Routledge
Taylor & Francis Group

The Journal of Experimental Education

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/vjxe20

Design and Analysis of Clustered Regression Discontinuity Designs for Probing Mediation Effects

Fangxing Bai, Ben Kelcey, Yanli Xie & Kyle Cox

To cite this article: Fangxing Bai, Ben Kelcey, Yanli Xie & Kyle Cox (29 Nov 2023): Design and Analysis of Clustered Regression Discontinuity Designs for Probing Mediation Effects, The Journal of Experimental Education, DOI: 10.1080/00220973.2023.2287445

To link to this article: https://doi.org/10.1080/00220973.2023.2287445

Published online: 29 Nov 2023.

THE JOURNAL OF EXPERIMENTAL EDUCATION
https://doi.org/10.1080/00220973.2023.2287445

Routledge
Taylor & Francis Group

Check for updates

Design and Analysis of Clustered Regression Discontinuity Designs for Probing Mediation Effects

Fangxing Bai[a], Ben Kelcey[a], Yanli Xie[b], and Kyle Cox[c]

[a]University of Cincinnati, Cincinnati, OH; [b]FL State University, Tallahassee, FL; [c]University of NC at Charlotte, Charlotte, NC

ABSTRACT
Prior research has suggested that clustered regression discontinuity designs are a formidable alternative to cluster randomized designs because they provide targeted treatment assignment while maintaining a high-quality basis for inferences on local treatment effects. However, methods for the design and analysis of clustered regression discontinuity designs have not been fully developed to address the array of core effects (e.g., main, moderation and mediation) typically examined in education studies. In this study, we complement prior design literature by developing principles of estimation, sampling variability, and closed-form expressions to predict the statistical power to detect mediation effects in clustered regression discontinuity designs. The results suggest that sample sizes typically seen in educational intervention studies (e.g., about 50 schools) can be sufficient to detect a mediation effect under some conditions when studies are carefully designed. We implement the results in software and a Shiny App (BLINDED FOR REVIEW).

KEYWORDS
Mediation; regression discontinuity; power; multilevel models; sample size determination; indirect effects

Introduction

Prior research has suggested that regression discontinuity designs are a formidable alternative to cluster randomized designs because they provide targeted treatment assignment while maintaining a high-quality basis for local inferences (Cook, 2008). In education, regression discontinuity designs often leverage cluster-level running variables to assign treatment conditions at the school-level to accommodate policy initiatives and/or the school-wide scope and implementation of many interventions. When correctly implemented and specified, these types of clustered regression discontinuity designs facilitate unbiased inferences concerning the local area effects of a treatment (Cook, 2008).

Despite the widespread use of clustered regression discontinuity designs in education, prior literature has not fully developed methods to address the more comprehensive sets of effects that are typically used to supplement evidence of whether an intervention works on average (i.e., main effect). For example, contemporary research routinely supplements evidence on main effects by further examining evidence for the underlying theory of action through mediation analyses. Such mediational analyses provide complementary evidence by probing the mechanisms through which the intervention operates on the outcome. More generally, such investigations of mediation effects play an essential role in testing and refining teaching and learning theories.

In this study, we advance regression discontinuity designs by developing principles of estimation, sampling variability, and closed-form expressions to predict the statistical power with which

CONTACT Fangxing Bai baifg@mail.uc.edu School of Education, University of Cincinnati, 638 Teachers-Dyer Complex, Cincinnati, OH 45219, USA.

we can detect mediation effects in clustered regression discontinuity designs. The results provide tools intended to inform and guide researchers in planning clustered regression discontinuity designs with cluster- and individual-level mediators. Below, we first detail the methods for cluster- and individual-level mediators using a working example. We then follow with a simulation assessing our results and finish with an illustration of the methods.

Methods

To explicate the models and methods, let us consider a working example from the literature (e.g., Bonell et al., 2018): The Learning Together intervention. This intervention aims to improve school environments in ways that reduce incidents of bullying and aggression and increase student health and wellbeing (Bonell et al., 2018). To achieve these goals, the Learning Together intervention focuses on modifying school policies and systems, increasing restorative practice, and processing social and emotional education (Bonell et al., 2018). The underlying theory of the intervention suggests that improvement of the school environment (i.e., a school-level mediator) and/or improvement of student opinion (i.e., a student-level mediator) about learning and the school community are key mechanisms through which we can improve student mental health (i.e., an outcome) (Bonell et al., 2019).

Within this context, consider a study that draws on a clustered regression discontinuity design such that schools are assigned to participate in the Learning Together intervention or business as usual based on a continuous school-level variable such as the number of mental health referrals at each school during the prior year. For example, the discontinuity assignment may assign schools above the 50th percentile (i.e., those schools with a high number of mental health referrals) to the Learning Together intervention and those at or below the 50th percentile to continue without any changes. Further assume that we are interested in examining the extent to which the Learning Together intervention improves student mental health (outcome) by operating through changes in (a) the school environment (school-level mediator) or (b) student opinion on learning and school community (student-level mediator).

Cluster-level mediator

Let us first consider analysis of a cluster-level mediator using a 2-2-1 mediation framework where the 2-2-1 numeric acronym represents the respective levels of the intervention (school-level, 2), mediator (school-level, 2) and outcome (student-level, 1). We develop our regression discontinuity design models within the context of flexible linear, quadratic and cubic functional forms and assume that the conditional density functions of both the mediator and the outcome exhibit continuity at the designated cutoff point. More specifically, prior literature has drawn on this approach to ensure the smoothness and uninterrupted flow in the distributions of both the mediator and the outcome variables at the threshold (e.g., Imbens & Lemieux, 2008; McCrary, 2008). Let us further assume that the running variable (school mental health) follows an approximate normal distribution

$$S_j = \varepsilon_j^S \qquad \varepsilon_j^S \sim N(0, \sigma_S^2) \tag{1}$$

Where S_j is the value of the running variable for school j with mean zero and variance σ_S^2.

To obtain inferences regarding mediation, we draw on the potential outcomes framework and take up the following additional requisite assumptions (e.g., VanderWeele, 2010): (1) stable unit treatment value assumption (SUTVA), (2) sequential ignorability, (3) consistency, (4) no downstream confounders, and (5) no treatment-by-mediator interaction (Kelcey et al., 2017). Although we draw on these assumptions to develop expressions to guide design and analysis, we note that their validity in practice should be thoroughly considered and evaluated (see, for example, Imbens & Lemieux, 2008). For the (cluster-level only) mediator model, we adopt the following model

$$M_j = \pi_0 + aT_j + f(S_j) + \pi_1 W_j + \pi_2 \overline{X}_j + \varepsilon_j^M \qquad \varepsilon_j^M \sim N(0, \sigma_{M|}^2) \tag{2}$$

Where M_j is the mediator value for school j, π_0 is the intercept, T_j is the treatment condition (i.e., 1 as treatment and 0 as control) with a as the path coefficient of the treatment-mediator effect, $f(S_j)$ is a (nonlinear) function of the running variable S_j, W_j is a school-level covariate with coefficient π_1, $\overline{X}_j$ is a school-level aggregate of a student-level covariate (X_{ij}) with coefficient π_2, and ε_j^M is the error term.

For the (multilevel) outcome model, we have

$$\text{Student−level:} \quad Y_{ij} = \beta_{0j} + \beta_1 \left(X_{ij} - \overline{X}_j\right) + \pi_2 V_{ij} + \varepsilon_{ij}^Y \qquad \varepsilon_{ij}^Y \sim N(0, \sigma_{Y|}^2)$$

$$\text{School−level:} \quad \beta_{0j} = \gamma_{00} + bM_j + F(S_j) + c'T_j + \gamma_{01} W_j + \gamma_{02} \overline{X}_j + u_{0j}^Y \qquad u_{0j}^Y \sim N(0, \tau_|^2) \tag{3}$$

Here, Y_{ij} represents the outcome for student i in school j, $F(S_j)$ is a (non)linear function of the running variable S_j (similar but different to $f(S_j)$), β_{0j} is the school-specific intercept, X_{ij} is a student-level covariate with school-level average $\overline{X}_j$ and coefficient β_1, V_{ij} is a student-level covariate that varies only across individuals (no variation among clusters) with the coefficient π_2, ε_{ij}^Y is the individual level error term, γ_{00} is the overall intercept of the model, b is the path coefficient of the mediator, c' is the path coefficient of the treatment, γ_{01} and γ_{02} are the coefficients of the covariates, and u_{0j}^Y is the cluster level random effect.

Under this model specification and aforementioned assumptions, the mediation effect (ME) can be described as

$$ME = ab \tag{4}$$

Where a and b are the path coefficients obtained from expressions (2) and (3).

Error variance

Under the aforementioned specification, the error variance of the mediation effect in a clustered regression discontinuity design is (Bollen, 1987; Kelcey et al., 2017; MacKinnon, 2012; Mackinnon et al., 2007)

$$\sigma_{ab}^2 = b^2\sigma_a^2 + a^2\sigma_b^2 + \sigma_a^2\sigma_b^2 + 2ab\sigma_{ab} + \sigma_{ab}^2 \approx b^2\sigma_a^2 + a^2\sigma_b^2 \tag{5}$$

The full expression capturing the variance can be simplified given the independence of the a and b paths (i.e., $\sigma_{ab} = 0$; Kelcey et al., 2017). Likewise, prior research has consistently found that the product of the error variances ($\sigma_a^2\sigma_b^2$) is approximately zero such that we can safely assume $\sigma_a^2\sigma_b^2 \approx 0$.

Under the maximum likelihood, the respective error variances can be obtained as a function of common summary statistics or design parameters

$$\sigma_a^2 = \frac{\sigma_{M|}^2}{n_2\sigma_T^2} = \frac{\sigma_M^2\left(1 - R_M^2\right)}{n_2 P(1-P)\left(1 - \rho_{f(S),T}^2\right)}$$

$$\sigma_b^2 = \frac{\tau_|^2 + \sigma_{Y|}^2/n_1}{n_2\sigma_M^2} = \frac{\tau^2\left(1 - R_{YL2}^2\right) + \sigma_Y^2\left(1 - R_{YL1}^2\right)/n_1}{n_2\sigma_M^2\left(1 - R_M^2\right)} \tag{6}$$

Here, σ_M^2 is the unconditional variance of the mediator, σ_T^2 is the unconditional variance of the treatment (i.e., P(1-P) where P is the proportion of clusters assigned to the treatment condition), $\rho_{f(S),T}$ is the correlation between the treatment assignment and the function of the running variable $f(S)T$ (Schochet, 2009), τ^2 is the unconditional school-level outcome variance, σ_Y^2 is the unconditional student-level outcome variance, R^2_{YL2} and R^2_{YL1} are the outcome variance explained by the school- and student-level covariates, and R^2_M is the mediator variance explained by the covariates in the models, and n_1 and n_2 represent the number of students per school and the number of school. Further, we can unpack the correlation between the function of the running variable ($f(S)$) and the treatment condition $\rho_{f(S),T}$ as

$$\rho_{f(S),T} = \frac{\sigma_{f(S),T}}{\sqrt{P(1-P)}\sigma_{f(S)}} = \frac{E\left(T_j^{RD} f(S_j)\right) - P\mu_{f(S)}}{\sqrt{P(1-P)}\sigma_{f(S)}} = \frac{P\left[E(f(S_j)|S_j \leq K) - \mu_{f(S)}\right]}{\sqrt{P(1-P)}\sigma_{f(S)}} \tag{7}$$

Where $\sigma_{f(S),T}$ is the covariance between the function of the running variable ($f(S)$) and the treatment condition (T), $\mu_{f(S)}$ is the mean of $f(S)$, K is the cutoff value on the running variable S such that $K = \frac{\Phi^{-1}(P,0,\sigma_S^2)}{\sigma_S^2}$ with Φ as the cumulative normal density function, and σ_S is the standard deviation of the running variable S, and $\sigma_{f(S)}$ is the standard deviation of $f(S)$.

Path formulation

When expressions tracking the sampling variability of an effect are intended for design decisions, studies have often reparametrized them in terms of parameters for which there is readily accessible empirical information from in the literature (e.g., Kelcey and Shen, 2017). In this way, researchers can draw on prior empirical values for key parameters that govern power when planning studies. For this reason, we decomposed the variance explained parameters (R-squared) into components attributable to the primary path coefficients (i.e., a, b, c') of the variables that are specific to this intervention and study (i.e., T, M) and components attributable to control covariates (e.g., X, W, $\overline{X}$) that tend to be more commonly reported in the literature. For the total variance in the outcome explained at the school level, the expected value of R^2_{YL2} can be expressed as

$$R^2_{YL2} = R^2_{YL2|\vec{Z}} + \frac{b^2\sigma^2_{f(S)} + \sigma^2_{F(S)} + 2b\sigma_{f(S),F(s)}}{\tau^2} + \frac{(ab+c')^2 P(1-P)}{\tau^2} + \frac{2\left(b\sigma_{f(S),T} + \sigma_{F(S),T}\right)(ab+c')}{\tau^2} + \frac{b^2\sigma_M^2\left(1-R^2_{ML2}\right)}{\tau^2} \tag{8}$$

Where we use the covariances between the treatment and the function of the running variable in the mediator model, $\sigma_{f(S),T}$, and outcome model, $\sigma_{F(S),T}$, and the covariance between $f(S)$ and $F(S)$, $\sigma_{f(S),F(s)}$, in this equation. This expression contains the variance explained by the covariates that can be obtained from prior studies ($R^2_{YL2|\vec{Z}}$) and the variance explained by the key focal variables in this study (i.e., T, M).

A similar decomposition of the variance explained in the mediator in the mediation model results in

$$R^2_M = R^2_{M|\vec{Z}} + \frac{\sigma^2_{f(S)}}{\tau^2_M} + \frac{a^2P(1-P)}{\tau^2_M} + \frac{2a\sigma_{f(S),T}}{\tau^2_M}. \tag{9}$$

Where where $R^2_{M|\vec{Z}}$ represents the variance in the mediator explained by (control) covariates that can be obtained from prior literature.

Student-level mediators

We next examined clustered regression discontinuity designs that probe a student-level mediator (e.g., student opinion). We again approximate the distribution of the running variable S_j using expression (1). In turn, we now draw on a multilevel structure to decompose variation among students and among schools in the mediator

$$\text{Student-level:}\quad M_{ij} = \pi_{0j} + \pi_1\left(X_{ij} - \overline{X}_j\right) + \pi_2 V_{ij} + \varepsilon_{ij}^M \qquad \varepsilon_{ij}^M \sim N(0, \sigma_{M|}^2)$$

$$\text{School-level:}\ \pi_{0j} = \zeta_{00} + aT_j + f(S_j) + \zeta_{01}W_j + \zeta_{02}\overline{X}_j + u_{0j}^M \qquad u_{0j}^M \sim N\left(0, \tau_{M|}^2\right) \tag{11}$$

Here, M_{ij} is the mediator value of student i in school j, π_{0j} is the school-specific intercept for school j, X_{ij} is a student-level covariate with the coefficient π_1, and $\overline{X}_j$ is its school-level mean, V_{ij} is a student-level covariate that varies only across students (no cluster variation) with the coefficient π_2, ε_{ij}^M is the individual level error term, ζ_{00} is the overall intercept, T_j is the treatment assignment coded as 0/1 with associated path coefficient a, $f(S_j)$ is the (non)linear function of S_j in the mediator model, ζ_{01} and ζ_{02} are the coefficients of the covariates, and u_{0j}^M is the cluster level random effect.

For the outcome model, we have

$$\text{Individual-level:}\quad Y_{ij} = \beta_{0j} + b_1\left(M_{ij} - \overline{M}_j\right) + \beta_1\left(X_{ij} - \overline{X}_j\right) + \beta_2 V_{ij} + \varepsilon_{ij}^Y \qquad \varepsilon_{ij}^Y \sim N(0, \sigma_{Y|}^2)$$

$$\text{Cluster-level:}\quad \beta_{0j} = \gamma_{00} + B\overline{M}_j + F(S_j) + c'T_j + \gamma_{01}W_j + \gamma_{02}\overline{X}_j + u_{0j}^Y \qquad u_{0j}^Y \sim N(0, \tau_{Y|}^2) \tag{12}$$

Where Y_{ij} is the outcome of student i's in school j, β_{0j} is the school-specific intercept, M_{ij} is the student-level mediator with the path coefficient b_1, $\overline{M}_j$ is the school-level aggregate mean of that mediator with B as its coefficient at the school-level, β_1 and β_2 are the coefficient of the student-level covariates, ε_{ij}^Y is the individual level error term, c' is the path coefficient of the treatment, $F(S_j)$ is the function of S_j in the outcome model, γ_{01} and γ_{02} are the coefficients of the school-level covariates, and u_{0j}^Y is the school-level random effect.

In this formulation the path coefficient B captures the total (conditional) association between the mediator and outcome (including the student- and school-level relationships; Kelcey et al. 2017). The path coefficient b_1 captures the student- level (conditional) association between the mediator and outcome (Kreft et al., 1995; Raudenbush & Bryk, 2002). As a result, the product of path coefficients a and B captures the total mediation effects including both the student-level mediator and the school-level averaged mediator values (Pituch & Stapleton, 2012; Zhang et al., 2009). Under the model specification and the aforementioned assumptions, the mediation effect is as follows

$$ME = aB \tag{13}$$

Error variance

The error variance of the mediation effect in clustered regression discontinuity design is

$$\sigma_{aB}^2 = B^2\sigma_a^2 + a^2\sigma_B^2 \tag{14}$$

Under the maximum likelihood, we can again express the error variances of path a and path b as follows as functions of summary statistics as

$$\sigma_a^2 = \frac{\tau_{M|}^2 + \sigma_{M|}^2/n_1}{n_2\sigma_T^2} = \frac{\tau_M^2\left(1 - R_{M^{L2}}^2\right) + \left(1 - R_{M^{L1}}^2\right)\sigma_M^2/n_1}{n_2P(1-P)\left(1 - \rho_{f(S)T}^2\right)}$$

$$\sigma_B^2 = \frac{\tau_{Y|}^2 + \sigma_{Y|}^2/n_1}{n_2\left(\tau_M^2 + \sigma_M^2/n_1\right)} = \frac{\tau_Y^2\left(1 - R_{YL2}^2\right) + \left(1 - R_{YL1}^2\right)\sigma_Y^2/n_1}{\mathrm{n}_2\left(\tau_M^2\left(1 - R_{ML2}^2\right) + \left(1 - R_{ML1}^2\right)\sigma_M^2/n_1\right)} \tag{15}$$

Where τ_M^2 is the unconditional variance of the cluster level aggregated mediation effect, σ_M^2 is the unconditional variance of the individual level mediator, σ_T^2 is the unconditional variance of the treatment, τ^2 is the unconditional variance of the cluster level aggregated outcome, σ_Y^2 is the unconditional variance of the individual outcome, and n_1 and n_2 represent the number of individuals per cluster and the number of clusters.

