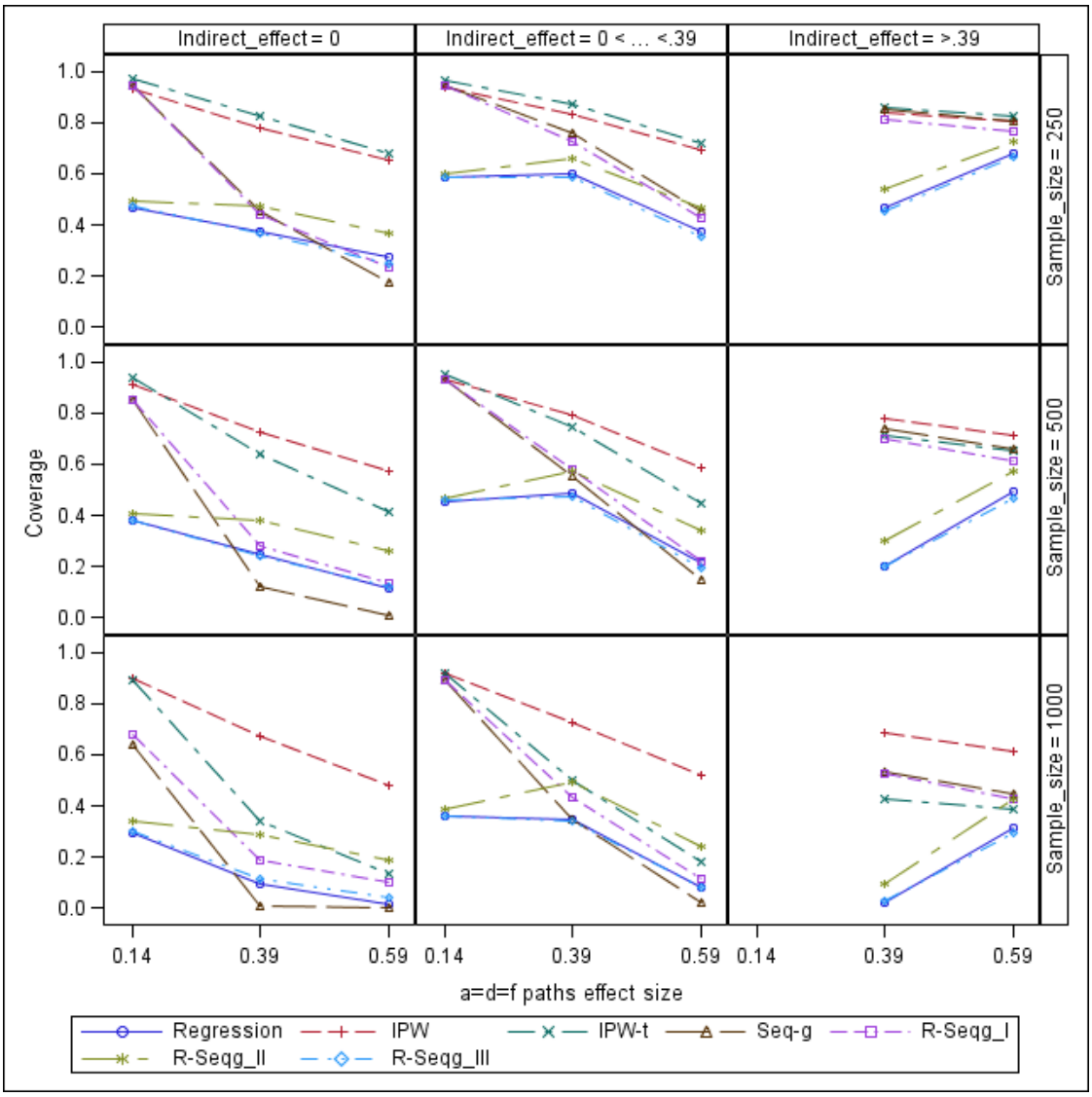
Figures 41 and 42 investigate whether the confidence interval coverage is influenced by the confounder effect size and the *a* path effect size. For the two-confounders estimation models, linear regression did not have coverage greater .90 in any of the simulation conditions. For all other methods except for sequential g-estimation, coverage decreased as the size of the *d* and *f* paths increased. For the IPW, IPW-truncated, and the doubly robust g-estimation methods, coverage was greater than .90 only when the confounder effect size was equal to .14. The truncated IPW method did not perform better than the conventional IPW method in terms of coverage. Coverage of the sequential g-estimation was robust to the size of the confounder effect and greater than .90 across conditions.

For the one-confounder estimation models in Figure 42, linear regression did not have coverage greater than .90 in any of the conditions. IPW methods, sequential g-estimation and doubly robust sequential g-estimation with one-confounder estimation mediator model had coverage greater than .90 only when the confounder effect size was equal to .14 for the sample sizes 250 and 500. The coverage results for the truncated IPW method were similar to the conventional IPW methods. The doubly robust sequential g-estimation methods with one-confounder estimation outcome model and one-confounder estimation mediator and outcome model did not reach coverage greater than .90 in any conditions.



