Current Psychology
https://doi.org/10.1007/s12144-021-02034-6

Equivalence of two least-squares estimators for indirect effects

WenWu Wang[1,2] · Ping Yu[3] · Yuejin Zhou[4] · Tiejun Tong[5] · Zhonghua Liu[2]

Accepted: 22 June 2021

Abstract
In social and behavioral sciences, the mediation test based on the indirect effect is an important topic. There are many methods to assess intervening variable effects. In this paper, we focus on the difference method and the product method in mediation models. Firstly, we analyze the regression functions in the simple mediation model, and provide an expectation-consistent condition. We further show that the difference estimator and the product estimator are numerically equivalent based on the least-squares regression regardless of the error distribution. Secondly, we generalize the equivalence result to the three-path model and the multiple mediators model, and prove a general equivalence result in a class of restricted linear mediation models. Thirdly, we investigate the empirical distributions of the indirect effect estimators in the simple mediation model by simulations, and show that the indirect effect estimators are normally distributed as long as one multiplicand of the product estimator is large. Finally, we introduce some popular R packages for mediation analysis and also provide some useful suggestions on how to correctly conduct mediation analysis.

Keywords Bootstrap · Difference in coefficients · Indirect effects · Least-squares regression · Mediation analysis · Product of coefficients

✉ WenWu Wang
wangwenwu@qfnu.edu.cn; wwwang@hku.hk

Ping Yu
pingyu@hku.hk

Yuejin Zhou
yjchow@163.com

Tiejun Tong
tongt@hkbu.edu.hk

Zhonghua Liu
zhhliu@hku.hk

1 School of Statistics, Qufu Normal University, Shandong, China

2 Department of Statistics and Actuarial Science, The University of Hong Kong, Pok Fu Lam, Hong Kong

3 HKU Business School, The University of Hong Kong, Pok Fu Lam, Hong Kong

4 School of Mathematics and Big Data, Anhui University of Science & Technology, Anhui, China

5 Department of Mathematics, Hong Kong Baptist University, Kowloon Tong, Hong Kong

Introduction

In many disciplines, the effect of an exposure on the outcome variable is often mediated by an intermediate variable. Mediation analysis aims to identify the direct effect of the predictor on the outcome and the indirect effect between the same predictor and outcome via the change in a mediator (MacKinnon, 2008). Since the seminal paper by Baron and Kenny (1986), mediation analysis has become one of the most popular statistical methods in social sciences. For basic information on mediation analysis, one may refer to the recent textbooks including, for example, Hayes (2013) and MacKinnon (2008), and VanderWeele (2015). Empirical applications of mediation analysis have dramatically expanded in sociology, psychology, epidemiology, and medicine (Lockhart et al., 2011; Newland et al., 2013; Ogden et al., 2010; Richiardi et al., 2013; Rucker et al., 2011). Meanwhile, modern scientific investigations require sophisticated models for conducting mediation analysis (Frölich & Huber, 2017; Lachowicz et al., 2018; VanderWeele & Tchetgen Tchetgen, 2017).

One important issue in mediation studies is to derive the statistical inference of the mediated effects, for which three main approaches are available in the literature. The first approach is the causal steps approach (Baron & Kenny, 1986), which specifies a series of tests of links in a causal chain. Some variants of this method that test three different hypotheses have been proposed (Allison, 1995; Kenny et al., 1998). The second approach is the difference in coefficients approach (McGuigan & Langholtz, 1988), which takes the difference between a regression coefficient before and after being adjusted by the intervening variable. The third approach is the product of coefficients approach which involves paths in a path model (MacKinnon et al., 1998; MacKinnon & Lockwood, 2001; Sobel, 1982). To evaluate their performance, MacKinnon et al. (2002) gave a summary and comparison of these existing methods. For more simulation comparison, see, for example, MacKinnon et al. (2004), Preacher and Hayes (2008), and Preacher and Selig (2012).

In this paper, we focus on the total indirect effect based on the least-squares regression. Firstly, we review the simple mediation model and some basic inference methods, provide an expectation-consistent condition for the model, and prove the equivalence between the difference and product estimators using the closed-form expressions. Meanwhile, we introduce some popular R packages for mediation analysis and also provide some useful suggestions on how to correctly conduct mediation analysis. Secondly, we prove the equivalence between the difference and product estimators in the three-path model (Taylor et al., 2008) and in the multiple mediators model (Daniel et al., 2015; Taguri et al., 2018). Thirdly, we prove a general result on the numerical equivalence between the two estimators in a general linear mediation model with restriction. Fourthly, we report some empirical distributions of the indirect effect estimators by simulations, and point out some limitations of the existing inference methods by analyzing real data on DNA methylation. Finally, we also call for the preregistration of all mediation analyses prior to data collection where the model specification should take into account practical considerations, e.g., the mediation factors must play their roles (in time) after the treatment is conducted, or some factors should be included based on expert suggestions. Meanwhile, we note that the practice of searching for significant indirect effects in the lack of significant direct effects is essentially the same as a data snooping problem; see, e.g., White (2000) where the uncertainty of model searching is explicitly incorporated.