Because the treatment is assigned based on the cutoff value on the running variable, P is the proportion of clusters receiving treatment and T is following Bernoulli distribution, then we have $\sigma_T^2 = P(1-P)$. Based on Schochet (2009), the conditional $\sigma_{T|}^2 = P(1-P)(1-\rho_{f(S)T}^2)$, where $\rho_{f(S)T}$ is the correlation between the treatment assignment and the express of the running variable $f(S)$. R_{YL2}^2 and R_{YL1}^2 are the outcome variance explained by the school- and student-level covariates, and R_{ML2}^2 and R_{ML1}^2 are the mediator variance explained by the school- and student- level covariates in the models.

Path formulation

Similar to the approach applied for the 2-2-1 mediation, we decomposed R-squared parameters in 2-1-1 mediation into components specific to the intervention and study, and control covariates more commonly reported in the literature. The total variance of the second level explained in the outcome model is the assembling of all parts of the equations as follows.

$$R_{YL2}^2 = R_{YL2|\vec{Z}}^2 + \frac{B\sigma_{f(S)}^2 + \sigma_{F(S)}^2 + 2B\sigma_{f(S),F(S)}}{\tau^2} + \frac{(aB+c')^2P(1-P)}{\tau^2} + \frac{2\left(B\sigma_{f(S),T} + \sigma_{F(S),T}\right)(ab+c')}{\tau^2} + \frac{B^2\left(\tau_M^2\left(1-R_{ML2}^2\right) + \frac{\sigma_M^2\left(1-R_{ML1}^2\right)}{n_1}\right)}{\tau^2} \tag{16}$$

For the R squared of the individual level outcome, the expected expression is as follows

$$R_{YL1}^2 = R_{YL1|\vec{Z}}^2 + \frac{{b_1}^2\sigma_M^2\left(1-R_{ML1}^2\right)}{\sigma_Y^2} \tag{17}$$

With the same approach, we can construct the R squared for the mediation effect as

$$R_{ML2}^2 = R_{ML2|\vec{Z}}^2 + \frac{\sigma_{f(S)}^2}{\tau_M^2} + \frac{a^2P(1-P)}{\tau_M^2} + \frac{2a\sigma_{f(S),T}}{\tau_M^2} \tag{18}$$

Functions of the running variable

Our specification of the running variable S in shaping the mediator ($f(S)$) and outcome ($F(S)$) values allows for a broad set of flexible functions including linear, quadratic and cubic in the regressions. For example, consider a function $q(S)$ such that

$$q(S) = m(S) + \varepsilon^{q(S)} \quad \varepsilon^{q(S)} \sim N\left(0, \sigma_{q(S)|}^2\right) \tag{19}$$

Where $m(S)$ is the deterministic function of S and $\varepsilon^{q(S)}$ is the error term. The resulting covariance between the treatment and this function ($\sigma_{q(S),T}$) is

$$\sigma_{q(S),T} = \sigma_{m(S),T} \tag{20}$$

Where $\sigma_{m(S),T}$ represents the covariance between the treatment T and $m(S)$.

Returning to the construction of R squared of the cluster level outcome model, the covariance of $f(S)$ and $F(S)$ can be estimated as follows

$$\sigma_{f(S),F(s)} = \sigma_{m_f(S)}\sigma_{m_F(S)} \tag{21}$$

Where $\sigma_{m_f(S)}$ is the deterministic function of S in $f(S)$ and $\sigma_{m_F(S)}$ is the deterministic function of S in $F(S)$.

The correlation between the treatment T and $q(S)$ can then be obtained as follows

$$\rho_{q(S),T} = \frac{\sigma_{q(S),T}}{\sqrt{P(1-P)\sigma^2_{q(S)}}} = \frac{\sigma_{m(S),T}}{\sqrt{P(1-P)(\sigma^2_{q(S)|} + \sigma^2_{m(S)})}} \tag{22}$$

Where $\sigma^2_{m(S)}$ is the variance of $m(S)$.

For example, when considering a cubic function, the model is as follows

$$q(S) = m(S) + \varepsilon^{q(S)} = \lambda_1 S + \lambda_2 S^2 + \lambda_3 S^3 + \varepsilon^{q(S)} \;\; \varepsilon^{q(S)} \sim N\left(0, \sigma^2_{q(S)|}\right) \tag{23}$$

Where $m(S) = \lambda_1 S + \lambda_2 S^2 + \lambda_3 S^3$, and λ_1, λ_2, and λ_3 are the coefficients of the S, S^2, and S^3. We have $\sigma^2_{S^2} = 2\sigma^4_S$, $\sigma^2_{S^3} = 15\sigma^6_S$, and $\sigma_{S,S^3} = 3\sigma^4_S$, which yields

$$\sigma^2_{m(S)} = {\lambda_1}^2\sigma^2_S + {\lambda_2}^2 2\sigma^4_S + {\lambda_3}^2 15\sigma^6_S + 6\lambda_1\lambda_3\sigma^4_S \tag{24}$$

For the covariance between T and $m(S)$, we have

$$\sigma_{m(S),T} = \lambda_1\sigma_{S,T} + \lambda_2\sigma_{S^2,T} + \lambda_3\sigma_{S^3,T} \tag{25}$$

Where $\sigma_{S,T}$ is the covariance between T and S, $\sigma_{S^2,T}$ is the covariance between T and S^2, and σ_{S^3T} is the covariance between T and S^3. Using K ($K = \frac{\Phi^{-1}\left(P,0,\sigma^2_S\right)}{\sigma^2_S}$; Φ is the cumulative density function of normal distribution) as the cutoff value on the running variable S, the cluster with a value below K will be assigned to the treatment condition and the other clusters will be assigned to the control condition. The resulting covariance between S and T is

$$\sigma_{S,T} = -\phi(K,0,1)\sigma_S$$

$$\sigma_{S^2T} = 0.5\left(\sigma^2 + sgn(P-0.5)\left(\sigma^2\mathrm{erf}\left(\frac{\sqrt{y}}{\sqrt{2}\sigma}\right) - \frac{\sigma\sqrt{2y}e^{-y/(2\sigma^2)}}{\sqrt{\pi}}\right)\right) - \Phi\left(K,\;0,\sigma^2_s\right)\sigma^2_s$$

$$\sigma_{S^3T} = -\frac{\sigma_S\left(|K^3|^{\frac{2}{3}} + 2\sigma^2_S\right)e^{-\frac{|K^3|^{\frac{2}{3}}}{2\sigma^2_S}}}{\sqrt{2\pi}} \tag{26}$$

Assembling these results, the covariance between the function and the treatment can be expressed as

$$\begin{aligned}\sigma_{m(S),T} = {} & \lambda_1(-\phi(K,0,1)\sigma_S) \\ & + \lambda_2\left(0.5\left(\sigma^2 + sgn(P-0.5)\left(\sigma^2\mathrm{erf}\left(\frac{\sqrt{K^2}}{\sqrt{2}\sigma}\right)/2 - \frac{\sigma\sqrt{K^2}e^{-K^2/(2\sigma^2)}}{\sqrt{2\pi}}\right)\right) - \Phi\left(K,\;0,\sigma^2_s\right)\sigma^2_s\right) \\ & + \lambda_3\left(-\frac{\sigma_S\left(|K^3|^{\frac{2}{3}} + 2\sigma^2_S\right)e^{-\frac{|K^3|^{\frac{2}{3}}}{2\sigma^2_S}}}{\sqrt{2\pi}}\right)\end{aligned} \tag{27}$$

Similarly, the correlation between the treatment T and $q(S)$ can be obtained as follows

$$\rho_{q(S),T} = \frac{\sigma_{m(S)T}}{\sqrt{P(1-P)(\sigma^2_{q(S)|} + \lambda_1{}^2\sigma^2_S + \lambda_2{}^2 2\sigma^4_S + \lambda_3{}^2 15\sigma^6_S + 6\lambda_1\lambda_3\sigma^4_S)}} \tag{28}$$

Because we can obtain the R-squared of the $\sigma^2_{q(S)}$ from the prior studies, the formula of the correlation can be rewritten as follows

$$\rho_{f(S),T} = \frac{\sigma_{m(S),T}}{\sqrt{P(1-P)(\lambda_1{}^2\sigma^2_S + \lambda_2{}^2 2\sigma^4_S + \lambda_3{}^2 15\sigma^6_S + 6\lambda_1\lambda_3\sigma^4_S)/R^2_{q(S)}}} \tag{29}$$

Where $R^2_{q(S)}$ is the variance explained by $m(S)$ in $q(S)$.

For the expressions of S are quadratic regressions (i.e., $q(S) = \lambda_1 S + \lambda_2 S^2 + \varepsilon^{q(S)}$, $\varepsilon^{q(S)} \sim N\left(0, \sigma^2_{q(S)|}\right)$) or linear regressions (i.e., $q(S) = \lambda_1 S + \varepsilon^{q(S)}$, $\varepsilon^{q(S)} \sim N\left(0, \sigma^2_{q(S)|}\right)$), we can apply the formulas we developed for the cubic regression to the quadratic regressions with setting $\lambda_3 = 0$ or to the linear regressions with setting $\lambda_3 = \lambda_2 = 0$.

Statistical power

An important consideration in the design of clustered regression discontinuity studies is the statistical power with which we can detect the targeted mediation effect if it exists (Cohen, 1988). For this reason, we developed methods to predict power by extending several common mediation tests that are suitable for the design phase (i.e., Sobel test, joint test, and Monte Carlo interval test).

The Sobel test

One classic approach to testing mediation effects is the Sobel test based on the asymptotic normality of the sampling distribution of the mediation effect (Sobel, 1982). The Sobel test compares the ratio of the estimated indirect effect to its estimated standard error to a normal distribution (which tends to be a poor approximation in samples of less than say 100 or 200 clusters). Given our prior results, the forms of the Sobel test statistics (z^{Sobel}) for cluster-level mediators (2-2-1) and individual-level mediators (2-1-1) can be expressed as

$$z^{Sobel}_{RDD_{2-2-1}} = \frac{ab}{\sigma^2_{ab}} = \frac{ab}{a^2\frac{\tau^2\left(1-R^2_{YL2}\right)+\sigma^2_Y\left(1-R^2_{YL1}\right)/n_1}{n_2\sigma^2_M\left(1-R^2_M\right)} + b^2\frac{\sigma^2_M\left(1-R^2_M\right)}{n_2P(1-P)\left(1-\rho^2_{f(S)T}\right)}}$$

$$z^{Sobel}_{RDD_{2-1-1}} = \frac{ab}{\sqrt{a^2\frac{\tau^2_Y\left(1-R^2_{YL2}\right)+\left(1-R^2_{YL1}\right)\sigma^2_Y/n_1}{n_2\left(\tau^2_M\left(1-R^2_{ML2}\right)+\left(1-R^2_{ML1}\right)\sigma^2_M/n_1\right)} + B^2\frac{\tau^2_M\left(1-R^2_{ML2}\right)+\left(1-R^2_{ML1}\right)\sigma^2_M/n_1}{n_2P(1-P)\left(1-\rho^2_{f(S)T}\right)}}} \tag{30}$$

In turn, the power of the Sobel test can be estimated as:

$$P\left(|z^{Sobel}| > z_{critical}\right) = 1-\Phi\left(z_{critical} - z^{Sobel}\right) + \Phi\left(-z_{critical} - z^{Sobel}\right) \tag{31}$$

Where Φ represents the normal distribution with $z_{critical}$ as the chosen critical value (e.g., 1.96) corresponding to a nominal type I error rate.

The joint test

A common alternative test for mediation is the joint test. In the joint test, we can relax the normality assumption of the distribution of the mediation effect by examining the constituent paths independently (MacKinnon et al., 2002). The test statistics for the a paths with cluster- (2-2-1) and individual-level (2-1-1) mediators are

$$t_{a\_2-2-1} = a/\sigma_a^2 = a/\frac{\sigma_M^2\left(1-R_M^2\right)}{n_2P(1-P)\left(1-\rho_{TS}^2\right)}$$

$$t_{a\_2-1-1} = a/\sigma_a^2 = a/\frac{\tau_M^2\left(1-R_{M^{L2}}^2\right)+\left(1-R_{M^{L1}}^2\right)\sigma_M^2/n_1}{n_2P(1-P)\left(1-\rho_{TS}^2\right)} \tag{32}$$

Similarly, the test statistics for the path b under a cluster- and individual-level mediator are

$$t_{b\_2-2-1} = b/\sigma_b^2 = b/\frac{\tau^2\left(1-R_{Y^{L2}}^2\right)+\sigma_Y^2\left(1-R_{Y^{L1}}^2\right)/n_1}{n_2\sigma_M^2\left(1-R_M^2\right)}$$

$$t_{B\_2-2-1} = B/\sigma_B^2 = B/\frac{\tau^2\left(1-R_{Y^{L2}}^2\right)+\sigma_Y^2\left(1-R_{Y^{L1}}^2\right)/n_1}{n_2\left(\tau_M^2\left(1-R_{M^{L2}}^2\right)+\left(1-R_{M^{L1}}^2\right)\sigma_M^2/n_1\right)} \tag{33}$$

Using these tests, the statistical power of the joint test can be obtained as

$$\begin{aligned} &P(|t_a| > t_{critical}\ \&\ |t_b| > t_{critical}) \\ &\quad = (1-\Phi(t_{critical}-t_a)+\Phi(-t_{critical}-t_a))*(1-\Phi(t_{critical}-t_b)+\Phi(-t_{critical}-t_b)) \end{aligned} \tag{34}$$

Where t represents the t density function with corresponding degrees of freedom with $t_{critical}$ as the chosen critical value (e.g., 1.96 in large sample sizes) corresponding to a t distribution type one error rate.

The Monte Carlo test

A modern approach is the resampling-based Monte Carlo test (Preacher & Selig, 2012). In the Monte Carlo test, random samples are drawn from a joint distribution formed by the product of the with the a path and b (or B) path coefficients. The distribution of each path is approximated with a normal distribution centered at the maximum likelihood point estimate and variance set to the sampling variability of the coefficient.

In practice, draws are taken from

$$\begin{pmatrix} a^* \\ b^* \end{pmatrix} \sim MVN\left(\begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix}, \begin{pmatrix} \hat{\sigma}_{\hat{a}}^2 & \hat{\sigma}_{\hat{a},\ \hat{b}} \\ \hat{\sigma}_{\hat{a},\ \hat{b}} & \hat{\sigma}_{\hat{b}}^2 \end{pmatrix}\right) \tag{35}$$

In turn, we use the product of sampled path coefficients, a^*b^*, to approximate the sampling distribution of a mediation effect. Statistical power is the proportion of the asymmetric confidence intervals (e.g., 95%) that exclude zero.

Simulation

We conducted Monte Carlo simulations to assess the accuracy of our derivations regarding the error variance of the path coefficients, the R-squared path formulations and the statistical power to detect a mediation effect. The simulations followed the guidance and principles in prior literature in this area (Bloom, 2012; Calcagno & Long, 2008; Imbens & Lemieux, 2008; McCrary, 2008; Schochet et al., 2010). We drew on 5000 random samples and compared the empirical values with those predicted by our derived formulas. We simulated data based on Equations (1)–(3) for the 2-2-1 mediation and Equations (16)–(18) for the 2-1-1 mediation using cluster sample sizes ranging from 20 to 200.

Table 1. Average absolute bias in sampling variability for cluster-level mediator.

	Sample size		Parameters											Mediation error variance			
	n1	n2	a	b	c	λ_1	λ_2	λ_3	Λ_1	Λ_2	Λ_3	P	ICC	Empirical variance	Predicted variance	Absolute bias	Average absolute bias
Con 1	10	40	0.3	0.1	0.15	0.2	0.4	0.2	0.1	0.2	0.1	0.3	0.2	0.0031	0.0020	0.0011	0.0004
	10	60	0.3	0.1	0.15	0.2	0.4	0.2	0.1	0.2	0.1	0.3	0.2	0.0017	0.0014	0.0003	
	10	100	0.3	0.1	0.15	0.2	0.4	0.2	0.1	0.2	0.1	0.3	0.2	0.0009	0.0008	0.0001	
Con 2	10	40	0.5	0.3	0.25	0.1	0.2	0.1	0.4	0.1	0.2	0.5	0.05	0.0143	0.0108	0.0035	0.0013
	10	60	0.5	0.3	0.25	0.1	0.2	0.1	0.4	0.1	0.2	0.5	0.05	0.0087	0.0070	0.0017	
	10	100	0.5	0.3	0.25	0.1	0.2	0.1	0.4	0.1	0.2	0.5	0.05	0.0047	0.0044	0.0003	
Con 3	10	40	0.8	0.5	0.35	0.4	0.1	0.2	0.2	0.4	0.1	0.7	0.5	0.0652	0.0551	0.0101	0.0040
	10	60	0.8	0.5	0.35	0.4	0.1	0.2	0.2	0.4	0.1	0.7	0.5	0.0400	0.0368	0.0032	
	10	100	0.8	0.5	0.35	0.4	0.1	0.2	0.2	0.4	0.1	0.7	0.5	0.0225	0.0223	0.0002	
Con 4	20	40	0.3	0.1	0.15	0.4	0.2	0.1	0.2	0.1	0.2	0.3	0.05	0.0022	0.0016	0.0006	0.0003
	20	60	0.3	0.1	0.15	0.4	0.2	0.1	0.2	0.1	0.2	0.3	0.05	0.0014	0.0011	0.0003	
	20	100	0.3	0.1	0.15	0.4	0.2	0.1	0.2	0.1	0.2	0.3	0.05	0.0008	0.0006	0.0001	
Con 5	20	40	0.3	0.5	0.25	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0357	0.0288	0.0069	0.0071
	20	60	0.3	0.5	0.25	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0321	0.0245	0.0076	
	20	100	0.3	0.5	0.25	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0181	0.0154	0.0027	
Con 6	20	40	0.8	0.3	0.35	0.2	0.1	0.1	0.1	0.4	0.2	0.5	0.5	0.0357	0.0288	0.0069	0.0023
	20	60	0.8	0.3	0.35	0.2	0.1	0.1	0.1	0.4	0.2	0.5	0.5	0.0210	0.0191	0.0019	
	20	100	0.8	0.3	0.35	0.2	0.1	0.1	0.1	0.4	0.2	0.5	0.5	0.0118	0.0111	0.0007	

Note. Full version of Table 1 can be found in Appendix B.

In the first simulation we examined the error variance of the mediation effect by considering 114 different conditions for the cluster-level mediator (see Table 1) and another 114 conditions for the individual-level mediator (see Table 2). In each simulation, we varied the individual- and cluster-level sample sizes, each of the path coefficients, the intraclass correlation coefficient and the proportion of clusters receiving treatment. In addition, the analysis examined situations that involve highly nonlinear functions of the running variable ($f(S)$ and $F(S)$ were cubic) in terms of the relationships with the mediator and outcome. We estimated the empirical sampling variance of the mediation effect across the draws under each condition using

$$\sigma_{ME}^2 = \frac{1}{K}\sum_{k=1}^{K}(ME_k - \overline{ME})^2 \tag{36}$$

Similarly, we predicted the sampling variability of the mediation effect using the formulas derived above.