Note: Regression: linear regression adjusting for covariates; IPW: inverse propensity weighting; IPW-t.: IPW with truncated weights; Seq-g: sequential g-estimation; R-Seqg: robust sequential g-estimation.

Figure 41. Indirect effect *c-c'* confidence interval coverage by *a*, *d*, and *f paths* effect size and sample size - two-confounders estimation model

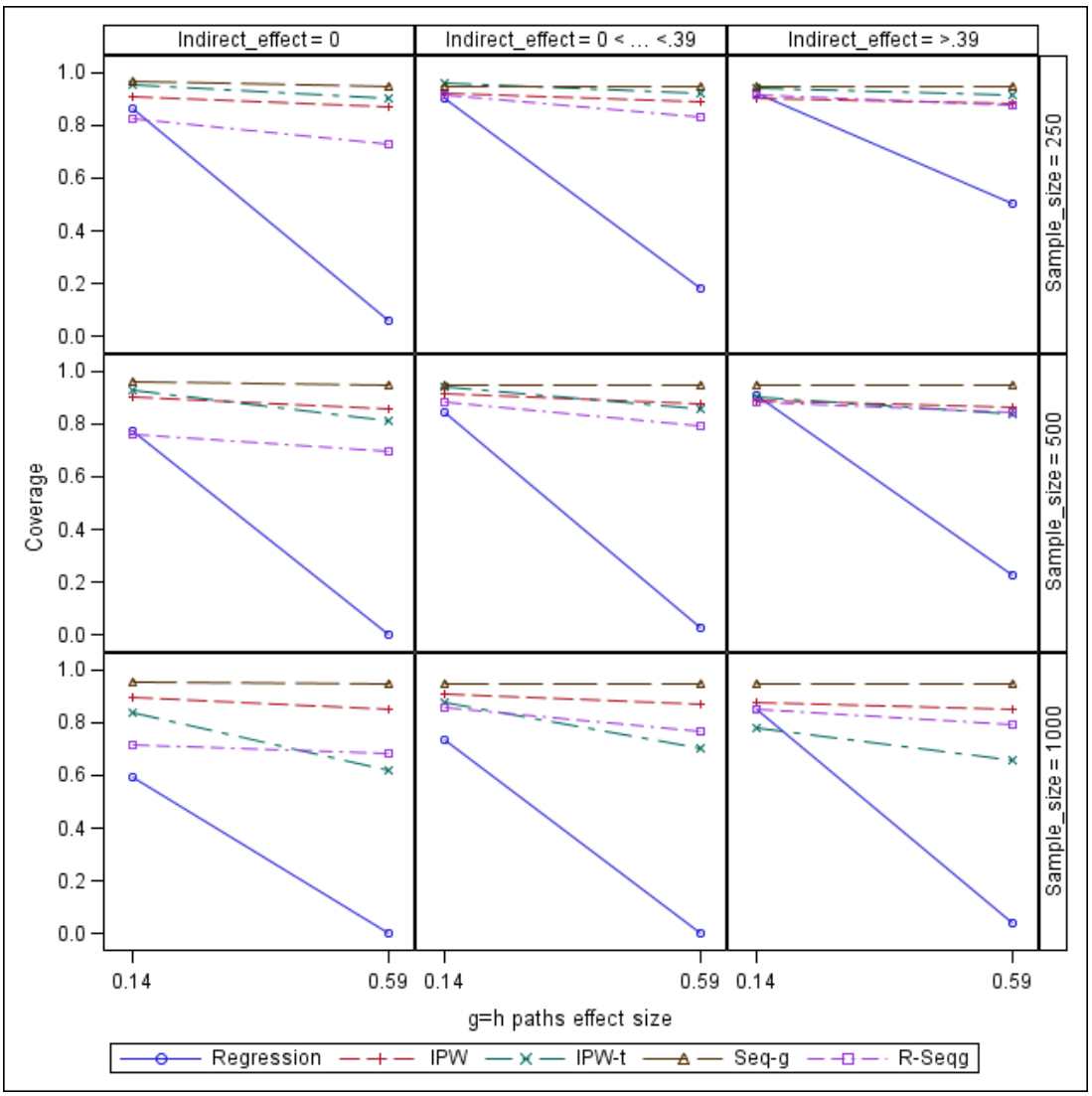


Note: Regression: linear regression adjusting for covariates; IPW: inverse propensity weighting; IPW-t.: IPW with truncated weights; Seq-g: sequential g-estimation; R-Seqg_I: doubly robust sequential g-estimation with one-confounder estimation mediator model; R-Seqg_II: doubly robust sequential g-estimation with one-confounder estimation outcome model; R-Seqg_III: doubly robust sequential g-estimation with one-confounder estimation mediator and outcome models.

Figure 42. Indirect effect *c-c'* confidence interval coverage by *a*, *d*, and *f* paths effect size and sample size - one-confounder estimation model

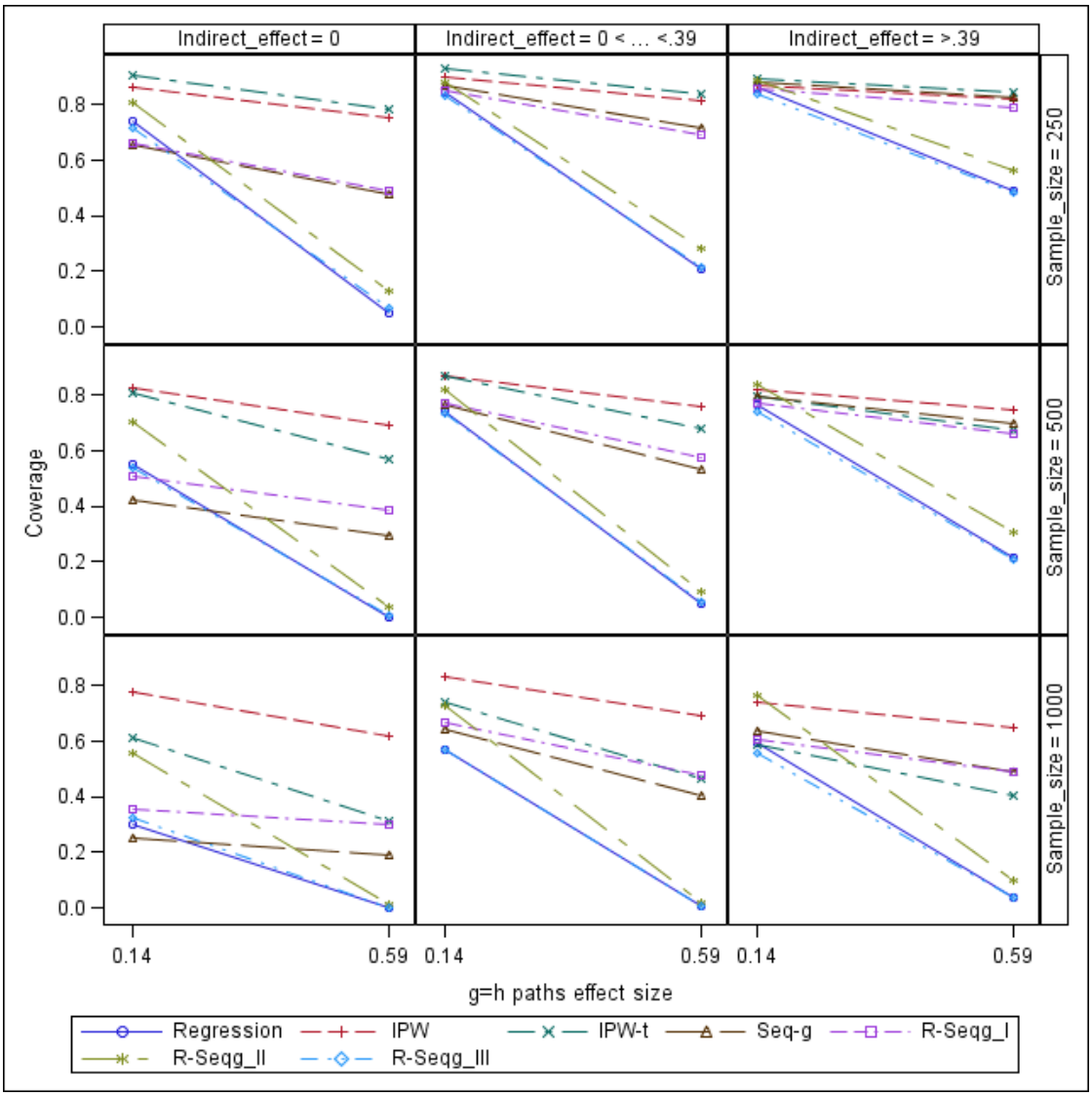
Figures 43-44 investigate if the confidence interval coverage of the indirect effect estimates is influenced by how much the confounder is influenced by the treatment (the *g* and *h* paths). For the two-confounders estimation models, the coverage of the linear regression method decreases sharply when the size of the *g* and *h* paths increase. The coverage of the linear regression method reaches .90 for the sample sizes 250 and 500 when the indirect effect is greater than .39 and the effect size for relation between the treatment and confounder is equal to .14. The coverage of the IPW, truncated-IPW, and robust g-estimation methods is greater than .90 for the sample sizes 250 and 500 when the effect of the treatment on confounders is .14. Coverage of the sequential g-estimation was robust to the size of the confounder effect and greater than .90 across conditions.

For the one-confounder estimation models in Figure 44, coverage of the linear regression, the doubly robust sequential g-estimation methods with one-confounder estimation outcome model and one-confounder estimation mediator and outcome models decreased sharply as the effect of the treatment on the confounders increased and the methods did not have coverage greater than .90 in any of the conditions. The IPW methods had coverage greater than .90 only for the sample size of 250. Similarly, sequential g-estimation reached coverage greater than .90 only for the sample size of 250 in the case of an indirect effect greater than zero.



Note: Regression: linear regression adjusting for covariates; IPW: inverse propensity weighting; IPW-t.: IPW with truncated weights; Seq-g: sequential g-estimation; R-Seqg: robust sequential g-estimation.