For each cluster sample size condition, we then compared the empirical sampling variance with that predicted by our formulas using absolute bias. We summarize the absolute bias of the predicted error variance across those draws and sample sizes using the average absolute bias

$$Average\ absolute\ bias(\hat{\sigma}^2) = \frac{1}{L}\sum_{l}^{L}|\hat{\sigma}_l^2 - \sigma_l^2| \tag{37}$$

Where l represents the cluster sample spanning from 20 to 200, σ_l^2 represents the empirical variance of the mediation effect within the sample size of 5000, and $\hat{\sigma}_l^2$ represents the predicted variance of the mediation effect across the 5000 draws using the Equation (8) for the 2-2-1 mediation and Equation (23) for the 2-1-1 mediation.

In the second simulation, we examined our formula-based predictions of statistical power while varying the functional form (i.e., linear, quadratic, cubic) of the running variable using 54 different conditions (18 conditions for linear function of running variables, 18 conditions for quadratic function of running variables, and 18 conditions for cubic function of running variables) for the cluster-level mediator setting (see Table 3) and 54 conditions for the individual-level

Table 2. Average absolute bias in sampling variability for individual-level mediator.

	Sample Size		Parameters												Mediation Error Variance			
	n1	n2	a	B	b1	c	λ_1	λ_2	λ_3	Λ_1	Λ_2	Λ_3	P	ICC	Empirical variance	Predicted variance	Absolute bias	Average absolute bias
Con 1	10	40	0.3	0.1	0.1	0.15	0.2	0.4	0.2	0.1	0.2	0.1	0.3	0.2	0.0048	0.0028	0.0020	0.0008
	10	60	0.3	0.1	0.1	0.15	0.2	0.4	0.2	0.1	0.2	0.1	0.3	0.2	0.0026	0.0019	0.0008	
	10	100	0.3	0.1	0.1	0.15	0.2	0.4	0.2	0.1	0.2	0.1	0.3	0.2	0.0014	0.0011	0.0003	
Con 2	10	40	0.5	0.3	0.05	0.25	0.1	0.2	0.1	0.4	0.1	0.2	0.5	0.05	0.0092	0.0068	0.0024	0.0012
	10	60	0.5	0.3	0.05	0.25	0.1	0.2	0.1	0.4	0.1	0.2	0.5	0.05	0.0058	0.0041	0.0016	
	10	100	0.5	0.3	0.05	0.25	0.1	0.2	0.1	0.4	0.1	0.2	0.5	0.05	0.0034	0.0026	0.0008	
Con 3	10	40	0.8	0.5	0.2	0.35	0.4	0.1	0.2	0.2	0.4	0.1	0.7	0.5	0.0775	0.0612	0.0162	0.0068
	10	60	0.8	0.5	0.2	0.35	0.4	0.1	0.2	0.2	0.4	0.1	0.7	0.5	0.0476	0.0398	0.0078	
	10	100	0.8	0.5	0.2	0.35	0.4	0.1	0.2	0.2	0.4	0.1	0.7	0.5	0.0273	0.0240	0.0032	
Con 4	20	40	0.3	0.1	0.05	0.15	0.4	0.2	0.1	0.2	0.1	0.2	0.3	0.05	0.0032	0.0021	0.0010	0.0005
	20	60	0.3	0.1	0.05	0.15	0.4	0.2	0.1	0.2	0.1	0.2	0.3	0.05	0.0019	0.0015	0.0004	
	20	100	0.3	0.1	0.05	0.15	0.4	0.2	0.1	0.2	0.1	0.2	0.3	0.05	0.0011	0.0007	0.0003	
Con 5	20	40	0.3	0.5	0.2	0.25	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0184	0.0136	0.0048	0.0025
	20	60	0.3	0.5	0.2	0.25	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0111	0.0088	0.0023	
	20	100	0.3	0.5	0.2	0.25	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0063	0.0053	0.0010	
Con 6	20	40	0.8	0.3	0.1	0.35	0.2	0.1	0.1	0.1	0.4	0.2	0.5	0.5	0.0384	0.0290	0.0094	0.0044
	20	60	0.8	0.3	0.1	0.35	0.2	0.1	0.1	0.1	0.4	0.2	0.5	0.5	0.0229	0.0185	0.0045	
	20	100	0.8	0.3	0.1	0.35	0.2	0.1	0.1	0.1	0.4	0.2	0.5	0.5	0.0127	0.0111	0.0016	

Note. Full version of Table 2 can be found in Appendix B.

Table 3. Empirical versus predicted statistical power for cluster-level mediators with linear function of the running variable.

Sample Size		Parameters													Power	
n1	n2	a	b	c	λ_1	λ_2	λ_3	Λ_1	Λ_2	Λ_3	P	ICC	Empirical R^2_{YL2}	Predicted R^2_{YL2}	Empirical rejection rate	Predicted power
10	30	0.1	0.1	0.25	0.1	0	0	0.4	0	0	0.7	0.2	0.567	0.560	0.008	0.003
10	40	0.3	0.3	0.35	0.4	0	0	0.2	0	0	0.3	0.15	0.698	0.681	0.120	0.082
10	50	0.3	0.1	0.15	0.2	0	0	0.1	0	0	0.3	0.2	0.364	0.363	0.019	0.017
10	80	0.5	0.5	0.15	02	0	0	0.1	0	0	0.5	0.05	0.927	0.911	0.403	0.370
10	100	0.5	0.3	0.25	0.1	0	0	0.4	0	0	0.5	0.05	0.900	0.882	0.469	0.447
10	300	0.8	0.5	0.35	0.4	0	0	0.2	0	0	0.7	0.5	0.566	0.565	0.998	0.998
20	30	0.8	0.3	0.25	0.2	0	0	0.1	0	0	0.5	0.15	0.661	0.642	0.351	0.322
20	40	0.8	0.5	0.35	0.1	0	0	0.2	0	0	0.3	0.3	0.682	0.666	0.478	0.478
20	50	0.3	0.5	0.25	0.1	0	0	0.4	0	0	0.7	0.2	0.793	0.777	0.154	0.105
20	80	0.1	0.1	0.15	0.4	0	0	0.4	0	0	0.7	0.05	0.865	0.856	0.048	0.035
20	100	0.8	0.3	0.35	0.2	0	0	0.1	0	0	0.5	0.5	0.428	0.428	0.644	0.661
20	300	0.3	0.1	0.15	0.4	0	0	0.2	0	0	0.3	0.05	0.706	0.703	0.445	0.442
50	30	0.5	0.5	0.35	0.4	0	0	0.2	0	0	0.3	0.2	0.772	0.748	0.237	0.160
50	40	0.3	0.3	0.25	0.1	0	0	0.1	0	0	0.5	0.15	0.643	0.630	0.141	0.105
50	50	0.5	0.3	0.35	0.4	0	0	0.2	0	0	0.3	0.2	0.636	0.627	0.277	0.241
50	80	0.8	0.1	0.15	0.2	0	0	0.4	0	0	0.7	0.2	0.594	0.590	0.236	0.235
50	100	0.5	0.5	0.25	0.1	0	0	0.1	0	0	0.5	0.05	0.927	0.914	0.458	0.434
50	300	0.3	0.1	0.15	0.2	0	0	0.4	0	0	0.7	0.5	0.400	0.400	0.141	0.140

mediator setting (see Table 4). Further, we probed the accuracy of our path formulations of the R-squared parameters by also comparing our formula-based predictions (expression 8 and 16 above) with the empirical counterparts from the simulation. In these analyses, we considered a broad range of parameter values and combinations (see Tables 3 and 4) including different non-linear forms of the running variable. We compared the empirical rejection rates using the Monte Carlo test across 5000 simulation draws and the predicted rejection rate using our formulas.

In the third simulation, we examined the accuracy of our power predictions when the sequential ignorability assumption is violated and when the no treatment by mediator interaction assumption is violated. For violations of the sequential ignorability assumption, we considered unobserved confounding such that there were (residual) conditional correlations between the mediator and the outcome, the running variable and the outcome, or the running variable and

Table 4. Empirical versus predicted statistical power for individual-level mediators with linear function of the running variable.

Sample size		Parameters														Power	
n1	n2	a	B	b1	c	λ_1	λ_2	λ_3	Λ_1	Λ_2	Λ_3	P	ICC	Empirical R^2_{YL2}	Predicted R^2_{YL2}	Empirical rejection rate	Predicted power
10	30	0.1	0.1	0.1	0.25	0.1	0	0	0.4	0	0	0.7	0.2	0.538	0.530	0.006	0.002
10	40	0.3	0.3	0.2	0.35	0.4	0	0	0.2	0	0	0.3	0.15	0.545	0.551	0.093	0.087
10	50	0.3	0.1	0.1	0.15	0.2	0	0	0.1	0	0	0.3	0.2	0.292	0.295	0.022	0.020
10	80	0.5	0.5	0.05	0.15	0.2	0	0	0.1	0	0	0.5	0.05	0.759	0.758	0.971	0.976
10	100	0.5	0.3	0.05	0.25	0.1	0	0	0.4	0	0	0.5	0.05	0.835	0.824	0.828	0.838
10	300	0.8	0.5	0.2	0.35	0.4	0	0	0.2	0	0	0.7	0.5	0.571	0.572	0.996	0.995
20	30	0.8	0.3	0.05	0.25	0.2	0	0	0.1	0	0	0.5	0.15	0.494	0.496	0.325	0.309
20	40	0.8	0.5	0.2	0.35	0.1	0	0	0.2	0	0	0.3	0.3	0.582	0.584	0.561	0.589
20	50	0.3	0.5	0.2	0.25	0.1	0	0	0.4	0	0	0.7	0.2	0.679	0.673	0.243	0.246
20	80	0.1	0.1	0.1	0.15	0.4	0	0	0.4	0	0	0.7	0.05	0.847	0.834	0.008	0.014
20	100	0.8	0.3	0.1	0.35	0.2	0	0	0.1	0	0	0.5	0.5	0.433	0.434	0.639	0.677
20	300	0.3	0.1	0.05	0.15	0.4	0	0	0.2	0	0	0.3	0.05	0.635	0.633	0.384	0.412
50	30	0.5	0.5	0.2	0.35	0.4	0	0	0.2	0	0	0.3	0.2	0.626	0.624	0.302	0.295
50	40	0.3	0.3	0.1	0.25	0.1	0	0	0.1	0	0	0.5	0.15	0.444	0.448	0.134	0.122
50	50	0.5	0.3	0.1	0.35	0.4	0	0	0.2	0	0	0.3	0.2	0.514	0.513	0.325	0.317
50	80	0.8	0.1	0.05	0.15	0.2	0	0)	0.4	0	0	0.7	0.2	0.574	0.572	0.147	0.141
50	100	0.5	0.5	0.05	0.25	0.1	0	0	0.1	0	0	0.5	0.05	0.701	0.702	0.999	0.999
50	300	0.3	0.1	0.2	0.15	0.2	0	0	0.4	0	0	0.7	0.5	0.398	0.398	0.135	0.141

the mediator. Similarly, although the expressions can be adapted to accommodate a treatment by mediator interaction (e.g., Kelcey et al., 2017), we examined sensitivity of our predictions to a treatment by mediator interaction. By systematically introducing these controlled violations of assumptions, we indexed the practical accuracy and sensitivity of our formula-based predictions.

Results

Tables 1 and 2 summarize the average absolute bias in the sampling variability of the mediation effect for each condition when considering cluster- and individual-level mediators (see Appendix B for full results). The results suggested that our formulas demonstrated good accuracy across conditions. For conditions with cluster sample size greater than about 50, the absolute bias of the predicted error variance was near zero. Below 50 clusters the accuracy was still quite good in most conditions but was slightly elevated when the running variable had a strong nonlinear relationship with the mediator or outcome. We illustrate these results for cluster- and individual-level mediators by plotting the six parameter combinations listed in Tables 1 and 2 for cluster samples ranging between 20 and 200 (Figures 1 and 2). The plotted comparisons demonstrate the small sample bias (underestimation) in the formula-based predictions of the sample variability of the mediation effect when cluster-level samples are less than about 50. That bias is most pronounced when the running variable is highly nonlinear in its relationship with the mediator or outcome (e.g., conditions 5 and 6) but that bias decreases quickly with cluster-level sample sizes greater that about 50.

Similarly, we report the empirical versus predicted statistical power and R-squared values for cluster- (Table 3) and individual-level mediators (Table 4) under linear functions of the running variable (see Appendix B for quadratic and cubic functions of the running variable). Our results demonstrated excellent and consistent accuracy for both quantities even with small sample sizes and highly nonlinear relationships between the running variable and the mediator/outcome. Figure 3 illustrates and expands on the tabled results by plotting the empirical and predicted power curves as a function of cluster-level sample size for the Monte Carlo interval, joint and Sobel tests under condition 9 with cubic function of the running variables ($n_1 = 20,\ ICC = 0.50,\ a = 0.80,\ b = 0.30,\ c' = 0.35,\ \lambda_1 = 0.20,\ \lambda_2 = 0.10,\ \lambda_3 = 0.10,\ \Lambda_1 = 0.10,\ \Lambda_2 = 0.40,\ \Lambda_3 = 0.20,\ R^2_{M|\vec{Z}} = R^2_{Y^{L1}|\vec{Z}} = R^2_{Y^{L2}|\vec{Z}} = 0.25,\ and\ R^2_{f(S)} = R^2_{F(S)} = 0.75$ for 2-2-1

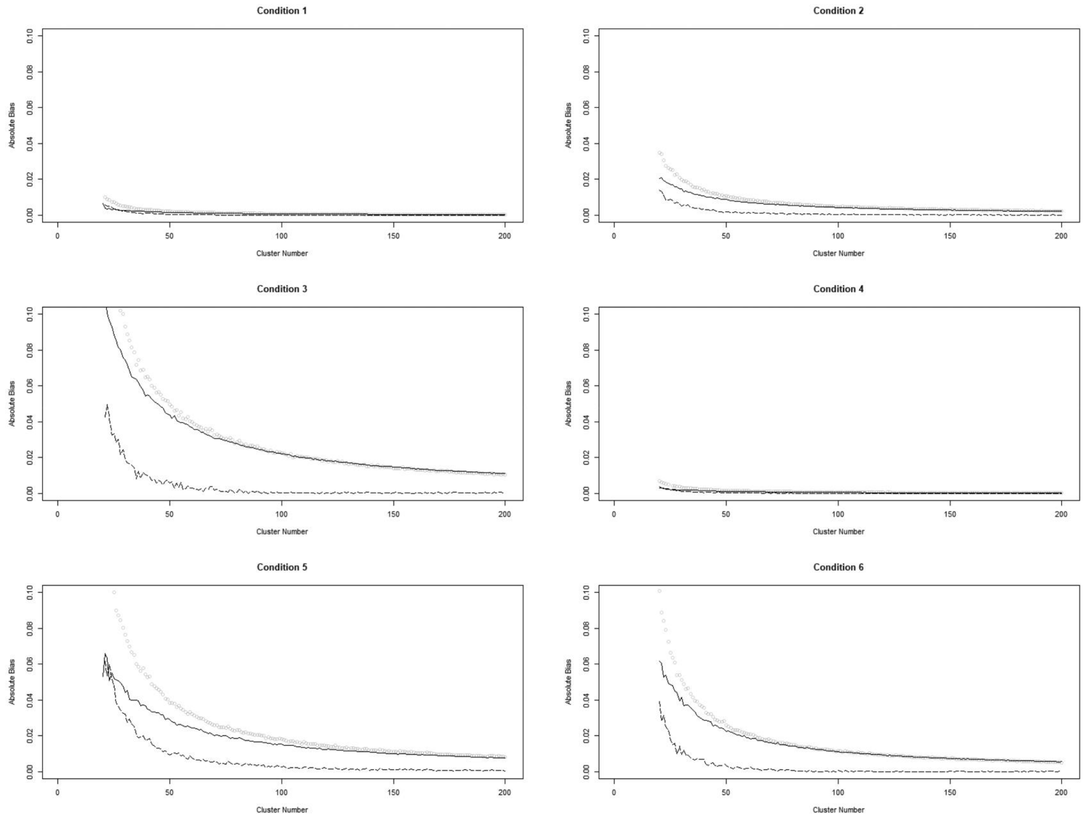


Figure 1. True empirical (gray) and predicted (black) error variance and absolute bias (dashed) as a function of group-level sample size for 2-2-1 mediation.

mediation; $n_1 = 20$, $ICC = 0.50$, $a = 0.80$, $B = 0.30$, $b_1 = 0.10$, $c' = 0.35$, $\lambda_1 = 0.20$, $\lambda_2 = 0.10$, $\lambda_3 = 0.10$, $\Lambda_1 = 0.10$, $\Lambda_2 = 0.40$, $\Lambda_3 = 0.20$, $R^2_{M^{L1}|\vec{Z}} = R^2_{M^{L2}|\vec{Z}} = R^2_{Y^{L1}|\vec{Z}} = R^2_{Y^{L2}|\vec{Z}} = 0.25$, *and* $R^2_{f(S)} = R^2_{F(S)} = 0.75$ for 2-1-1 mediation). Evident from these figures, the formula-based power predictions for the Monte Carlo and joint tests have excellent accuracy across all sample sizes. By contrast, the Sobel test incurs significant inaccuracies across sample sizes and will generally misestimate power and requisite sample sizes.

Tables 5a–c summarize the misprediction associated with different types and levels of violations of the sequential ignorability or the no treatment by mediator interaction assumptions. Collectively, the results demonstrated increased errors in power prediction when there is unaccounted for confounding among the treatment, mediator and outcome or when there is a treatment by mediator interaction. However, the results also demonstrated that under the circumstances considered, our formulas still provided a reasonable prediction of power for planning purposes.