Figure 43. Indirect effect *c-c'* confidence interval coverage by *g* and *h* paths effect size and sample size - two-confounders estimation model



Note: Regression: linear regression adjusting for covariates; IPW: inverse propensity weighting; IPW-t.: IPW with truncated weights; Seq-g: sequential g-estimation; R-Seqg_I: doubly robust sequential g-estimation with one-confounder estimation mediator model; R-Seqg_II: doubly robust sequential g-estimation with one-confounder estimation outcome model; R-Seqg_III: doubly robust sequential g-estimation with one-confounder estimation mediator and outcome models.

Figure 44. Indirect effect *c-c'* confidence interval coverage by *g* and *h* paths effect size and sample size - one-confounder estimation model

Chapter 4

DISCUSSION

Mediation analysis is widely used in social and medical sciences for its key benefits to test and develop theory. The goal of mediation analysis is to investigate the causal mechanisms of a phenomenon. The purpose of this dissertation was to describe the causal inference challenge that mediation analysis confronts and describe solutions to overcome it. Randomized studies are accepted as the best approach to infer causality for X to Y relation. Yet, randomized X is not sufficient in the presence of a post-randomization variable such as a mediator for causal inference about mediation because potential confounders may exist for the M to Y relation. Recent research in mediation has led to many advanced methods for estimating causal mediator effects, among those are inverse propensity weighting and g-estimation that have been reviewed in this dissertation. The purpose of this dissertation was (1) to investigate how much traditional mediation methods are affected by confounding variables and (2) to assess the statistical performance of modern methods to address confounding variable effects in mediation analysis. A large simulation study was designed to evaluate how robust the OLS and causal inference methods estimators of direct and indirect effects are to different confounder effect sizes and to the violation of no omitted variables assumption in terms of bias, statistical power and confidence interval coverage.

4.1 Summary and discussion of results

To test the research question of how robust the methods are to different confounder effect sizes, models including the two post-treatment confounders were

estimated for each simulation condition across different mediation path effect sizes, confounder effect sizes and sample sizes (i.e., two-confounders estimation models). Results for the direct and indirect effects indicated that bias increased as the effect of the post-treatment confounder on the mediator increased for all methods except for linear regression and sequential g-estimation.

In general sequential g-estimation was the best method compared to other methods in terms of bias, power and coverage. This finding was expected in this simulation study because the sequential g-estimation method is specifically designed to handle models with post-treatment confounders whereas linear regression and IPW methods are not. This study shows that that sequential g-estimation performs well in general with relative bias less than .10, Type I error rates around the .05 nominal value, power greater than .80 and coverage greater than .90 across different confounder effect size, direct effect size, and sample size conditions.

The linear regression performance was poor in general, as the method had high bias, Type I error rates that were out of bounds, low statistical power and coverage. The main factor that influenced the performance of the linear regression method was the size of the relation between the treatment and the confounders. The bias of the linear regression estimates increased drastically as the effect of the treatment on the confounders increased showing that linear regression was failing when the confounders were post-treatment. This study did not have the condition in which the effect of the treatment on confounders was equal to zero, but I would expect linear regression to perform much better if the confounders of the M to Y relation were not post-treatment.

As a side note, I ran one condition where all confounder effects were equal to zero in order to check the simulation code, and in that situation all methods including linear regression were unbiased across conditions.

The IPW method's performance was mainly influenced by the confounder effect size. In general, it had acceptable bias, power, and coverage when the confounder effect size was small. In this study, I used a weighting strategy for the propensity score approach to causal mediation analysis; however other strategies such as matching or stratification may have performed better (Rubin & Thomas, 1992; Rosenbaum & Rubin, 1984). Even though there is no study showing how matching would perform in the case of mediation analysis, propensity score studies addressing the X to Y relation indicate that matching works better than weighting to achieve unbiased causal estimates (Frolich, 2004). Weighting estimates can be problematic since the estimates can be highly influenced by the assigned weights of individuals with propensity scores that are close to values 0 or 1 (Kang & Schafer, 2007; Schafer & Kang, 2008). In order to avoid extreme weights, the IPW method with truncated weights was also included in this dissertation. However, results showed that in general, the IPW-truncated method did not perform better than the conventional IPW method. This finding may be due to the trimming rule used in this study (weights were trimmed at the 1^{st} and 99^{th} percentile of the weight distribution) and some other trimming strategies may yield better results. Simulation studies point out that trimming may optimize propensity score weights by decreasing variability in the weights, the optimal level of trimming may be difficult to determine and may have adverse effects in estimation (Lee et al., 2011; Freedman & Berk 2008).

The pattern of results for the doubly robust sequential g-estimation method's performance was in general similar to the IPW method rather than sequential g-estimation. The doubly robust g-estimation method employed in this dissertation used IPW estimation, and results suggest that the performance of the doubly robust method was influenced heavily by the IPW part of the estimation. Doubly robust g-estimation method has been suggested in the literature as a superior method to g-estimation but has never been tested in simulation studies, making this finding important for researchers.