Illustration

Returning to our working example, consider a theory that the Learning Together intervention is hypothesized impact student mental health through changes in the school environment (cluster-level mediator) or/and student opinion on learning and school community (student-level mediator) (MacKinnon, 2008). Assume researchers draw on a clustered regression discontinuity design such that schools with an average pretest score below the 50th percentile will be assigned to

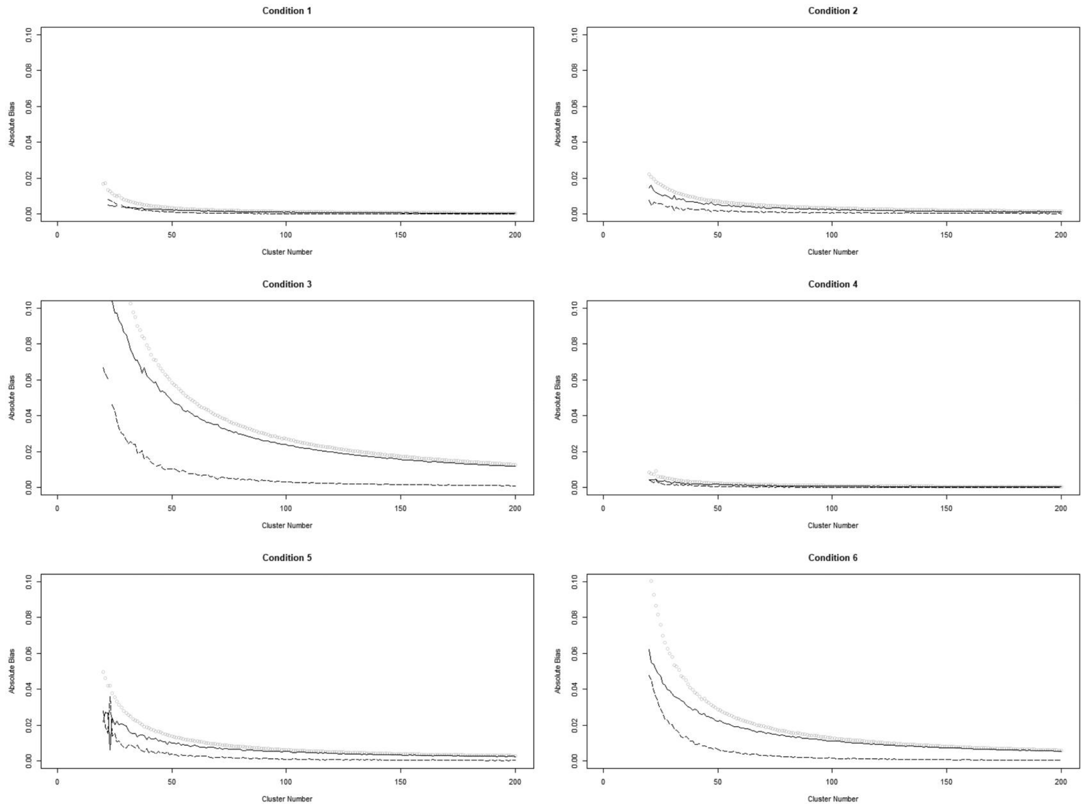


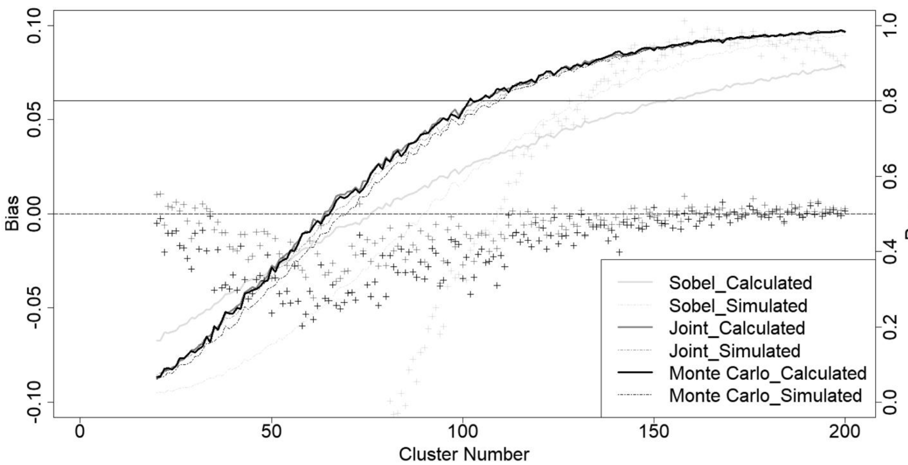
Figure 2. True empirical (gray) and predicted (black) error variance and absolute bias (dashed) as a function of group-level sample size for 2-1-1 mediation.

receive the Learning Together intervention (treatment), and the other schools will be assigned to receive no new intervention (control).

Let us assume that we intend to sample 50 students per school (n_1) and the intraclass correlation coefficient (ICC) for the outcome is 0.20. For the R squared, we assume the selected covariates can explain 25% of the variance in the mediator and the outcome on both the student- and school-level, and the deterministic functions can explain 75% of the variance of the cubic functions of the running variable. For the path coefficients, we assume the approximated treatment-mediator difference between the treatment and control condition (a) is 0.50 standard deviations, the approximated mediator-outcome conditional association (b) is 0.30 standard deviations, and the direct effect (c') is 0.35 standard deviations. For the running variable, we adopt a cubic relationships such that the coefficients predicting the mediator are linear (λ_1), quadratic (λ_2), and cubic (λ_3) term are all equal to 0.12; the coefficients predicting the outcome model are linear (Λ_1), quadratic (Λ_2), and cubic (Λ_3) are all equal to 0.20. In brief, we assume $n_1 = 50,\ ICC = 0.20,\ a = 0.50,\ b = 0.30,\ c' = 0.35,\ \lambda_1 = \lambda_2 = \lambda_3 = 0.12,\ \Lambda_1 = \Lambda_2 = \Lambda_3 = 0.02,\ R^2_{M|\vec{Z}} = R^2_{Y^{L1}|\vec{Z}} = R^2_{Y^{L2}|\vec{Z}} = 0.25,\ and\ R^2_{f(S)} = R^2_{F(S)} = 0.75.$

With the given parameter values, we can obtain an appropriate school-level sample size providing an 80% chance to detect the school-level mediation effect with the Type I error rate as 0.05. We implemented this in the accompanying Shiny App (BLINDED FOR REVIEW). As shown in Figure 4, we plotted the statistical power curves of Sobel (light grey dash curve), joint (middle grey dot curve), and Monte Carlo interval (black curve) tests calculated by our developed power formulas. Our results showed that under the Monte Carlo test a school sample size of 57

(a)



(b)

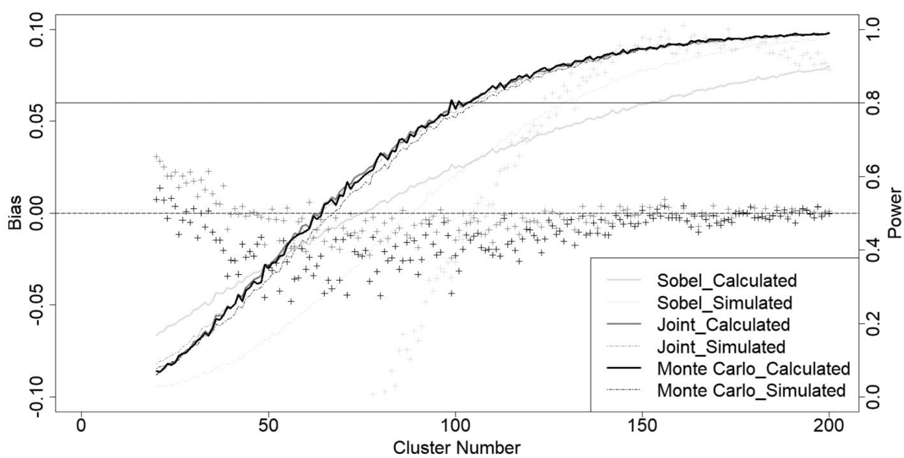


Figure 3. Bias (stars) and power (curves) as functions of group sample size (n2) when utilizing the parameters in condition 9 with cubic function of the running variables for (a) cluster- and (b) individual-level mediators.

with 50 students each can provide a 0.8 statistical power. Under the same scenario, if we are intended to test the main effect, the school sample size 34 with 50 students each would provide a 0.8 statistical power. While switching to the cluster randomized trial and using the running variable as a covariate uncorrelated to the treatment, 32 schools with 50 students each would provide 0.8 for detecting the school-level mediation effect.

Similarly, for a clustered regression discontinuity study planning to detect the effect of student opinion (student-level mediator; 2-1-1 mediation), assume we sample 20 students per school (n_1) with the ICCs for the mediation and outcome are both 0.20, the selected covariates can explain 25% of the variance in both the mediator and outcome on the student- and school-level, the deterministic functions can explain 75% of the variance of the cubic functions of the running variable, the approximated treatment-mediator difference between the treatment and control condition (a) is 0.50 standard deviations, the approximated total mediator-outcome effect (B) is 0.30 standard deviations, the student-level approximated mediator-outcome effect (b_1) is 0.20 standard deviations, and the direct effect (c') is 0.35 standard deviations. For the expression of running

Table 5. Empirical versus predicted statistical power for individual-level mediators with linear function of the running variable when the Assumption (2) sequential ignorability is violated (treatment-outcome and mediator-outcome relationship are both confounded), Assumption (2) sequential ignorability is violated (treatment-mediator and mediator-outcome relationship are both confounded) and Assumption (5) no treatment-by-mediator interaction is violated.

(a) Empirical versus predicted statistical power for individual-level mediators with linear function of the running variable when the Assumption (2) sequential ignorability is violated (treatment-outcome and mediator-outcome relationship are both confounded).

Sample size		Parameters											Power	
n1	n2	a	b	c	λ_1	λ_2	λ_3	Λ_1	Λ_2	Λ_3	P	ICC	Empirical rejection rate	Predicted power
10	30	0.1	0.1	0.25	0.1	0	0	0.4	0	0	0.7	0.2	0.072	0.032
10	40	0.3	0.3	0.35	0.4	0	0	0.2	0	0	0.3	0.15	0.152	0.103
10	50	0.3	0.1	0.15	0.2	0	0	0.1	0	0	0.3	0.2	0.144	0.111
10	80	0.5	0.5	0.15	02	0	0	0.1	0	0	0.5	0.05	0.406	0.378
10	100	0.5	0.3	0.25	0.1	0	0	0.4	0	0	0.5	0.05	0.463	0.438
10	300	0.8	0.5	0.35	0.4	0	0	0.2	0	0	0.7	0.5	0.998	0.998
20	30	0.8	0.3	0.25	0.2	0	0	0.1	0	0	0.5	0.15	0.456	0.385
20	40	0.8	0.5	0.35	0.1	0	0	0.2	0	0	0.3	0.3	0.502	0.448
20	50	0.3	0.5	0.25	0.1	0	0	0.4	0	0	0.7	0.2	0.158	0.118
20	80	0.1	0.1	0.15	0.4	0	0	0.4	0	0	0.7	0.05	0.068	0.052
20	100	0.8	0.3	0.35	0.2	0	0	0.1	0	0	0.5	0.5	0.833	0.828
20	300	0.3	0.1	0.15	0.4	0	0	0.2	0	0	0.3	0.05	0.443	0.442
50	30	0.5	0.5	0.35	0.4	0	0	0.2	0	0	0.3	0.2	0.243	0.159
50	40	0.3	0.3	0.25	0.1	0	0	0.1	0	0	0.5	0.15	0.161	0.111
50	50	0.5	0.3	0.35	0.4	0	0	0.2	0	0	0.3	0.2	0.285	0.240
50	80	0.8	0.1	0.15	0.2	0	0	0.4	0	0	0.7	0.2	0.717	0.718
50	100	0.5	0.5	0.25	0.1	0	0	0.1	0	0	0.5	0.05	0.477	0.452
50	300	0.3	0.1	0.15	0.2	0	0	0.4	0	0	0.7	0.5	0.434	0.418

Note. The running variable and outcome are correlated with correlation as 0.5 conditional on all other variables, which means the treatment-outcome relationship is confounded; the mediator and outcome are correlated with correlation as 0.5 conditional on all other variables, which means the mediator-outcome relationship is confounded.

(b). Empirical versus predicted statistical power for individual-level mediators with linear function of the running variable when the Assumption (2) sequential ignorability is violated (treatment-mediator and mediator-outcome relationship are both confounded).

Sample size		Parameters											Power	
n1	n2	a	b	c	λ_1	λ_2	λ_3	Λ_1	Λ_2	Λ_3	P	ICC	Empirical rejection rate	Predicted power
10	30	0.1	0.1	0.25	0.1	0	0	0.4	0	0	0.7	0.2	0.036	0.010
10	40	0.3	0.3	0.35	0.4	0	0	0.2	0	0	0.3	0.15	0.089	0.060
10	50	0.3	0.1	0.15	0.2	0	0	0.1	0	0	0.3	0.2	0.087	0.058
10	80	0.5	0.5	0.15	02	0	0	0.1	0	0	0.5	0.05	0.244	0.293
10	100	0.5	0.3	0.25	0.1	0	0	0.4	0	0	0.5	0.05	0.295	0.377
10	300	0.8	0.5	0.35	0.4	0	0	0.2	0	0	0.7	0.5	0.978	0.998
20	30	0.8	0.3	0.25	0.2	0	0	0.1	0	0	0.5	0.15	0.367	0.359
20	40	0.8	0.5	0.35	0.1	0	0	0.2	0	0	0.3	0.3	0.385	0.490
20	50	0.3	0.5	0.25	0.1	0	0	0.4	0	0	0.7	0.2	0.092	0.079
20	80	0.1	0.1	0.15	0.4	0	0	0.4	0	0	0.7	0.05	0.031	0.016
20	100	0.8	0.3	0.35	0.2	0	0	0.1	0	0	0.5	0.5	0.690	0.785
20	300	0.3	0.1	0.15	0.4	0	0	0.2	0	0	0.3	0.05	0.162	0.219
50	30	0.5	0.5	0.35	0.4	0	0	0.2	0	0	0.3	0.2	0.180	0.168
50	40	0.3	0.3	0.25	0.1	0	0	0.1	0	0	0.5	0.15	0.101	0.073
50	50	0.5	0.3	0.35	0.4	0	0	0.2	0	0	0.3	0.2	0.198	0.229
50	80	0.8	0.1	0.15	0.2	0	0	0.4	0	0	0.7	0.2	0.573	0.680
50	100	0.5	0.5	0.25	0.1	0	0	0.1	0	0	0.5	0.05	0.283	0.374
50	300	0.3	0.1	0.15	0.2	0	0	0.4	0	0	0.7	0.5	0.163	0.236

Note. The running variable and outcome are correlated with correlation as 0.5 conditional on all other variables, which means the treatment-outcome relationship is confounded; the running variable and mediator are correlated with correlation as 0.2 conditional on all other variables, which means the mediator-outcome relationship is confounded.

(c). Empirical versus predicted statistical power for individual-level mediators with linear function of the running variable when the Assumption (5) no treatment-by-mediator interaction is violated.

Sample size		Parameters											Power	
n1	n2	a	b	c	λ_1	λ_2	λ_3	Λ_1	Λ_2	Λ_3	P	ICC	Empirical rejection rate	Predicted power
10	30	0.1	0.1	0.25	0.1	0	0	0.4	0	0	0.7	0.2	0.072	0.033
10	40	0.3	0.3	0.35	0.4	0	0	0.2	0	0	0.3	0.15	0.143	0.098
10	50	0.3	0.1	0.15	0.2	0	0	0.1	0	0	0.3	0.2	0.086	0.067
10	80	0.5	0.5	0.15	02	0	0	0.1	0	0	0.5	0.05	0.396	0.372
10	100	0.5	0.3	0.25	0.1	0	0	0.4	0	0	0.5	0.05	0.473	0.447
10	300	0.8	0.5	0.35	0.4	0	0	0.2	0	0	0.7	0.5	0.997	0.997
20	30	0.8	0.3	0.25	0.2	0	0	0.1	0	0	0.5	0.15	0.442	0.367
20	40	0.8	0.5	0.35	0.1	0	0	0.2	0	0	0.3	0.3	0.497	0.416
20	50	0.3	0.5	0.25	0.1	0	0	0.4	0	0	0.7	0.2	0.148	0.122
20	80	0.1	0.1	0.15	0.4	0	0	0.4	0	0	0.7	0.05	0.073	0.053
20	100	0.8	0.3	0.35	0.2	0	0	0.1	0	0	0.5	0.5	0.825	0.822
20	300	0.3	0.1	0.15	0.4	0	0	0.2	0	0	0.3	0.05	0.447	0.446
50	30	0.5	0.5	0.35	0.4	0	0	0.2	0	0	0.3	0.2	0.248	0.173
50	40	0.3	0.3	0.25	0.1	0	0	0.1	0	0	0.5	0.15	0.152	0.092
50	50	0.5	0.3	0.35	0.4	0	0	0.2	0	0	0.3	0.2	0.290	0.234
50	80	0.8	0.1	0.15	0.2	0	0	0.4	0	0	0.7	0.2	0.713	0.701
50	100	0.5	0.5	0.25	0.1	0	0	0.1	0	0	0.5	0.05	0.471	0.458
50	300	0.3	0.1	0.15	0.2	0	0	0.4	0	0	0.7	0.5	0.441	0.432

Note. The interaction of the treatment and mediator is appeared in the outcome model with a coefficient of 0.5.

variable, we assume for the mediation model, the coefficient for the linear (λ_1), quadratic (λ_2), and cubic (λ_3) term is 0.05; for the outcome model, the coefficient for the linear (Λ_1), quadratic (Λ_2), and cubic (Λ_3) term is 0.04. That is, we have $n_1 = 20$, $ICC = 0.20$, $a = 0.50$, $B = 0.30$, $b_1 = 0.05$, $c' = 0.35$, $\lambda_1 = \lambda_2 = \lambda_3 = 0.05$, $\Lambda_1 = \Lambda_2 = \Lambda_3 = 0.04$, $R^2_{M^{L1}|\vec{Z}} = R^2_{M^{L2}|\vec{Z}} = R^2_{Y^{L1}|\vec{Z}} = R^2_{Y^{L2}|\vec{Z}} = 0.25,\ and\ R^2_{f(S)} = R^2_{F(S)} = 0.75$. The resulting power curves are plotted in the second panel of Figure 4. The analyses suggested that the required school-level sample size for 80% power is about 84 under the Monte Carlo test. That is, we need to sample 84 schools to achieve 80% chance to detect the existence of the mediation effect. Under the same scenario, for detecting the main effect, we need to sample only 48 schools to achieve a 0.8 statistical power. Similarly, drawing on a cluster randomized trial (as opposed to a clustered regression discontinuity assignment) and using the running variable as a covariate, we would need 76 schools to reach a power level of 0.8 to detect the school-level mediation effect.

Discussion

Prior research has repeatedly detailed the versatility and utility of the regression discontinuity design across a broad range of disciplines, interventions and policy initiatives because it allows for targeted treatment assignment while retaining a strong basis of inference. For example, prior research has suggested that the clustered regression discontinuity design can be an effective approach for interventions aimed at promoting equity in schools and communities (Hahn et al., 2001). By selectively providing treatment only to disadvantaged schools, the design allows researchers to focus and concentrate resources on those that need them most while maintaining a basis for inference about the program under study (e.g., Angrist & Pischke, 2008/2009).

Similarly, designing studies with the capacity to test the mechanisms underlying the program theory has become a prominent and critical aim of research studies. To date there is little guidance for power calculations and design strategies for probing mediation effects when considering clustered regression discontinuity designs. Our work here is intended to establish such power calculations and streamline the careful planning of clustered regression discontinuity studies to detect mediation effects.