To test the research question of how robust the methods are to the omission of potential confounders, the OLS regression and causal inference methods were applied to the generated data by omitting one of the confounders from the estimation (i.e., the one-confounder estimation model). In general, none of the methods had acceptable relative bias in the simulation study. Omitting one of the confounders from estimation corresponded to the common case in mediation studies where no measure of a confounder is available but a confounder may affect the analysis. Failing to measure potential post-treatment confounder variables in a mediation model leads to biased estimates regardless of the analysis method used and emphasize the importance of sensitivity analysis for causal mediation analysis.

For the doubly robust sequential g-estimation method, three one-confounder estimation models were obtained: (a) by omitting confounder C_2 in only the mediator propensity model, (b) by omitting confounder C_2 in only the outcome model, (c) by omitting confounder C_2 in both the mediator propensity and outcome models. This allowed for testing if the doubly robust method fails when both parts or one part of the

estimated model omits the confounder C_2. None of the three misspecifications for the doubly robust method had in general acceptable relative bias or Type I error rates, and the doubly robust method had the highest bias when both parts of its estimation process omitted the confounder C_2. This finding was consistent with the warnings from the literature on the use of doubly robust methods (Schafer & Kang, 2008).

4.2 Recommendations

Based on the current study, the following recommendations for applied researchers can be offered.

1. Researchers should carefully consider the potential confounder variables for their mediation model when designing the study and make an effort to measure the confounder variables. Failing to accommodate confounder variables in a mediation model leads to biased estimates of the direct and indirect effects.

2. It is important to identify the types of confounders to choose the analysis method to be implemented. In a mediation analysis, there may exist M to Y confounders. The simulation study in this dissertation shows that when these confounders are influenced by the treatment, sequential g-estimation produces unbiased estimates.

3. Each method has a distinct specification of the direct and indirect effects that may not directly translate to the linear regression approach to mediation because the methods differ on how they control for confounders and assumptions. Researchers should pay attention to which effects they are interested in estimating and the assumptions made by the analysis method they choose. For instance, the IPW method has assumptions such

as no XM interaction on Y that is not tested in the current simulation study, but its presence can bias estimates (Coffman & Zhong, 2012).

4. When using IPW methods, researchers need to be careful about extreme weights. In the case of applying weight trimming to avoid variability in the weight distribution, they should evaluate which trimming option best suits their data using evaluation criteria such as least mean square error (Potter, 1993). However literature suggests that researchers rather focus on correct specifications of their propensity score model rather than relying on trimming methods (Lee et al., 2011).

5. Sensitivity analysis methods are highly recommended to evaluate how robust the mediated effect is to unmeasured third variables because researchers usually fail to assess all potential confounders. Sensitivity analysis has been an important area of research to improve causality in treatment effects when randomization has not been possible (Rosenbaum & Rubin, 1983). For example, Cornfield et al. (1959) found that the relationship between smoking and lung cancer can be significantly weakened if a confounder variable for that relationship would be nine times more frequent in heavy smokers compared to nonsmokers. Consequently, one can even argue that a statistical analysis is not complete without sensitivity analysis (Imai et al., 2010). Current literature suggests several sensitivity analysis methods for mediation analysis (Cox, Kisbu-Sakarya, Miocevic, & MacKinnon, 2013). For example, an approach described by VanderWeele (2010) is based on the relation of the confounder to Y and the difference in proportion of persons with the confounder prevalence between treatment groups at the same level of the mediator. Another method presents confounder bias as correlated error

terms between the error in the mediator model and the error in the outcome model (Imai, Keele, & Yamamoto, 2010). The Imai et al. and VanderWeele methods both use potential outcomes definitions of mediated effects as described by Robins & Greenland (1992) and Pearl (2001; 2012). A third method by Mauro (1990) is based on the correlations of a potential confounder with study variables.

6. In addition to the quantitative methods to deal with the sequential ignorability assumption, researchers can also improve their results by using several research designs that attempt to manipulate the mediator when it is both practically and ethically possible. Examples of these experimental designs are the enhancement design where exposure to the mediator is manipulated by enhancing the dose of the mediator, and the blockage design in which the mediator is blocked in one condition but not in another condition to investigate if the effect of the treatment depends on the mediator (Robins & Greenland, 1992; Imai, Tingley, & Yamamoto, 2013; MacKinnon & Pirlott, 2013; Spencer, Zanna, & Fong, 2005; Bullock, Green, & Ha, 2010). These experimental designs can be combined with quantitative methods and sensitivity analysis to improve causal estimation of direct and indirect effects.

4.3 Limitations

The present research has a number of limitations. A first limitation was that I did not have a simulation condition where the effect of the treatment on the confounders was equal to zero (i.e., the confounders in the simulation were all post-treatment). This condition was not included because I needed to limit the number of conditions because of time considerations (e.g., the simulation was using bootstrapping technique that took a

considerable amount of time to run). The study was also focused on the performance of the methods in the case of post-treatment confounders. The results of the study showed a big effect of the size of the relation between the treatment and confounders on linear regression estimates. Therefore, it would be important to see how methods would perform when confounders were not influenced by the treatment.