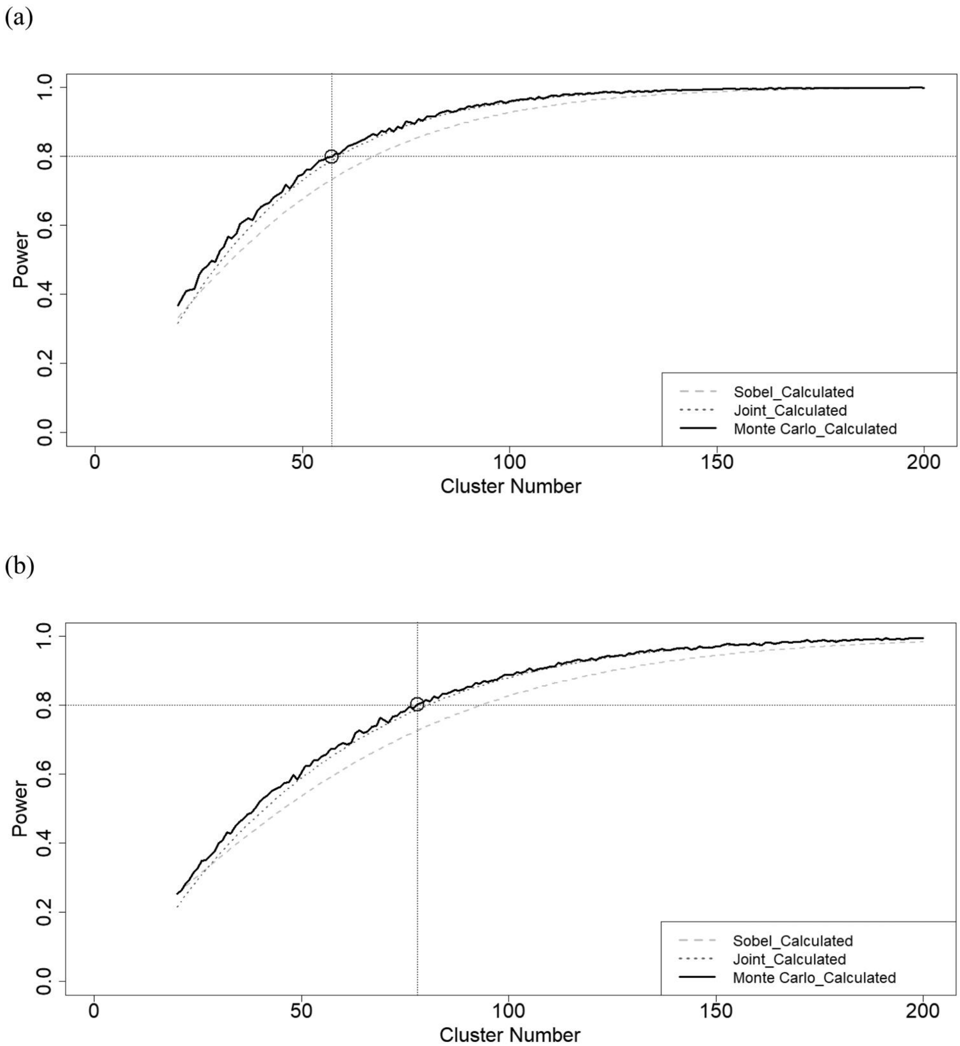


Figure 4. Power to detect (a) cluster- and (b) individual-level mediation.

Our initial analysis of the results suggests several planning considerations that largely parallel those well-documented in the literature for other designs (e.g., experiments, individual-level regression discontinuity designs; e.g., Kelcey et al. 2017). First, a balanced design in which the treatment and control groups have equal cluster-level sample sizes will generally yield the highest statistical power compared to designs with imbalanced cluster-level sample sizes. However, the loss of power resulting from unbalanced designs will typically be minimal in clustered regression discontinuity studies probing mediation. To illustrate the loss of power, Figures 5 plot the power curves as a function of the proportion of clusters assigned to the treatment under the Monte Carlo interval test (in condition 11 in Tables 3 and 4 for both the cluster- and individual-level mediator with cubic function of running variables). The plots suggest that small to medium deviations from an even split or balanced assigned have a negligible influence on power. However, once treatment assignment dips below 0.2 or above 0.8, statistical power declines quickly.

A second well-worn design strategy that is useful in clustered regression discontinuity designs is conditioning on predictive covariates. Most importantly, the tenability of the sequential ignorability assumption and unbiased estimates of mediation effects will typically require adjustment for

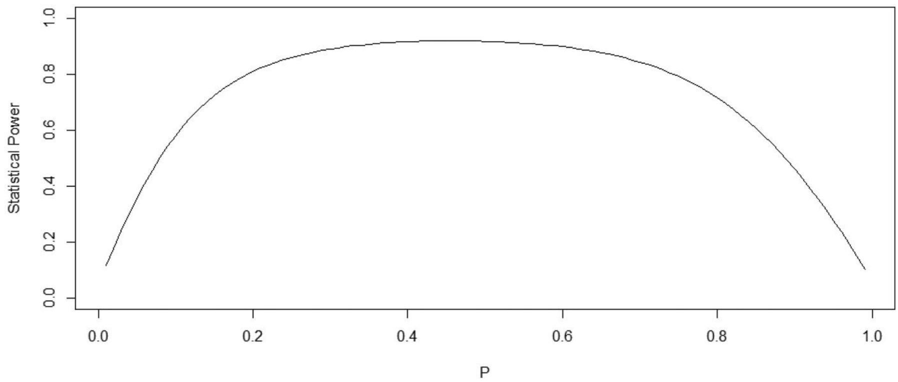


Figure 5. Statistical power with the selection of the proportion of clusters assigned to treatment (P).

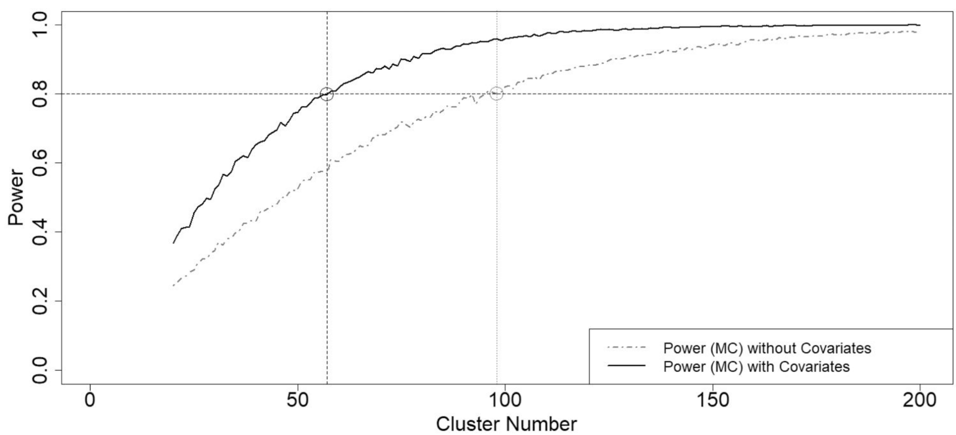
prognostic covariates regardless of whether or not they improve precision. Still, prior research has demonstrated some gains in precision and power when adjusting for prognostic covariates. However, the impact of such adjustments is much more complicated with mediation effects. Because mediation is principally composed of multiple paths, covariance adjustment can improve or reduce the precision with which we can detect mediation. Conceptually, this phenomenon arises because although explained variability in the mediator improves the power to detect the treatment-mediator *a* path, that same adjustment has the potential to deflate the power to detect the mediator-outcome *b* path because its standard error involves the ratio of the residual outcome and mediator variances.

To illustrate the differences in power that accompany covariate adjustment, Figure 6 plots the power curves as a function of cluster-level sample size for a model that controls for covariates ($R^2 = 0.25$ at both levels) versus a model that does not adjust for covariates. The plots demonstrate the increased power that accrues from adjusting for covariates when considering (a) cluster- (Figure 6(a,b) individual-level mediators (Figure 6(b)). To attain a statistical power of 0.8 (i.e., an 80% probability of detecting a significant effect if one exists) in the 2-2-1 mediation model depicted in Figure 6(a), the cluster-level sample size would need to increase by approximately 40 when not adjusting for covariates, given the same conditions as in the illustration. Conversely, in the 2-1-1 mediation model illustrated in Figure 6(b), the statistical power would only experience a slight reduction when not adjusting for covariates. Unlike the consistent power benefits gained when adjusting for covariates in detecting main effects, the gains in power of covariate adjustment in mediation effects are more complicated. In some instances when covariates explain a substantial amount of the variability in the mediator but not the outcome, the power can decrease when controlling for covariates. Figures 7(a,b) provide examples when varying the proportion of variance explained in the mediator by covariates when considering a cluster-level mediator (with 50 clusters) and an individual-level mediator (with 100 clusters). In the example with the cluster-level mediator (In Figure 7(a)), power is not monotonic and is maximized when covariates explain about 35% of the mediator variation. By contrast, the example with the individual-level demonstrated a monotonic decrease in power as the proportion of variance explained in the mediator increased holding other factors constant.

A third conclusion suggested by our results was the use of Monte Carlo test over more conventional tests. This contemporary test consistently provided the highest level of power compared to the Sobel and joint tests and is consistent with the literature (e.g., Kelcey et al., 2020). That said, the joint test provided very similar estimates of power under most conditions and is much less computationally intense.

A fourth consideration offers a cautionary design principle for nonlinear relationships. Specifically, if researchers anticipate a that the running variable has a nonlinear impact on the

(a)



(b)

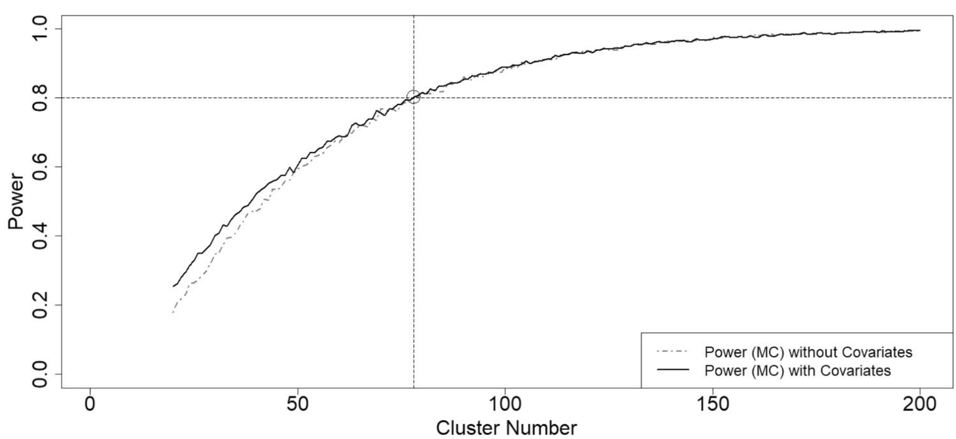
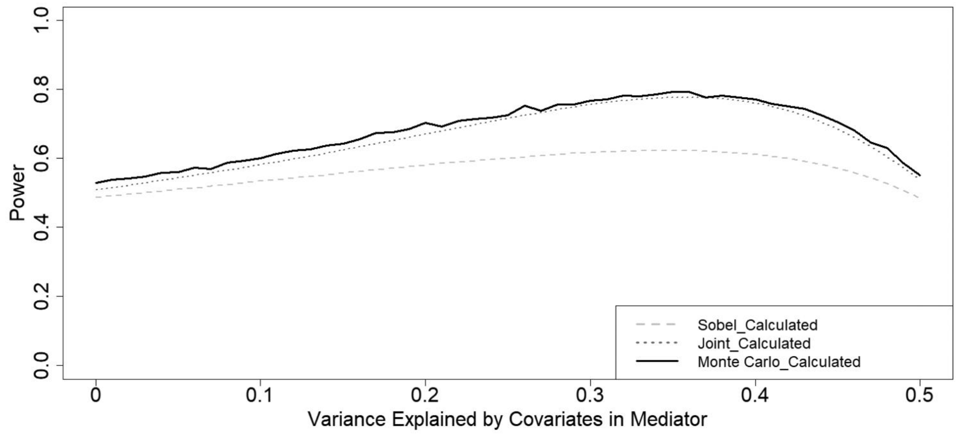


Figure 6. Comparison of statistical power for (a) cluster- and (b) individual-level mediators when adjusting and not adjusting for covariates.

mediator/outcome, the predicted power may be overestimated when the cluster-level sample size is small (e.g., less than 50). As a result, when anticipating a nonlinear relationship between the running variable and mediator/outcome, a prudent strategy would be to sample slightly more clusters than suggested by the power analysis (e.g., up to 10% more when samples include less than 50 clusters). That said, it is also important to note that while the use of a cubic function in regression discontinuity design is intended to improve the accuracy of estimated treatment effects by accounting for non-linearity in the relationship between the running variable and the outcome, it also increases the complexity of the model and may result in overfitting or unstable estimates when the sample size is small.

A fifth consideration is the sequential ignorability assumption. Our sensitivity analyses examined the predictive efficacy of our power formulas to violations of this assumption. Although the results suggested that in many cases the power predictions were still reasonable when this assumption is violated, a more important consideration is the degree of bias introduced into the estimated mediation effect through such violations. Prior literature has widely examined this from a theoretical perspective and concluded that approximating sequential ignorability to critical

(a)



(b)

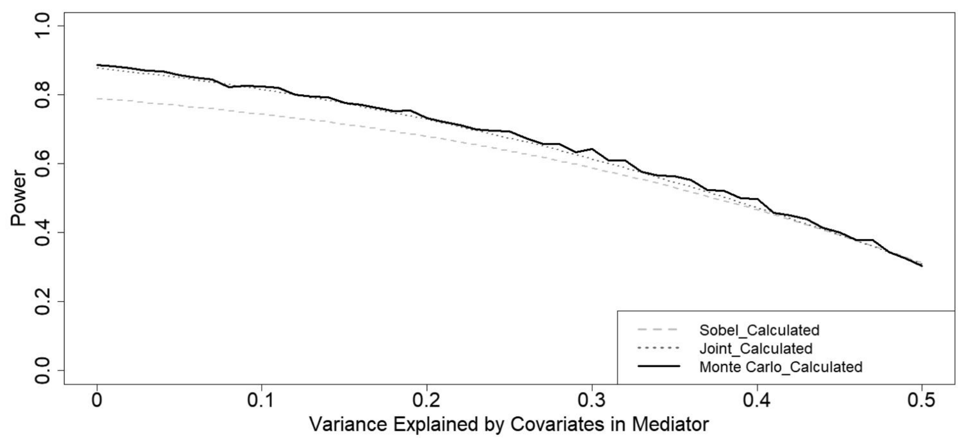


Figure 7. The impact of the variance explained by a covariate in the mediator on the statistical power to detect mediation with (a) cluster- and (b) individual-level mediators.

to the validity of inferences (e.g., VanderWeele et al., 2014). When amalgamated within the larger context of mediation analysis, our results simply suggested that we can precisely predict the power to detect a biased mediation effect. As a result, we emphasize the need for designs to proactively incorporate confounding variables and to carefully consider the plausibility of the sequentially ignorability assumption in a study. A final consideration is to consider a range of plausible values for key design parameters. For example, the value and accuracy of power predictions is contingent upon how well the assumed parameters values approximate the true values. Even though recent literature has developed an increasingly large collection of empirical estimates of these values for a broad range of outcomes and mediators, in most instances these values will not be without error (e.g., (Jacob et al., 2012); Kelcey and Phelps, 2013). As a result, considering plausible ranges of parameter values can help probe the sensitivity of design choices to parameter value misspecifications and provide a more comprehensive assessment of requisite sample sizes.

Although our results provide some initial tools and strategies to design clustered regression discontinuity studies, there are some limitations of our work and directions for future research. First, our study only considered a sharp clustered regression discontinuity design. Tools for fuzzy clustered regression discontinuity designs are a potentially important direction for future research because they often map onto practice better. Second, we considered two-level clustered regression discontinuity designs in this study, but many studies in education involve three or more levels. For example, many educational studies directly involve teachers as the primary vehicle or mechanism through which a program is delivered. In these settings, investigating mediation through a teacher mediator variable such as instruction may require the introduction of an intermediate level. Third, our analysis considered only a single mediator with no treatment by mediator interaction. Exploring interactions between treatment and mediators in future research could uncover valuable insights into the nuanced pathways through which the treatment effect unfolds. Multiple mediator and interactive models are also important considerations and directions for future research. Investigating the combined and decomposed effects of multiple mediators can offer an even richer comprehension of the underlying mechanisms driving treatment outcomes.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This study was supported by the National Science Foundation [grant numbers 1552535,1760884]. The National Science Foundation had no role or involvement in the conduct of the research or the preparation of the manuscript.

References

Angrist, J. D., & Pischke, J. (2008/2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press. https://doi.org/10.1515/9781400829828

Bloom, H. S. (2012). Modern regression discontinuity analysis. *Journal of Research on Educational Effectiveness*, *5*(1), 43–82. https://doi.org/10.1080/19345747.2011.578707

Bollen, K. A. (1987). Total, Direct, and Indirect Effects in Structural Equation Models. *Sociological Methodology*, *17*, 37. https://doi.org/10.2307/271028

Bonell, C., Allen, E., Warren, E., McGowan, J., Bevilacqua, L., Jamal, F., Legood, R., Wiggins, M., Opondo, C., Mathiot, A., Sturgess, J., Fletcher, A., Sadique, Z., Elbourne, D., Christie, D., Bond, L., Scott, S., & Viner, R. M. (2018). Effects of the learning together intervention on bullying and aggression in english secondary schools (INCLUSIVE): A cluster randomised controlled trial. *Lancet (London, England)*, *392*(10163), 2452–2464. https://doi.org/10.1016/S0140-6736(18)31782-3

Bonell, C., Allen, E., Opondo, C., Warren, E., Elbourne, D. R., Sturgess, J., Bevilacqua, L., McGowan, J., Mathiot, A., & Viner, R. M. (2019). Examining intervention mechanisms of action using mediation analysis within a randomised trial of a whole-school health intervention. *Journal of Epidemiology and Community Health*, *73*(5), 455–464. https://doi.org/10.1136/jech-2018-211443

Calcagno, J. C., & Long, B. T. (2008). *The impact of postsecondary remediation using a regression discontinuity approach: Addressing endogenous sorting and noncompliance*. National Bureau of Economic Research. https://doi.org/10.3386/w14194

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* (2nd ed.) L. Erlbaum Associates.

Cook, T. D. (2008). "Waiting for life to arrive": A history of the regression-discontinuity design in psychology, statistics and economics. *Journal of Econometrics*, *142*(2), 636–654. https://doi.org/10.1016/j.jeconom.2007.05.002

Hahn, J., Todd, P., & Klaauw, W. (2001). Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica*, *69*(1), 201–209. https://doi.org/10.1111/1468-0262.00183

Imbens, G. W., & Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, *142*(2), 615–635. https://doi.org/10.1016/j.jeconom.2007.05.001

Jacob, Robin, Zhu, Pei, Somers, Marie-Andrée, Bloom, Howard. (2012). *A practical guide to regression discontinuity*. MDRC.