A second limitation of the simulation is that the results are limited to linear models. All effects in the model were simulated to be linear, although situations such as nonlinear confounder effects may happen with real data. Another limitation was that parameter values to generate the data were not based on Cohen's small, medium, and large effect values. The parameter values for mediation paths were selected based on published papers for single mediator models without confounders, and the parameter values for confounder paths were then selected to be consistent with the values of the mediation paths. Attempts have been made to determine small, medium and large effect size values for the parameters by using covariance algebra; yet the large number of variables in the model made it difficult to choose the best combination of values.

4.4 Conclusions and future directions

The current study has shown that in the case of post-treatment confounders sequential g-estimation performs the best and other methods are heavily influenced by confounder effect size and how much the confounders are influenced by the treatment. Moreover, this dissertation demonstrated that failing to measure pot-treatment confounders of the M to Y relation may lead to bias in estimates.

Causal inference in the presence of mediating variables is an important area of research with recent advancements, yet there is need for future work on the investigation of the methods. The current study as well as recent studies in causal mediation analysis focuses on single mediator models. Future research should extent this framework to more complex situations such as multiple mediator, longitudinal and multilevel mediation models. For instance, it is assumed that there is no interference between study participants; thus the violation of this independence assumption should be investigated by studying multilevel models of causal mediation analysis. Even though applications in multilevel causal modeling exist, there is considerable need for analytical work and simulation studies (VanderWeele, 2010b; Hong & Raudenbush, 2006; Hong, 2010). Also, as stated in the study limitations, investigating nonlinear relations is an important area of research for causal mediation analysis. Future work should also focus more on sensitivity analysis methods. For instance, research shows that for the propensity score weighting approach, bias in the indirect effect increases as the number of confounders included in the model decreases (Coffman, 2011). Yet, a sensitivity analysis for the propensity score weighting method to assess the robustness of results to the number of confounders included in the propensity model is lacking.

In this dissertation, I focused on the quantitative methods to analyze mediation models. However, another important area of work for causal mediation is the development of alternative experimental designs in which researchers manipulates the mediator as described above. There has been limited literature on new experimental designs for mediation. Future work is crucial on evaluating the advantages and

disadvantages of the proposed designs, clarifying the assumptions, and developing and illustrating the analysis of such designs. Such work in both experimental and quantitative approaches to mediation would encourage substantive researchers to apply causal mediation methods to real data. This dissertation has explored quantitative methods to improve causal inference in mediation analysis and has shown that causal analysis is critical in the presence of post-treatment confounders.

REFERENCES

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.

Bradley J.V. (1978). Robustness? *British Journal of Mathematical and Statistical Psychology, 31*, 144–152.

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.

Claessens, B. J., Van Eerde, W., Rutte, C. G., & Roe, R. A. (2004). Planning behavior and perceived control of time at work. *Journal of Organizational Behavior*, *25*, 937-950.

Coffman, D. L. (2011). Estimating causal effects in mediation analysis using propensity scores. *Structural Equation Modeling, 18*, 357-369.

Coffman, D. L., & Zhong, W. (2012). Assessing Mediation Using Marginal Structural Models in the Presence of Confounding and Moderation. *Psychological Methods*, *17*, 642.

Cohen J.(1988). *Statistical power analysis for the behavioral sciences*. Lawrence Erlbaum Associates: Hillsdale, NJ.

Cohen, J., Cohen, P., West, S.G., & Aiken, L.S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences.* Lawrence Erlbaum Associates: New Jersey.

Cole, S. R., & Hernán, M. A. (2008). Constructing inverse probability weights for marginal structural models. *American journal of epidemiology, 168*, 656-664.

Collins, L. M., Schafer, J. L., & Kam, C. M. (2001). A comparison of inclusive and restrictive strategies in modern missing data procedures. *Psychological Methods, 6,* 330–351.

Cornfield, J., Haenszel, W., Hammond, E.C., Lilien-Feld, A.M., Shimkin, M.B., & Wynder, E.L. (1959). Smoking and lung cancer: Recent evidence and a discussion of some questions. *Journal of the National Cancer Institute, 22*, 173- 203.

Cox, M. G., Kisbu-Sakarya, Y., Miocevic, M., &MacKinnon, D.P. (2013). Sensitivity plots for confounder bias in the single mediator model. Manuscript under review.

Davidson, O. B., & Eden, D. (2000). Remedial self-fulfilling prophecy: Two field experiments to prevent Golem effects among disadvantaged women. *Journal of Applied Psychology, 85,* 386-398.

Elovainio, M., Ferrie, J. E., Gimeno, D., De Vogli, R., Shipley, M., Brunner, E. J., Kumari, M., Vahtera, J., Marmot, M.G., & Kivimäki, M. (2009). Organizational justice and sleeping problems: the Whitehall II study. *Psychosomatic Medicine, 71,* 334-340.

Flora, D. B., & Curran, P. J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods, 9,* 466–491.

Freedman, D. A., & Berk, R. A. (2008). Weighting regressions by propensity scores. *Evaluation Review, 32,* 392-409.

Fritz, M. S., & MacKinnon, D. P. (2007). Required sample size to detect the mediated effect. *Psychological Science, 18,* 233-239.

Frölich, M. (2004). Finite-sample properties of propensity-score matching and weighting estimators. *Review of Economics and Statistics, 86,* 77-90.

Glynn, A.N., & Quinn, K.M. (2009). Estimation of causal effects with generalized additive models. Vienna, Austria: R Foundation for Statistical Computing.

Glynn, A. N., & Quinn, K. M. (2010). An introduction to the augmented inverse propensity weighted estimator. *Political Analysis, 18,* 36-56.

Goetgeluk, S., Vansteelandt, S., & Goetghebeur, E. (2008). Estimation of controlled direct effects. *Journal of the Royal Statistical Society: Series B (Statistical Methodology), 70,* 1049-1066.

Hirano, K., & Imbens, G. (2001). Estimation of causal effects using propensity score weighting: An application to data on right heart catheterization. *Health Services and Outcomes Research Methodology, 2,* 259–278.

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

Holland, P. W. (1988). Causal inference, path analysis, and recursive structural equation models. *Sociological Methodology, 18,* 449–484.

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

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

Hyman, H. H. (1955). *Survey design and analysis: Principles, cases, and procedures*. Glencoe, IL: Free Press.

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

Imai, K., Tingley, D., & Yamamoto, T. (2013). Experimental designs for identifying causal mechanisms. *Journal of the Royal Statistical Society, Series A: Statistics in Society*, 5-51.

Imbens, G. W. (2004). Nonparametric estimation of average treatment effects under exogeneity: A review. *The Review of Economics and Statistics, 86*, 4–29.

Joffe, M.M., & Greene, T. (2009). Related causal frameworks for surrogate outcomes. *Biometrics, 65*, 530-538.

Judd, C. M., & Kenny, D. A. (1981). Process analysis: Estimating mediation in treatment evaluations. *Evaluation Review, 5*, 602–619.

Kang, J. D., & Schafer, J. L. (2007). Demystifying double robustness: A comparison of alternative strategies for estimating a population mean from incomplete data. *Statistical science*, *22*, 523-539.

King, G., & Zeng, L. (2006). When can history be our guide? The pitfalls of counterfactual inference. *International Studies Quarterly, 51*, 183–210.

Lee, B. K., Lessler, J., & Stuart, E. A. (2011). Weight trimming and propensity score weighting. *PloS one*, *6*, e18174.

Lepage, B., Dedieu, D., Savy, N., & Lang, T. (2012). Estimating controlled direct effects in the presence of intermediate confounding of the mediator–outcome relationship: Comparison of five different methods. *Statistical methods in medical research*.

Loeys, T., Moerkerke, B., Raes, A., Rosseel, Y., & Vansteelandt, S. (2013). Estimation of controlled direct effects in the presence of exposure-induced confounding and latent variables. *Structural Equation Modeling*.

Lunceford, J. K., & Davidian, M. (2004). Stratification and weighting via the propensity score in estimation of causal treatment effects: a comparative study. *Statistics in Medicine, 23*, 2937-2960.

MacKinnon, D. P. (2008). *Introduction to statistical mediation analysis.* Mahwah, NJ: Erlbaum.

MacKinnon, D. P., Fritz, M. S., Williams, J., & Lockwood, C. M. (2007). Distribution of the product confidence limits for the indirect effect: Program PRODCLIN. *Behavior Research Methods, 39,* 384-389.

MacKinnon, D.P., Kisbu-Sakarya, Y., & Gottschall, A. (2013). Developments in mediation analysis. In T.D. Little (Ed.), *Oxford Handbook of Quantitative Methods*. New York: Oxford University Press.

MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods, 7,* 83.

MacKinnon, D. P., Lockwood C. M., & Williams, J. (2004). Confidence limits for the indirect effect: Distribution of the product and resampling methods. *Multivariate Behavioral Research, 39,* 99-128.

MacKinnon, D. P., & Pirlott, A. (2013). *The Unbearable lightness of b: Design approaches to causal interpretation of the M to Y relation*. Manuscript under review.

Mauro, R. (1990). Understanding L.O.V.E. (left out variables error): A method for estimating the effects of omitted variables. *Psychological Bulletin, 108*, 314-329.

Meeker, W. Q., Jr., Cornwell, L. W., & Aroian, L. A. (1981). The product of two normally distributed random variables. In W. J. Kennedy & R. E. Odeh (Eds.), *Selected Tables in Mathematical Statistics, Volume VII* (pp. 1-256*)*. Providence, RI: American Mathematical Society.

Morgan, S. L., & Winship, C. (2007). *Counterfactuals and causal inference.* New York: Cambridge University Press.

Pearl, J. (2001). Direct and indirect effects. *Proceedings of the 17th conference on uncertainty in artificial intelligence (pp 411–420). San Francisco, CA: Morgan Kaufmann.*

Pearl, J. (2009). *Causality: Models, reasoning, and inference.* Cambridge, MA: Cambridge University Press.

Pearl, J. (2012). The causal mediation formula – a guide to the assessment of pathways and mechanisms. *Prevention Science*, 1-11.

Pirlott, A. G., Kisbu-Sakarya, Y., DeFrancesco, C. A., Elliot, D. L., & MacKinnon, D.P. (2012). Mechanisms of motivational interviewing in health promotion: A Bayesian mediation analysis. *International Journal of Behavioral Nutrition and Physical Activity, 9(69),* 1-11.