Kelcey, Ben., & Phelps, G. (2013). Strategies for Improving Power in School-Randomized Studies of Professional Development. *Evaluation Review*, *37*(6), 520–554. https://doi.org/10.1177/0193841X14528906

Kelcey, Ben., & Shen, Z. (2017). Planning community-based assessments of HIV educational intervention programs in sub-Saharan Africa. *Health Education Research*, *32*(4), 353–363. https://doi.org/10.1093/her/cyx04628854572

Kelcey, B., Dong, N., Spybrook, J., & Cox, K. (2017). Statistical power for causally defined indirect effects in group-randomized trials with individual-level mediators. *Journal of Educational and Behavioral Statistics*, *42*(5), 499–530. https://doi.org/10.3102/1076998617695506

Kelcey, B., Dong, N., Spybrook, J., & Shen, Z. (2017). Experimental power for indirect effects in group-randomized studies with group-level mediators. *Multivariate Behavioral Research*, *52*(6), 699–719. https://doi.org/10.1080/00273171.2017.1356212

Kelcey, B., Spybrook, J., Dong, N., & Bai, F. (2020). Cross-level mediation in school-randomized studies of teacher development: Experimental design and power. *Journal of Research on Educational Effectiveness*, *13*(3), 459–487. https://doi.org/10.1080/19345747.2020.1726540

Kelcey, B., Xie, Y., Spybrook, J., & Dong, N. (2020). Power and sample size determination for multilevel mediation in three-level cluster-randomized trials. *Multivariate Behavioral Research*, *56*(3), 496–513. https://doi.org/10.1080/00273171.2020.1738910

Kreft, I. G. G., de Leeuw, J., & Aiken, L. S. (1995). The effect of different forms of centering in hierarchical linear models. *Multivariate Behavioral Research*, *30*(1), 1–21. https://doi.org/10.1207/s15327906mbr3001_1

MacKinnon, D. P., Lockwood, C. M., Brown, C. H., Wang, W., & Hoffman, J. M. (2007). The intermediate endpoint effect in logistic and probit regression. *Clinical Trials (London, England)*, *4*(5), 499–513. https://doi.org/10.1177/174077450708343417942466

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*(1), 83–104. https://doi.org/10.1037/1082-989X.7.1.83

MacKinnon, D. P. (2012). *Introduction to statistical mediation analysis*. Routledge.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, *142*(2), 698–714. https://doi.org/10.1016/j.jeconom.2007.05.005

Pituch, K. A., & Stapleton, L. M. (2012). Distinguishing between cross- and cluster-level mediation processes in the cluster randomized trial. *Sociological Methods & Research*, *41*(4), 630–670. https://doi.org/10.1177/0049124112460380

Preacher, K. J., & Selig, J. P. (2012). Advantages of monte carlo confidence intervals for indirect effects. *Communication Methods and Measures*, *6*(2), 77–98. https://doi.org/10.1080/19312458.2012.679848

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods.* (Vol. 1). sage.

Schochet, P. Z. (2009). Statistical power for regression discontinuity designs in education evaluations. *Journal of Educational and Behavioral Statistics*, *34*(2), 238–266. https://doi.org/10.3102/1076998609332748

Schochet, P., Cook, T., Deke, J., Imbens, G., Lockwood, J. R., Porter, J., Smith, J., & What Works Clearinghouse (ED) (2010). *Standards for regression discontinuity designs*. What Works Clearinghouse.

Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. *Sociological Methodology*, *13*, 290–312. https://doi.org/10.2307/270723

VanderWeele, T. J., Vansteelandt, S., & Robins, J. M. (2014). Effect Decomposition in the Presence of an Exposure-Induced Mediator-Outcome Confounder. *Epidemiology*, *25*(2), 300–306. https://doi.org/10.1097/EDE.0000000000000034

VanderWeele, T. J. (2010). Direct and indirect effects for neighborhood-based clustered and longitudinal data. *Sociological Methods & Research*, *38*(4), 515–544. https://doi.org/10.1177/0049124110366236

Zhang, Z., Zyphur, M. J., & Preacher, K. J. (2009). Testing multilevel mediation using hierarchical linear models: Problems and solutions. *Organizational Research Methods*, *12*(4), 695–719. https://doi.org/10.1177/1094428108327450

Appendix A. Path decomposition for cluster-level mediator

We sequentially increase the conditional partitions, and finally, the formula will be able to express the R squared of total used variables. For the cluster level R squared of outcome, the formula will be as follows

$$R^2_{Y^{L2}} = R^2_{Y^{L2}|\vec{Z}} + R^2_{Y^{L2}T^{RD}S|\vec{Z}} + R^2_{Y^{L2}M|T^{RD}S\vec{Z}} \tag{A1}$$

The right part of the expression is constructed as a sequential addition of R squared of the left subscript part given or controlling the right subscript part. The vector $\vec{Z}$ is an arbitrary vector which represents the covariates. Take $R^2_{Y^{L2}|\vec{Z}}$ as an example, it represents the outcome on the cluster level explained by the covariates, which can be obtained by prior studies. The second symbol on the right arm $R^2_{Y^{L2}T^{RD}S|\vec{Z}}$ represents the outcome on the cluster level explained by the function of the running variable and the treatment while controlling for the covariates. Because of the correlation between the treatment and the functions of the running variable, we consider them together. The expected value of $R^2_{Y^{L2}TS|\vec{Z}}$ can be reduced as follows

$$R^2_{Y^{L2}TS|\vec{Z}} = \frac{b\sigma^2_{f(S)} + \sigma^2_{F(S)} + 2b\sigma_{f(S)F(s)}}{\tau^2} + \frac{(ab+c')^2P(1-P)}{\tau^2} + \frac{2\left(b\sigma_{f(S)T} + \sigma_{F(S)T}\right)(ab+c')}{\tau^2} \tag{A2}$$

We include the covariances between the treatment and the function of the running variable for both the mediator model, $\sigma_{f(S)T}$, and outcome model, $\sigma_{F(S)T}$, and the covariance between $f(S)$ and $F(S)$, $\sigma_{f(S)F(s)}$, in this equation. The last term on the right hand $R^2_{Y^{L2}M|T^{RD}S\vec{Z}}$ is the level two outcome variance explained by the mediator while controlling for the treatment, the cutoff variable, and the covariates. The expected expression of it can be reconstructed as follows

$$R^2_{Y^{L2}M|T^{RD}S\vec{Z}} = \frac{b^2\sigma^2_M\left(1-R^2_M\right)}{\tau^2} \tag{A3}$$

The total variance of the second level explained in the outcome model is the assembling of all parts of the equations as follows

$$R^2_{Y^{L2}} = R^2_{Y^{L2}|\vec{Z}} + \frac{b\sigma^2_{f(S)} + \sigma^2_{F(S)} + 2b\sigma_{f(S)F(s)}}{\tau^2} + \frac{(ab+c')^2P(1-P)}{\tau^2} + \frac{2\left(b\sigma_{f(S)T} + \sigma_{F(S)T}\right)(ab+c')}{\tau^2} + \frac{b^2\sigma^2_M\left(1-R^2_{M^{L2}}\right)}{\tau^2}. \tag{A4}$$

This expression contains the variance explained by the covariates that can be obtained from prior studies and the variance explained by the other variables that we have reconstructed. With the same approach, we can construct the R squared for the mediation effect as follows

$$R^2_M = R^2_{M|\vec{Z}} + \frac{\sigma_{f(S)}}{\tau^2_M} + \frac{a^2P(1-P)}{\tau^2_M} + \frac{2a\sigma_{f(S)T}}{\tau^2_M}. \tag{A5}$$

We can use ICC and a (standardized) total error variance to replace the cluster and individual level output error variances, which are more accessible for researchers.

Appendix B. Path decomposition for student-level mediator

For the student-level (2-1-1) mediation, we followed the approach in the 2-2-1 mediation part to decompose the R squared with sequential addition of conditional R squared of the elements. The cluster level R squared of the outcome is the same as the 2-2-1 case, the formula will be as follows

$$R^2_{Y^{L2}} = R^2_{Y^{L2}|\vec{Z}} + R^2_{Y^{L2}T^{RD}S|\vec{Z}} + R^2_{Y^{L2}M|T^{RD}S\vec{Z}} \tag{16}$$

The expected value of $R^2_{Y^{L2}S|\vec{Z}}$ can be reduced as follows

$$R^2_{Y^{L2}T^{RD}S|\vec{Z}} = \frac{B\sigma^2_{f(S)} + \sigma^2_{F(S)} + 2B\sigma_{f(S)F(s)}}{\tau^2} + \frac{(aB+c')^2P(1-P)}{\tau^2} + \frac{2\left(B\sigma_{f(S)T} + \sigma_{F(S)T}\right)(ab+c')}{\tau^2} \tag{17}$$

The expected expression of $R^2_{Y^{L2}M|T^{RD}S\vec{Z}}$ can be reconstructed as follows

$$R^2_{Y^{L2}M|T^{RD}S\vec{Z}} = \frac{B^2\left(\tau^2_M\left(1 - R^2_{M^{L2}}\right) + \frac{\sigma^2_M\left(1-R^2_{M^{L1}}\right)}{n_1}\right)}{\tau^2} \tag{18}$$

The total variance of the second level explained in the outcome model is the assembling of all parts of the equations as follows.

$$R^2_{Y^{L2}} = R^2_{Y^{L2}|\vec{Z}} + \frac{B\sigma^2_{f(S)} + \sigma^2_{F(S)} + 2B\sigma_{f(S)F(s)}}{\tau^2} + \frac{(aB + c')^2 P(1-P)}{\tau^2} + \frac{2\left(B\sigma_{f(S)T} + \sigma_{F(S)T}\right)(ab + c')}{\tau^2} + \frac{B^2\left(\tau^2_M\left(1 - R^2_{M^{L2}}\right) + \frac{\sigma^2_M\left(1-R^2_{M^{L1}}\right)}{n_1}\right)}{\tau^2} \tag{19}$$

For the R squared of the individual level outcome, the expected expression is as follows

$$R^2_{Y^{L1}} = R^2_{Y^{L1}|\vec{Z}} + \frac{{b_1}^2\sigma^2_M\left(1 - R^2_{M^{L1}}\right)}{\sigma^2_Y} \tag{20}$$

With the same approach, we can construct the R squared for the cluster-level aggregated mediation effect,

$$R^2_{M^{L2}} = R^2_{M^{L2}|\vec{Z}} + \frac{\sigma^2_{f(S)}}{\tau^2_M} + \frac{a^2 P(1-P)}{\tau^2_M} + \frac{2a\sigma_{f(S)T}}{\tau^2_M} \tag{21}$$

Like 2-2-1 mediation, we can also use ICC and a (standardized) total error variance to replace the cluster and individual level output error variances, which is more accessible for researchers.

Appendix B. Detailed tables for average absolute biases.

Table B1. (Full Version).

Sample size		Parameters											Mediation error variance			
n1	n2	a	b	c	λ_1	λ_2	λ_3	Λ_1	Λ_2	Λ_3	P	ICC	Empirical variance	Predicted variance	Absolute bias	Average absolute bias
(a) Average absolute bias in sampling variability for cluster-level mediator for Condition 1.																
10	20	0.3	0.1	0.15	0.2	0.4	0.2	0.1	0.2	0.1	0.3	0.2	0.0108	0.0038	0.0069	0.0004
10	30	0.3	0.1	0.15	0.2	0.4	0.2	0.1	0.2	0.1	0.3	0.2	0.0050	0.0027	0.0023	
10	40	0.3	0.1	0.15	0.2	0.4	0.2	0.1	0.2	0.1	0.3	0.2	0.0031	0.0020	0.0011	
10	50	0.3	0.1	0.15	0.2	0.4	0.2	0.1	0.2	0.1	0.3	0.2	0.0022	0.0017	0.0006	
10	60	0.3	0.1	0.15	0.2	0.4	0.2	0.1	0.2	0.1	0.3	0.2	0.0017	0.0014	0.0003	
10	70	0.3	0.1	0.15	0.2	0.4	0.2	0.1	0.2	0.1	0.3	0.2	0.0014	0.0011	0.0002	
10	80	0.3	0.1	0.15	0.2	0.4	0.2	0.1	0.2	0.1	0.3	0.2	0.0011	0.0010	0.0001	
10	90	0.3	0.1	0.15	0.2	0.4	0.2	0.1	0.2	0.1	0.3	0.2	0.0010	0.0009	0.0001	
10	100	0.3	0.1	0.15	0.2	0.4	0.2	0.1	0.2	0.1	0.3	0.2	0.0009	0.0008	0.0001	
10	110	0.3	0.1	0.15	0.2	0.4	0.2	0.1	0.2	0.1	0.3	0.2	0.0008	0.0007	0.0000	
10	120	0.3	0.1	0.15	0.2	0.4	0.2	0.1	0.2	0.1	0.3	0.2	0.0007	0.0007	0.0000	
10	130	0.3	0.1	0.15	0.2	0.4	0.2	0.1	0.2	0.1	0.3	0.2	0.0007	0.0006	0.0000	
10	140	0.3	0.1	0.15	0.2	0.4	0.2	0.1	0.2	0.1	0.3	0.2	0.0006	0.0006	0.0000	
10	150	0.3	0.1	0.15	0.2	0.4	0.2	0.1	0.2	0.1	0.3	0.2	0.0005	0.0005	0.0000	
10	160	0.3	0.1	0.15	0.2	0.4	0.2	0.1	0.2	0.1	0.3	0.2	0.0005	0.0005	0.0000	
10	170	0.3	0.1	0.15	0.2	0.4	0.2	0.1	0.2	0.1	0.3	0.2	0.0005	0.0005	0.0000	
10	180	0.3	0.1	0.15	0.2	0.4	0.2	0.1	0.2	0.1	0.3	0.2	0.0004	0.0004	0.0000	
10	190	0.3	0.1	0.15	0.2	0.4	0.2	0.1	0.2	0.1	0.3	0.2	0.0004	0.0004	0.0000	
10	200	0.3	0.1	0.15	0.2	0.4	0.2	0.1	0.2	0.1	0.3	0.2	0.0004	0.0004	0.0000	
(b) Average absolute bias in sampling variability for cluster-level mediator for Condition 2.																
10	20	0.5	0.3	0.25	0.1	0.2	0.1	0.4	0.1	0.2	0.5	0.05	0.0344	0.0216	0.0128	0.0013
10	30	0.5	0.3	0.25	0.1	0.2	0.1	0.4	0.1	0.2	0.5	0.05	0.0198	0.0149	0.0049	
10	40	0.5	0.3	0.25	0.1	0.2	0.1	0.4	0.1	0.2	0.5	0.05	0.0143	0.0108	0.0035	
10	50	0.5	0.3	0.25	0.1	0.2	0.1	0.4	0.1	0.2	0.5	0.05	0.0107	0.0088	0.0019	
10	60	0.5	0.3	0.25	0.1	0.2	0.1	0.4	0.1	0.2	0.5	0.05	0.0087	0.0070	0.0017	
10	70	0.5	0.3	0.25	0.1	0.2	0.1	0.4	0.1	0.2	0.5	0.05	0.0071	0.0062	0.0010	
10	80	0.5	0.3	0.25	0.1	0.2	0.1	0.4	0.1	0.2	0.5	0.05	0.0063	0.0054	0.0010	
10	90	0.5	0.3	0.25	0.1	0.2	0.1	0.4	0.1	0.2	0.5	0.05	0.0056	0.0048	0.0007	
10	100	0.5	0.3	0.25	0.1	0.2	0.1	0.4	0.1	0.2	0.5	0.05	0.0047	0.0044	0.0003	
10	110	0.5	0.3	0.25	0.1	0.2	0.1	0.4	0.1	0.2	0.5	0.05	0.0044	0.0040	0.0004	
10	120	0.5	0.3	0.25	0.1	0.2	0.1	0.4	0.1	0.2	0.5	0.05	0.0039	0.0036	0.0003	
10	130	0.5	0.3	0.25	0.1	0.2	0.1	0.4	0.1	0.2	0.5	0.05	0.0036	0.0034	0.0002	
10	140	0.5	0.3	0.25	0.1	0.2	0.1	0.4	0.1	0.2	0.5	0.05	0.0037	0.0031	0.0005	
10	150	0.5	0.3	0.25	0.1	0.2	0.1	0.4	0.1	0.2	0.5	0.05	0.0033	0.0029	0.0004	
10	160	0.5	0.3	0.25	0.1	0.2	0.1	0.4	0.1	0.2	0.5	0.05	0.0031	0.0028	0.0003	
10	170	0.5	0.3	0.25	0.1	0.2	0.1	0.4	0.1	0.2	0.5	0.05	0.0028	0.0026	0.0002	
10	180	0.5	0.3	0.25	0.1	0.2	0.1	0.4	0.1	0.2	0.5	0.05	0.0026	0.0025	0.0001	
10	190	0.5	0.3	0.25	0.1	0.2	0.1	0.4	0.1	0.2	0.5	0.05	0.0025	0.0024	0.0001	
10	200	0.5	0.3	0.25	0.1	0.2	0.1	0.4	0.1	0.2	0.5	0.05	0.0023	0.0022	0.0001	
(c) Average absolute bias in sampling variability for cluster-level mediator for Condition 3.																
10	20	0.8	0.5	0.35	0.4	0.1	0.2	0.2	0.4	0.1	0.7	0.5	0.1796	0.1137	0.0659	0.0040
10	30	0.8	0.5	0.35	0.4	0.1	0.2	0.2	0.4	0.1	0.7	0.5	0.0931	0.0743	0.0188	
10	40	0.8	0.5	0.35	0.4	0.1	0.2	0.2	0.4	0.1	0.7	0.5	0.0652	0.0551	0.0101	
10	50	0.8	0.5	0.35	0.4	0.1	0.2	0.2	0.4	0.1	0.7	0.5	0.0496	0.0439	0.0057	
10	60	0.8	0.5	0.35	0.4	0.1	0.2	0.2	0.4	0.1	0.7	0.5	0.0400	0.0368	0.0032	
10	70	0.8	0.5	0.35	0.4	0.1	0.2	0.2	0.4	0.1	0.7	0.5	0.0327	0.0308	0.0019	
10	80	0.8	0.5	0.35	0.4	0.1	0.2	0.2	0.4	0.1	0.7	0.5	0.0286	0.0279	0.0008	
10	90	0.8	0.5	0.35	0.4	0.1	0.2	0.2	0.4	0.1	0.7	0.5	0.0252	0.0246	0.0006	
10	100	0.8	0.5	0.35	0.4	0.1	0.2	0.2	0.4	0.1	0.7	0.5	0.0225	0.0223	0.0002	
10	110	0.8	0.5	0.35	0.4	0.1	0.2	0.2	0.4	0.1	0.7	0.5	0.0200	0.0203	0.0002	
10	120	0.8	0.5	0.35	0.4	0.1	0.2	0.2	0.4	0.1	0.7	0.5	0.0188	0.0184	0.0003	
10	130	0.8	0.5	0.35	0.4	0.1	0.2	0.2	0.4	0.1	0.7	0.5	0.0165	0.0172	0.0007	
10	140	0.8	0.5	0.35	0.4	0.1	0.2	0.2	0.4	0.1	0.7	0.5	0.0160	0.0157	0.0003	
10	150	0.8	0.5	0.35	0.4	0.1	0.2	0.2	0.4	0.1	0.7	0.5	0.0141	0.0146	0.0005	
10	160	0.8	0.5	0.35	0.4	0.1	0.2	0.2	0.4	0.1	0.7	0.5	0.0133	0.0138	0.0005	
10	170	0.8	0.5	0.35	0.4	0.1	0.2	0.2	0.4	0.1	0.7	0.5	0.0123	0.0130	0.0007	
10	180	0.8	0.5	0.35	0.4	0.1	0.2	0.2	0.4	0.1	0.7	0.5	0.0117	0.0123	0.0006	

(*continued*)

Table B1. Continued.

Sample size		Parameters											Mediation error variance			
n1	n2	a	b	c	λ_1	λ_2	λ_3	Λ_1	Λ_2	Λ_3	P	ICC	Empirical variance	Predicted variance	Absolute bias	Average absolute bias
10	190	0.8	0.5	0.35	0.4	0.1	0.2	0.2	0.4	0.1	0.7	0.5	0.0113	0.0116	0.0003	
10	200	0.8	0.5	0.35	0.4	0.1	0.2	0.2	0.4	0.1	0.7	0.5	0.0106	0.0112	0.0007	
(d) Average absolute bias in sampling variability for cluster-level mediator for Condition 4.																
20	20	0.3	0.1	0.15	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0066	0.0032	0.0034	0.0003
20	30	0.3	0.1	0.15	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0034	0.0021	0.0013	
20	40	0.3	0.1	0.15	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0022	0.0016	0.0006	
20	50	0.3	0.1	0.15	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0018	0.0014	0.0004	
20	60	0.3	0.1	0.15	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0014	0.0011	0.0003	
20	70	0.3	0.1	0.15	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0011	0.0009	0.0002	
20	80	0.3	0.1	0.15	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0010	0.0008	0.0002	
20	90	0.3	0.1	0.15	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0008	0.0007	0.0001	
20	100	0.3	0.1	0.15	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0008	0.0006	0.0001	
20	110	0.3	0.1	0.15	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0007	0.0006	0.0001	
20	120	0.3	0.1	0.15	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0006	0.0005	0.0001	
20	130	0.3	0.1	0.15	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0005	0.0005	0.0000	
20	140	0.3	0.1	0.15	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0005	0.0005	0.0000	
20	150	0.3	0.1	0.15	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0005	0.0004	0.0000	
20	160	0.3	0.1	0.15	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0004	0.0004	0.0000	
20	170	0.3	0.1	0.15	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0004	0.0004	0.0000	
20	180	0.3	0.1	0.15	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0004	0.0004	0.0000	
20	190	0.3	0.1	0.15	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0004	0.0003	0.0000	
20	200	0.3	0.1	0.15	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0004	0.0003	0.0000	
(e) Average absolute bias in sampling variability for cluster-level mediator for Condition 5.																
20	20	0.3	0.5	0.25	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.1362	0.0637	0.0725	0.0071
20	30	0.3	0.5	0.25	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0765	0.0444	0.0322	
20	40	0.3	0.5	0.25	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0529	0.0349	0.0180	
20	50	0.3	0.5	0.25	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0384	0.0291	0.0094	
20	60	0.3	0.5	0.25	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0321	0.0245	0.0076	
20	70	0.3	0.5	0.25	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0262	0.0203	0.0059	
20	80	0.3	0.5	0.25	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0231	0.0185	0.0046	
20	90	0.3	0.5	0.25	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0207	0.0168	0.0039	
20	100	0.3	0.5	0.25	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0181	0.0154	0.0027	
20	110	0.3	0.5	0.25	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0165	0.0137	0.0028	
20	120	0.3	0.5	0.25	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0145	0.0125	0.0019	
20	130	0.3	0.5	0.25	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0128	0.0117	0.0011	
20	140	0.3	0.5	0.25	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0122	0.0110	0.0012	
20	150	0.3	0.5	0.25	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0112	0.0100	0.0012	
20	160	0.3	0.5	0.25	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0105	0.0096	0.0009	
20	170	0.3	0.5	0.25	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0097	0.0091	0.0006	
20	180	0.3	0.5	0.25	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0095	0.0086	0.0009	
20	190	0.3	0.5	0.25	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0087	0.0082	0.0005	
20	200	0.3	0.5	0.25	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0083	0.0078	0.0006	
(f) Average absolute bias in sampling variability for cluster-level mediator for Condition 6.																
20	20	0.8	0.3	0.35	0.2	0.1	0.1	0.1	0.4	0.2	0.5	0.5	0.0973	0.0627	0.0346	0.0023
20	30	0.8	0.3	0.35	0.2	0.1	0.1	0.1	0.4	0.2	0.5	0.5	0.0513	0.0415	0.0098	
20	40	0.8	0.3	0.35	0.2	0.1	0.1	0.1	0.4	0.2	0.5	0.5	0.0357	0.0288	0.0069	
20	50	0.8	0.3	0.35	0.2	0.1	0.1	0.1	0.4	0.2	0.5	0.5	0.0260	0.0226	0.0034	
20	60	0.8	0.3	0.35	0.2	0.1	0.1	0.1	0.4	0.2	0.5	0.5	0.0210	0.0191	0.0019	
20	70	0.8	0.3	0.35	0.2	0.1	0.1	0.1	0.4	0.2	0.5	0.5	0.0176	0.0162	0.0015	
20	80	0.8	0.3	0.35	0.2	0.1	0.1	0.1	0.4	0.2	0.5	0.5	0.0151	0.0143	0.0007	
20	90	0.8	0.3	0.35	0.2	0.1	0.1	0.1	0.4	0.2	0.5	0.5	0.0130	0.0125	0.0006	
20	100	0.8	0.3	0.35	0.2	0.1	0.1	0.1	0.4	0.2	0.5	0.5	0.0118	0.0111	0.0007	
20	110	0.8	0.3	0.35	0.2	0.1	0.1	0.1	0.4	0.2	0.5	0.5	0.0105	0.0103	0.0003	
20	120	0.8	0.3	0.35	0.2	0.1	0.1	0.1	0.4	0.2	0.5	0.5	0.0094	0.0094	0.0000	
20	130	0.8	0.3	0.35	0.2	0.1	0.1	0.1	0.4	0.2	0.5	0.5	0.0085	0.0086	0.0001	
20	140	0.8	0.3	0.35	0.2	0.1	0.1	0.1	0.4	0.2	0.5	0.5	0.0081	0.0080	0.0000	
20	150	0.8	0.3	0.35	0.2	0.1	0.1	0.1	0.4	0.2	0.5	0.5	0.0073	0.0076	0.0002	
20	160	0.8	0.3	0.35	0.2	0.1	0.1	0.1	0.4	0.2	0.5	0.5	0.0070	0.0071	0.0001	
20	170	0.8	0.3	0.35	0.2	0.1	0.1	0.1	0.4	0.2	0.5	0.5	0.0065	0.0065	0.0001	
20	180	0.8	0.3	0.35	0.2	0.1	0.1	0.1	0.4	0.2	0.5	0.5	0.0059	0.0062	0.0003	
20	190	0.8	0.3	0.35	0.2	0.1	0.1	0.1	0.4	0.2	0.5	0.5	0.0058	0.0059	0.0002	
20	200	0.8	0.3	0.35	0.2	0.1	0.1	0.1	0.4	0.2	0.5	0.5	0.0052	0.0056	0.0004	

Table B2. (Full Version).

Sample size		Parameters												Mediation error variance			
n1	n2	a	B	b1	c	λ_1	λ_2	λ_3	Λ_1	Λ_2	Λ_3	P	ICC	Empirical Variance	Predicted Variance	Absolute Bias	Average Absolute Bias
(a) Average absolute bias in sampling variability for individual-level mediator in Condition 1.																	
10	20	0.3	0.1	0.1	0.15	0.2	0.4	0.2	0.1	0.2	0.1	0.3	0.2	0.0168	0.0054	0.0114	0.0008
10	30	0.3	0.1	0.1	0.15	0.2	0.4	0.2	0.1	0.2	0.1	0.3	0.2	0.0078	0.0037	0.0041	
10	40	0.3	0.1	0.1	0.15	0.2	0.4	0.2	0.1	0.2	0.1	0.3	0.2	0.0048	0.0028	0.0020	
10	50	0.3	0.1	0.1	0.15	0.2	0.4	0.2	0.1	0.2	0.1	0.3	0.2	0.0035	0.0023	0.0012	
10	60	0.3	0.1	0.1	0.15	0.2	0.4	0.2	0.1	0.2	0.1	0.3	0.2	0.0026	0.0019	0.0008	
10	70	0.3	0.1	0.1	0.15	0.2	0.4	0.2	0.1	0.2	0.1	0.3	0.2	0.0021	0.0016	0.0005	
10	80	0.3	0.1	0.1	0.15	0.2	0.4	0.2	0.1	0.2	0.1	0.3	0.2	0.0018	0.0014	0.0004	
10	90	0.3	0.1	0.1	0.15	0.2	0.4	0.2	0.1	0.2	0.1	0.3	0.2	0.0015	0.0013	0.0003	
10	100	0.3	0.1	0.1	0.15	0.2	0.4	0.2	0.1	0.2	0.1	0.3	0.2	0.0014	0.0011	0.0003	
10	110	0.3	0.1	0.1	0.15	0.2	0.4	0.2	0.1	0.2	0.1	0.3	0.2	0.0012	0.0010	0.0002	
10	120	0.3	0.1	0.1	0.15	0.2	0.4	0.2	0.1	0.2	0.1	0.3	0.2	0.0011	0.0009	0.0002	
10	130	0.3	0.1	0.1	0.15	0.2	0.4	0.2	0.1	0.2	0.1	0.3	0.2	0.0010	0.0009	0.0001	
10	140	0.3	0.1	0.1	0.15	0.2	0.4	0.2	0.1	0.2	0.1	0.3	0.2	0.0009	0.0008	0.0001	
10	150	0.3	0.1	0.1	0.15	0.2	0.4	0.2	0.1	0.2	0.1	0.3	0.2	0.0009	0.0007	0.0001	
10	160	0.3	0.1	0.1	0.15	0.2	0.4	0.2	0.1	0.2	0.1	0.3	0.2	0.0008	0.0007	0.0001	
10	170	0.3	0.1	0.1	0.15	0.2	0.4	0.2	0.1	0.2	0.1	0.3	0.2	0.0007	0.0007	0.0001	
10	180	0.3	0.1	0.1	0.15	0.2	0.4	0.2	0.1	0.2	0.1	0.3	0.2	0.0007	0.0006	0.0001	
10	190	0.3	0.1	0.1	0.15	0.2	0.4	0.2	0.1	0.2	0.1	0.3	0.2	0.0006	0.0006	0.0001	
10	200	0.3	0.1	0.1	0.15	0.2	0.4	0.2	0.1	0.2	0.1	0.3	0.2	0.0006	0.0006	0.0001	
(b) Average absolute bias in sampling variability for individual-level mediator in Condition 2.																	
10	20	0.5	0.3	0.05	0.25	0.1	0.2	0.1	0.4	0.1	0.2	0.5	0.05	0.0224	0.0147	0.0077	0.0012
10	30	0.5	0.3	0.05	0.25	0.1	0.2	0.1	0.4	0.1	0.2	0.5	0.05	0.0130	0.0087	0.0044	
10	40	0.5	0.3	0.05	0.25	0.1	0.2	0.1	0.4	0.1	0.2	0.5	0.05	0.0092	0.0068	0.0024	
10	50	0.5	0.3	0.05	0.25	0.1	0.2	0.1	0.4	0.1	0.2	0.5	0.05	0.0072	0.0055	0.0017	
10	60	0.5	0.3	0.05	0.25	0.1	0.2	0.1	0.4	0.1	0.2	0.5	0.05	0.0058	0.0041	0.0016	
10	70	0.5	0.3	0.05	0.25	0.1	0.2	0.1	0.4	0.1	0.2	0.5	0.05	0.0049	0.0041	0.0008	
10	80	0.5	0.3	0.05	0.25	0.1	0.2	0.1	0.4	0.1	0.2	0.5	0.05	0.0042	0.0030	0.0013	
10	90	0.5	0.3	0.05	0.25	0.1	0.2	0.1	0.4	0.1	0.2	0.5	0.05	0.0037	0.0031	0.0007	
10	100	0.5	0.3	0.05	0.25	0.1	0.2	0.1	0.4	0.1	0.2	0.5	0.05	0.0034	0.0026	0.0008	
10	110	0.5	0.3	0.05	0.25	0.1	0.2	0.1	0.4	0.1	0.2	0.5	0.05	0.0030	0.0023	0.0007	
10	120	0.5	0.3	0.05	0.25	0.1	0.2	0.1	0.4	0.1	0.2	0.5	0.05	0.0027	0.0020	0.0007	
10	130	0.5	0.3	0.05	0.25	0.1	0.2	0.1	0.4	0.1	0.2	0.5	0.05	0.0025	0.0019	0.0006	
10	140	0.5	0.3	0.05	0.25	0.1	0.2	0.1	0.4	0.1	0.2	0.5	0.05	0.0023	0.0018	0.0005	
10	150	0.5	0.3	0.05	0.25	0.1	0.2	0.1	0.4	0.1	0.2	0.5	0.05	0.0022	0.0017	0.0004	
10	160	0.5	0.3	0.05	0.25	0.1	0.2	0.1	0.4	0.1	0.2	0.5	0.05	0.0020	0.0016	0.0004	
10	170	0.5	0.3	0.05	0.25	0.1	0.2	0.1	0.4	0.1	0.2	0.5	0.05	0.0019	0.0014	0.0005	
10	180	0.5	0.3	0.05	0.25	0.1	0.2	0.1	0.4	0.1	0.2	0.5	0.05	0.0018	0.0014	0.0004	
10	190	0.5	0.3	0.05	0.25	0.1	0.2	0.1	0.4	0.1	0.2	0.5	0.05	0.0017	0.0015	0.0002	
10	200	0.5	0.3	0.05	0.25	0.1	0.2	0.1	0.4	0.1	0.2	0.5	0.05	0.0016	0.0013	0.0003	
(c) Average absolute bias in sampling variability for individual-level mediator in Condition 3.																	
10	20	0.8	0.5	0.2	0.35	0.4	0.1	0.2	0.2	0.4	0.1	0.7	0.5	0.1935	0.1266	0.0669	0.0068
10	30	0.8	0.5	0.2	0.35	0.4	0.1	0.2	0.2	0.4	0.1	0.7	0.5	0.1113	0.0850	0.0263	
10	40	0.8	0.5	0.2	0.35	0.4	0.1	0.2	0.2	0.4	0.1	0.7	0.5	0.0775	0.0612	0.0162	
10	50	0.8	0.5	0.2	0.35	0.4	0.1	0.2	0.2	0.4	0.1	0.7	0.5	0.0586	0.0482	0.0104	
10	60	0.8	0.5	0.2	0.35	0.4	0.1	0.2	0.2	0.4	0.1	0.7	0.5	0.0476	0.0398	0.0078	
10	70	0.8	0.5	0.2	0.35	0.4	0.1	0.2	0.2	0.4	0.1	0.7	0.5	0.0400	0.0352	0.0048	
10	80	0.8	0.5	0.2	0.35	0.4	0.1	0.2	0.2	0.4	0.1	0.7	0.5	0.0343	0.0300	0.0043	
10	90	0.8	0.5	0.2	0.35	0.4	0.1	0.2	0.2	0.4	0.1	0.7	0.5	0.0303	0.0262	0.0041	
10	100	0.8	0.5	0.2	0.35	0.4	0.1	0.2	0.2	0.4	0.1	0.7	0.5	0.0273	0.0240	0.0032	
10	110	0.8	0.5	0.2	0.35	0.4	0.1	0.2	0.2	0.4	0.1	0.7	0.5	0.0243	0.0218	0.0025	
10	120	0.8	0.5	0.2	0.35	0.4	0.1	0.2	0.2	0.4	0.1	0.7	0.5	0.0222	0.0199	0.0024	
10	130	0.8	0.5	0.2	0.35	0.4	0.1	0.2	0.2	0.4	0.1	0.7	0.5	0.0203	0.0182	0.0021	
10	140	0.8	0.5	0.2	0.35	0.4	0.1	0.2	0.2	0.4	0.1	0.7	0.5	0.0188	0.0170	0.0018	
10	150	0.8	0.5	0.2	0.35	0.4	0.1	0.2	0.2	0.4	0.1	0.7	0.5	0.0173	0.0157	0.0016	
10	160	0.8	0.5	0.2	0.35	0.4	0.1	0.2	0.2	0.4	0.1	0.7	0.5	0.0163	0.0148	0.0015	
10	170	0.8	0.5	0.2	0.35	0.4	0.1	0.2	0.2	0.4	0.1	0.7	0.5	0.0152	0.0138	0.0014	
10	180	0.8	0.5	0.2	0.35	0.4	0.1	0.2	0.2	0.4	0.1	0.7	0.5	0.0144	0.0130	0.0014	
10	190	0.8	0.5	0.2	0.35	0.4	0.1	0.2	0.2	0.4	0.1	0.7	0.5	0.0136	0.0125	0.0011	
10	200	0.8	0.5	0.2	0.35	0.4	0.1	0.2	0.2	0.4	0.1	0.7	0.5	0.0128	0.0118	0.0010	

(*continued*)

Table B2. Continued.