Potter, F.J. (1993). The effect of weight trimming on nonlinear survey estimates. American Statistical Association: San Francisco, CA.

Ranby, K. W., MacKinnon, D. P., Fairchild, A. J., Elliot, D. L., Kuehl, K. S., & Goldberg, L. (2011). The PHLAME (Promoting Healthy Lifestyles: Alternative Models' Effects) firefighter study: Testing mediating mechanisms. *Journal of Occupational Health Psychology*, *16*, 501.

Robins, J.M. (1986). A new approach to causal inference in mortality studies with a sustained exposure period – applications to control of the healthy workers survivor effect. *Mathematical Modeling, 7,* 1393-1512.

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

Robins, J. M., Hernán, M., & Brumback, B. (2000). Marginal structural models and causal inference in epidemiology. *Epidemiology, 11,* 550– 560.

Robins, J. M., Rotnitzky, A., & Zhao, L. P. (1994). Estimation of regression coefficients when some regressors are not always observed. *Journal of the American Statistical Association, 89,* 846–866.

Robins, J. M., Rotnitzky, A., & Zhao, L. P. (1995). Analysis of semiparametric regression models for repeated outcomes in the presence of missing data. *Journal of the American Statistical Association, 90,* 106–121.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika, 70,* 41–55.

Rosenbaum, P. R., & Rubin, D. B. (1984). Reducing bias in observational studies using sub-classification on the propensity score. *Journal of the American Statistical Association, 79,* 516–524.

Rosenbaum, P. R., & Rubin, D. B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *American Statistician, 39,* 33–38.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology, 66*, 688–701.

Rubin, D. B. (1997). Estimating causal effects from large data sets using propensity scores. *Annals of Internal Medicine, 127,* 757–763.

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

Rubin, D. B. (2005). Causal inference using potential outcomes: Design, modeling, decisions. *Journal of the American Statistical Association, 100,* 322–331.

Rubin, D. B., & Thomas, N. (1992). Characterizing the effect of matching using linear propensity score methods with normal distributions. *Biometrika, 79*, 797–809.

Rubin, D. B., & Thomas, N. (1996). Matching using estimated propensity scores: Relating theory to practice. *Biometrics, 52*, 249–264.

Schafer, J. L., & Kang, J. (2008). Average causal effects from nonrandomized studies: a practical guide and simulated example. *Psychological methods*, *13*, 279.

Scharfstein, D. O., Rotnitzky, A., & Robins, J. M. (1999). Adjusting for nonignorable dropout using semiparametric nonresponse models. *Journal of the American Statistical Association, 94,*1096–1120 (with Rejoinder, 1135–1146).

Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. *Sociological Methodology*, *13*, 290-312.

Spencer, S. J., Zanna, M. P., & Fong, G. T. (2005). Establishing a causal chain: Why experiments are often more effective in examining psychological process than mediational analyses. *Journal of Personality and Social Psychology, 89,* 845–851.

Vansteelandt, S. (2009). Estimating Direct Effects in Cohort and Case–Control Studies. *Epidemiology, 20*, 851-860.

Vansteelandt, S., & Keiding, N. (2011). Invited commentary: G-computation–lost in translation? *American journal of epidemiology*, *173*, 739-742.

VanderWeele, T.J. (2010a). Bias formulas for sensitivity analysis for direct and indirect effects. *Epidemiology*, *21*, 540-551.

VanderWeele, T.J. (2010b). Direct and indirect effects for neighborhood-based clustered and longitudinal data. *Sociological Methods & Research, 38,* 515-544.

VanderWeele, T.J. (2011). Controlled direct and mediated effects: definition, identification and bounds. *Scandinavian Journal of Statistics*, *38*, 551-563.

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

Wang, H., Law, K. S., Hackett, R. D.,Wang, D., & Chen, Z. X. (2005). Leader-member exchange as a mediator of the relationship between transformational leadership and followers' performance and organizational citizenship behavior. *Academy of Management Journal*, *48*, 420-432.

Yuan, Y., & MacKinnon, D. P. (2009). Bayesian mediation analysis. *Psychological methods*, *14*, 301-322.

APPENDIX A

CORRELATIONS BETWEEN STUDY VARIABLES

Table below presents the correlations between study variables for the simulation condition with the following parameter values: *a*=.14, *b*=.59, *c'*=.59, *d*=.14, *f*=.14, *g*=.14, *h*=.14, *k*=1, *n*=1, *t*=1.

	X	M	Y	C1	C2	U
X	1	.17	.39	.10	.10	0
M		1	.23	.20	.20	.26
Y			1	.50	.50	.65
C1				1	.50	.70
C2					1	.70
U						1

Table below presents the correlations between study variables for the simulation condition with the following parameter values: *a*=.14, *b*=.59, *c'*=.59, *d*=.39, *f*=.39, *g*=.14, *h*=.14, *k*=1, *n*=1, *t*=1.

	X	M	Y	C1	C2	U
X	1	.18	.36	.10	.10	0
M		1	.46	.41	.41	.56
Y			1	.54	.54	.72
C1				1	.50	.70
C2					1	.70
U						1