Sample size		Parameters												Mediation error variance			
n1	n2	a	B	b1	c	λ_1	λ_2	λ_3	Λ_1	Λ_2	Λ_3	P	ICC	Empirical Variance	Predicted Variance	Absolute Bias	Average Absolute Bias
(d) Average absolute bias in sampling variability for individual-level mediator in Condition 4.																	
20	20	0.3	0.1	0.05	0.15	0.4	0.2	0.1	0.2	0.1	0.2	0.3	0.05	0.0086	0.0044	0.0043	0.0005
20	30	0.3	0.1	0.05	0.15	0.4	0.2	0.1	0.2	0.1	0.2	0.3	0.05	0.0047	0.0026	0.0021	
20	40	0.3	0.1	0.05	0.15	0.4	0.2	0.1	0.2	0.1	0.2	0.3	0.05	0.0032	0.0021	0.0010	
20	50	0.3	0.1	0.05	0.15	0.4	0.2	0.1	0.2	0.1	0.2	0.3	0.05	0.0024	0.0019	0.0005	
20	60	0.3	0.1	0.05	0.15	0.4	0.2	0.1	0.2	0.1	0.2	0.3	0.05	0.0019	0.0015	0.0004	
20	70	0.3	0.1	0.05	0.15	0.4	0.2	0.1	0.2	0.1	0.2	0.3	0.05	0.0016	0.0011	0.0005	
20	80	0.3	0.1	0.05	0.15	0.4	0.2	0.1	0.2	0.1	0.2	0.3	0.05	0.0014	0.0010	0.0004	
20	90	0.3	0.1	0.05	0.15	0.4	0.2	0.1	0.2	0.1	0.2	0.3	0.05	0.0012	0.0009	0.0003	
20	100	0.3	0.1	0.05	0.15	0.4	0.2	0.1	0.2	0.1	0.2	0.3	0.05	0.0011	0.0007	0.0003	
20	110	0.3	0.1	0.05	0.15	0.4	0.2	0.1	0.2	0.1	0.2	0.3	0.05	0.0010	0.0008	0.0002	
20	120	0.3	0.1	0.05	0.15	0.4	0.2	0.1	0.2	0.1	0.2	0.3	0.05	0.0009	0.0007	0.0001	
20	130	0.3	0.1	0.05	0.15	0.4	0.2	0.1	0.2	0.1	0.2	0.3	0.05	0.0008	0.0006	0.0002	
20	140	0.3	0.1	0.05	0.15	0.4	0.2	0.1	0.2	0.1	0.2	0.3	0.05	0.0007	0.0006	0.0002	
20	150	0.3	0.1	0.05	0.15	0.4	0.2	0.1	0.2	0.1	0.2	0.3	0.05	0.0007	0.0006	0.0001	
20	160	0.3	0.1	0.05	0.15	0.4	0.2	0.1	0.2	0.1	0.2	0.3	0.05	0.0006	0.0006	0.0001	
20	170	0.3	0.1	0.05	0.15	0.4	0.2	0.1	0.2	0.1	0.2	0.3	0.05	0.0006	0.0005	0.0001	
20	180	0.3	0.1	0.05	0.15	0.4	0.2	0.1	0.2	0.1	0.2	0.3	0.05	0.0006	0.0005	0.0001	
20	190	0.3	0.1	0.05	0.15	0.4	0.2	0.1	0.2	0.1	0.2	0.3	0.05	0.0005	0.0005	0.0001	
20	200	0.3	0.1	0.05	0.15	0.4	0.2	0.1	0.2	0.1	0.2	0.3	0.05	0.0005	0.0004	0.0001	
(e) Average absolute bias in sampling variability for individual-level mediator in Condition 5.																	
20	20	0.3	0.5	0.2	0.25	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0498	0.0218	0.0280	0.0025
20	30	0.3	0.5	0.2	0.25	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0270	0.0195	0.0075	
20	40	0.3	0.5	0.2	0.25	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0184	0.0136	0.0048	
20	50	0.3	0.5	0.2	0.25	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0139	0.0100	0.0039	
20	60	0.3	0.5	0.2	0.25	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0111	0.0088	0.0023	
20	70	0.3	0.5	0.2	0.25	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0093	0.0072	0.0021	
20	80	0.3	0.5	0.2	0.25	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0081	0.0066	0.0015	
20	90	0.3	0.5	0.2	0.25	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0070	0.0054	0.0016	
20	100	0.3	0.5	0.2	0.25	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0063	0.0053	0.0010	
20	110	0.3	0.5	0.2	0.25	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0057	0.0045	0.0011	
20	120	0.3	0.5	0.2	0.25	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0051	0.0043	0.0008	
20	130	0.3	0.5	0.2	0.25	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0047	0.0038	0.0009	
20	140	0.3	0.5	0.2	0.25	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0043	0.0037	0.0006	
20	150	0.3	0.5	0.2	0.25	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0040	0.0036	0.0005	
20	160	0.3	0.5	0.2	0.25	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0038	0.0032	0.0005	
20	170	0.3	0.5	0.2	0.25	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0035	0.0030	0.0005	
20	180	0.3	0.5	0.2	0.25	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0033	0.0028	0.0005	
20	190	0.3	0.5	0.2	0.25	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0031	0.0027	0.0004	
20	200	0.3	0.5	0.2	0.25	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.0030	0.0026	0.0004	
(f) Average absolute bias in sampling variability for individual-level mediator in Condition 6.																	
20	20	0.8	0.3	0.1	0.35	0.2	0.1	0.1	0.1	0.4	0.2	0.5	0.5	0.1101	0.0624	0.0477	0.0044
20	30	0.8	0.3	0.1	0.35	0.2	0.1	0.1	0.1	0.4	0.2	0.5	0.5	0.0579	0.0375	0.0205	
20	40	0.8	0.3	0.1	0.35	0.2	0.1	0.1	0.1	0.4	0.2	0.5	0.5	0.0384	0.0290	0.0094	
20	50	0.8	0.3	0.1	0.35	0.2	0.1	0.1	0.1	0.4	0.2	0.5	0.5	0.0287	0.0221	0.0066	
20	60	0.8	0.3	0.1	0.35	0.2	0.1	0.1	0.1	0.4	0.2	0.5	0.5	0.0229	0.0185	0.0045	
20	70	0.8	0.3	0.1	0.35	0.2	0.1	0.1	0.1	0.4	0.2	0.5	0.5	0.0195	0.0163	0.0032	
20	80	0.8	0.3	0.1	0.35	0.2	0.1	0.1	0.1	0.4	0.2	0.5	0.5	0.0166	0.0140	0.0026	
20	90	0.8	0.3	0.1	0.35	0.2	0.1	0.1	0.1	0.4	0.2	0.5	0.5	0.0144	0.0123	0.0021	
20	100	0.8	0.3	0.1	0.35	0.2	0.1	0.1	0.1	0.4	0.2	0.5	0.5	0.0127	0.0111	0.0016	
20	110	0.8	0.3	0.1	0.35	0.2	0.1	0.1	0.1	0.4	0.2	0.5	0.5	0.0116	0.0100	0.0015	
20	120	0.8	0.3	0.1	0.35	0.2	0.1	0.1	0.1	0.4	0.2	0.5	0.5	0.0105	0.0094	0.0011	
20	130	0.8	0.3	0.1	0.35	0.2	0.1	0.1	0.1	0.4	0.2	0.5	0.5	0.0096	0.0086	0.0010	
20	140	0.8	0.3	0.1	0.35	0.2	0.1	0.1	0.1	0.4	0.2	0.5	0.5	0.0088	0.0079	0.0009	
20	150	0.8	0.3	0.1	0.35	0.2	0.1	0.1	0.1	0.4	0.2	0.5	0.5	0.0082	0.0074	0.0008	
20	160	0.8	0.3	0.1	0.35	0.2	0.1	0.1	0.1	0.4	0.2	0.5	0.5	0.0077	0.0070	0.0007	
20	170	0.8	0.3	0.1	0.35	0.2	0.1	0.1	0.1	0.4	0.2	0.5	0.5	0.0072	0.0065	0.0007	
20	180	0.8	0.3	0.1	0.35	0.2	0.1	0.1	0.1	0.4	0.2	0.5	0.5	0.0067	0.0062	0.0005	
20	190	0.8	0.3	0.1	0.35	0.2	0.1	0.1	0.1	0.4	0.2	0.5	0.5	0.0063	0.0058	0.0005	
20	200	0.8	0.3	0.1	0.35	0.2	0.1	0.1	0.1	0.4	0.2	0.5	0.5	0.0060	0.0056	0.0005	

Sample size		Parameters													Power	
n1	n2	a	b	c	λ_1	λ_2	λ_3	Λ_1	Λ_2	Λ_3	P	ICC	Empirical R^2_{YL2}	Predicted R^2_{YL2}	Empirical rejection rate	Predicted power
(a) Empirical versus predicted statistical power for cluster-level mediators with quadratic function of the running variable.																
10	30	0.1	0.1	0.25	0.1	0.2	0	0.4	0.1	0	0.7	0.2	0.586	0.600	0.005	0.003
10	40	0.3	0.3	0.35	0.4	0.1	0	0.2	0.4	0	0.3	0.15	0.865	0.898	0.134	0.095
10	50	0.3	0.1	0.15	0.2	0.4	0	0.1	0.2	0	0.3	0.2	0.584	0.579	0.025	0.029
10	80	0.5	0.5	0.15	0.2	0.4	0	0.1	0.2	0	0.5	0.05	0.961	0.949	0.630	0.606
10	100	0.5	0.3	0.25	0.1	0.2	0	0.4	0.1	0	0.5	0.05	0.916	0.951	0.705	0.702
10	300	0.8	0.5	0.35	0.4	0.1	0	0.2	0.4	0	0.7	0.5	0.656	0.687	0.994	0.994
20	30	0.8	0.3	0.25	0.2	0.4	0	0.1	0.2	0	0.5	0.15	0.797	0.781	0.511	0.505
20	40	0.8	0.5	0.35	0.1	0.1	0	0.2	0.2	0	0.3	0.3	0.745	0.730	0.657	0.638
20	50	0.3	0.5	0.25	0.1	0.4	0	0.4	0.2	0	0.7	0.2	0.858	0.889	0.203	0.157
20	80	0.1	0.1	0.15	0.4	0.2	0	0.4	0.1	0	0.7	0.05	0.875	0.870	0.043	0.032
20	100	0.8	0.3	0.35	0.2	0.1	0	0.1	0.4	0	0.5	0.5	0.586	0.597	0.719	0.749
20	300	0.3	0.1	0.15	0.4	0.2	0	0.2	0.1	0	0.3	0.05	0.781	0.780	0.602	0.587
50	30	0.5	0.5	0.35	0.4	0.2	0	0.2	0.4	0	0.3	0.2	0.879	0.875	0.302	0.225
50	40	0.3	0.3	0.25	0.1	0.4	0	0.1	0.2	0	0.5	0.15	0.795	0.781	0.213	0.153
50	50	0.5	0.3	0.35	0.4	0.2	0	0.2	0.1	0	0.3	0.2	0.698	0.690	0.349	0.313
50	80	0.8	0.1	0.15	0.2	0.1	0	0.4	0.1	0	0.7	0.2	0.610	0.608	0.227	0.233
50	100	0.5	0.5	0.25	0.1	0.4	0	0.1	0.4	0	0.5	0.05	0.974	0.963	0.715	0.717
50	300	0.3	0.1	0.15	0.2	0.1	0	0.4	0.2	0	0.7	0.5	0.444	0.445	0.143	0.146
(b) Empirical versus predicted statistical power for cluster-level mediators with cubic function of the running variable.																
10	30	0.1	0.1	0.25	0.1	0.2	0.1	0.4	0.1	0.2	0.7	0.2	0.876	0.903	0.005	0.002
10	40	0.3	0.3	0.35	0.4	0.1	0.2	0.2	0.4	0.1	0.3	0.15	0.916	0.980	0.152	0.126
10	50	0.3	0.1	0.15	0.2	0.4	0.2	0.1	0.2	0.1	0.3	0.2	0.742	0.776	0.016	0.023
10	80	0.5	0.5	0.15	0.2	0.4	0.4	0.1	0.2	0.4	0.5	0.05	0.995	1.000	0.555	0.638
10	100	0.5	0.3	0.25	0.1	0.2	0.1	0.4	0.1	0.2	0.5	0.05	0.981	0.992	0.650	0.649
10	300	0.8	0.5	0.35	0.4	0.1	0.2	0.2	0.4	0.1	0.7	0.5	0.787	0.832	1.000	1.000
20	30	0.8	0.3	0.25	0.2	0.4	0.2	0.1	0.2	0.1	0.5	0.15	0.873	0.901	0.487	0.510
20	40	0.8	0.5	0.35	0.1	0.1	0.4	0.2	0.2	0.2	0.3	0.3	0.926	0.949	0.605	0.676
20	50	0.3	0.5	0.25	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.926	0.941	0.171	0.140
20	80	0.1	0.1	0.15	0.4	0.2	0.1	0.4	0.1	0.4	0.7	0.05	0.991	1.000	0.045	0.044
20	100	0.8	0.3	0.35	0.2	0.1	0.1	0.1	0.4	0.2	0.5	0.5	0.729	0.754	0.761	0.770
20	300	0.3	0.1	0.15	0.4	0.2	0.1	0.2	0.1	0.2	0.3	0.05	0.971	0.978	0.589	0.581
50	30	0.5	0.5	0.35	0.4	0.2	0.4	0.2	0.4	0.2	0.3	0.2	0.962	0.984	0.306	0.274
50	40	0.3	0.3	0.25	0.1	0.4	0.2	0.1	0.2	0.4	0.5	0.15	0.965	0.998	0.198	0.156
50	50	0.5	0.3	0.35	0.4	0.2	0.1	0.2	0.1	0.4	0.3	0.2	0.954	0.980	0.345	0.327
50	80	0.8	0.1	0.15	0.2	0.1	0.1	0.4	0.1	0.1	0.7	0.2	0.810	0.816	0.262	0.281
50	100	0.5	0.5	0.25	0.1	0.4	0.2	0.1	0.4	0.2	0.5	0.05	0.987	1.000	0.674	0.691
50	300	0.3	0.1	0.15	0.2	0.1	0.4	0.4	0.2	0.1	0.7	0.5	0.657	0.665	0.148	0.224

Sample size		Parameters														Power	
n1	n2	a	B	b1	c	λ_1	λ_2	λ_3	Λ_1	Λ_2	Λ_3	P	ICC	Empirical R^2_{YL2}	Predicted R^2_{YL2}	Empirical Rejection Rate	Predicted Power
(a) Empirical versus predicted statistical power for individual-level mediators with quadratic function of the running variable.																	
10	30	0.1	0.1	0.1	0.25	0.1	0.2	0	0.4	0.1	0	0.7	0.2	0.559	0.577	0.003	0.001
10	40	0.3	0.3	0.2	0.35	0.4	0.1	0	0.2	0.4	0	0.3	0.15	0.836	0.882	0.114	0.124
10	50	0.3	0.1	0.1	0.15	0.2	0.4	0	0.1	0.2	0	0.3	0.2	0.559	0.566	0.019	0.030
10	80	0.5	0.5	0.05	0.15	0.2	0.4	0	0.1	0.2	0	0.5	0.05	0.936	0.933	0.999	0.996
10	100	0.5	0.3	0.05	0.25	0.1	0.2	0	0.4	0.1	0	0.5	0.05	0.873	0.936	0.887	0.912
10	300	0.8	0.5	0.2	0.35	0.4	0.1	0	0.2	0.4	0	0.7	0.5	0.659	0.693	0.990	0.990
20	30	0.8	0.3	0.05	0.25	0.2	0.4	0	0.1	0.2	0	0.5	0.15	0.745	0.747	0.356	0.394
20	40	0.8	0.5	0.2	0.35	0.1	0.1	0	0.2	0.2	0	0.3	0.3	0.681	0.678	0.752	0.783
20	50	0.3	0.5	0.2	0.25	0.1	0.4	0	0.4	0.2	0	0.7	0.2	0.807	0.865	0.356	0.351
20	80	0.1	0.1	0.1	0.15	0.4	0.2	0	0.4	0.1	0	0.7	0.05	0.862	0.865	0.004	0.014
20	100	0.8	0.3	0.1	0.35	0.2	0.1	0	0.1	0.4	0	0.5	0.5	0.587	0.600	0.722	0.752
20	300	0.3	0.1	0.05	0.15	0.4	0.2	0	0.2	0.1	0	0.3	0.05	0.743	0.742	0.367	0.414
50	30	0.5	0.5	0.2	0.35	0.4	0.2	0	0.2	0.4	0	0.3	0.2	0.845	0.860	0.408	0.435
50	40	0.3	0.3	0.1	0.25	0.1	0.4	0	0.1	0.2	0	0.5	0.15	0.733	0.730	0.217	0.227
50	50	0.5	0.3	0.1	0.35	0.4	0.2	0	0.2	0.1	0	0.3	0.2	0.613	0.612	0.423	0.429
50	80	0.8	0.1	0.05	0.15	0.2	0.1	0	0.4	0.1	0	0.7	0.2	0.588	0.585	0.136	0.137
50	100	0.5	0.5	0.05	0.25	0.1	0.4	0	0.1	0.4	0	0.5	0.05	0.963	0.963	1.000	0.997
50	300	0.3	0.1	0.2	0.15	0.2	0.1	0	0.4	0.2	0	0.7	0.5	0.443	0.445	0.140	0.135
(b) Empirical versus predicted statistical power for individual-level mediators with cubic function of the running variable.																	
10	30	0.1	0.1	0.1	0.25	0.1	0.2	0.1	0.4	0.1	0.2	0.7	0.2	0.875	0.880	0.002	0.002
10	40	0.3	0.3	0.2	0.35	0.4	0.1	0.2	0.2	0.4	0.1	0.3	0.15	0.906	0.995	0.092	0.228
10	50	0.3	0.1	0.1	0.15	0.2	0.4	0.2	0.1	0.2	0.1	0.3	0.2	0.730	0.766	0.008	0.030
10	80	0.5	0.5	0.05	0.15	0.2	0.4	0.1	0.1	0.2	0.1	0.5	0.1	0.930	0.965	0.991	0.994
10	100	0.5	0.3	0.05	0.25	0.1	0.2	0.1	0.4	0.1	0.2	0.5	0.05	0.978	0.995	0.926	0.933
10	300	0.8	0.5	0.2	0.35	0.4	0.1	0.2	0.2	0.4	0.1	0.7	0.5	0.785	0.834	0.999	1.000
20	30	0.8	0.3	0.05	0.25	0.2	0.4	0.2	0.1	0.2	0.1	0.5	0.15	0.850	0.892	0.300	0.346
20	40	0.8	0.5	0.2	0.35	0.1	0.1	0.1	0.2	0.2	0.2	0.3	0.3	0.844	0.876	0.728	0.785
20	50	0.3	0.5	0.2	0.25	0.1	0.4	0.2	0.4	0.2	0.1	0.7	0.2	0.912	0.938	0.295	0.318
20	80	0.1	0.1	0.1	0.15	0.4	0.2	0.2	0.4	0.1	0.1	0.7	0.05	0.955	0.953	0.001	0.016
20	100	0.8	0.3	0.1	0.35	0.2	0.1	0.1	0.1	0.4	0.2	0.5	0.5	0.726	0.756	0.771	0.794
20	300	0.3	0.1	0.05	0.15	0.4	0.2	0.1	0.2	0.1	0.2	0.3	0.05	0.970	0.978	0.314	0.446
50	30	0.5	0.5	0.2	0.35	0.4	0.2	0.1	0.2	0.4	0.2	0.3	0.2	0.919	0.956	0.414	0.480
50	40	0.3	0.3	0.1	0.25	0.1	0.4	0.1	0.1	0.2	0.1	0.5	0.15	0.826	0.873	0.195	0.245
50	50	0.5	0.3	0.1	0.35	0.4	0.2	0.2	0.2	0.1	0.1	0.3	0.2	0.838	0.864	0.406	0.419
50	80	0.8	0.1	0.05	0.15	0.2	0.1	0.1	0.4	0.1	0.1	0.7	0.2	0.804	0.809	0.104	0.145
50	100	0.5	0.5	0.05	0.25	0.1	0.4	0.2	0.1	0.4	0.1	0.5	0.05	0.978	1.000	1.000	1.000
50	300	0.3	0.1	0.2	0.15	0.2	0.1	0.1	0.4	0.2	0.2	0.7	0.5	0.733	0.739	0.160	0.164