We emphasize that the main result of this paper is Theorem 6 in Section "General Linear Mediation Models". To better understand this result, we first use a simple mediation model and two more complex linear models to illustrate this result, and then state the general result in Theorem 6.

Simple Mediation Model

The simple mediation model is given in Fig. 1, where X is the independent variable, Y is the dependent variable, and M is the mediating variable that mediates the effects of X on Y. Given the observations (X_i, M_i, Y_i) for $i = 1, \ldots, n$, the simple mediation model consists of three regression equations:

$$Y_i = \beta_0 + cX_i + \epsilon_{0,i}, \tag{1}$$
$$M_i = \beta_1 + aX_i + \epsilon_{1,i}, \tag{2}$$
$$Y_i = \beta_2 + c'X_i + bM_i + \epsilon_{2,i}, \tag{3}$$

where c represents the total effect of X on Y, a represents the relation between X and M, c' represents the direct effect of X on Y after adjusting the effect of M, and b represents the relation between M and Y after adjusting the effect of X.

For the simple mediation model, the mediated effect, also called the indirect effect, can be defined in two different forms: ab or $c - c'$. In general, the main goal of mediation analysis is to test whether the null hypothesis $H_0 : ab = 0$ or $H_0 : c - c' = 0$ is true. In this section, we compare the two forms of indirect effect in the least-squares regression framework.

Zero-Mean Error Condition for Model Consistency

Note that the regression Eqs. 1–3 are interrelated in the simple mediation model. We substitute Eq. 2 into Eq. 3 to obtain the following equation:

$$\begin{aligned} Y_i &= \beta_2 + c'X_i + b(\beta_1 + aX_i + \epsilon_{1,i}) + \epsilon_{2,i} \\ &= (\beta_2 + b\beta_1) + (c' + ab)X_i + \epsilon_i, \end{aligned} \tag{4}$$

where $\epsilon_i = b\epsilon_{1,i} + \epsilon_{2,i}$. Assume also that $\epsilon_{1,i}$ and $\epsilon_{2,i}$ are zero-mean distributed, where "zero-mean" indicates

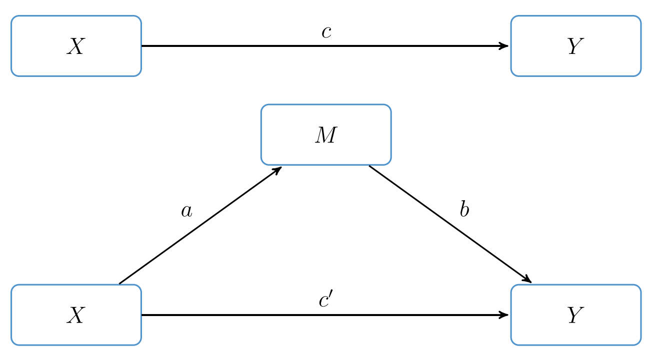

Fig. 1 Causal diagram of the simple mediation model, where X is the independent variable, M is the mediating variable (or mediator), and Y is the dependent variable. There is only one pathway through the mediator M ($X \to M \to Y$). For the regression parameters, c is the total effect of X on Y, c' is the direct effect of X on Y after adjusting the effect of M, a quantifies the relation between X and M, and b quantifies the relation between M and Y after adjusting the effect of X

their conditional means are zero, i.e., $E[\epsilon_{1,i}|X_i] = 0$ and $E[\epsilon_{2,i}|X_i] = 0$. Then consequently, ϵ_i is also zero-mean distributed by noting that $E[\epsilon_i|X_i] = bE[\epsilon_{1,i}|X_i] + E[\epsilon_{2,i}|X_i] = 0$. Further by Eqs. 1 and 4, we have

$$E[Y_i|X_i] = \beta_0 + cX_i,$$
$$E[Y_i|X_i] = (\beta_2 + b\beta_1) + (c' + ab)X_i.$$

This shows that $c = c' + ab$. The two expressions of $E[Y_i|X_i]$ also imply that ϵ_i in Eq. 4 is equal to $\epsilon_{0,i}$ in Eq. 1.

Theorem 1 *For the simple mediation model, assume that $\epsilon_{j,i}$ are zero-mean distributed with $E[\epsilon_{j,i}|X_i] = 0$ for $j = 1, 2$. Then we have the equality*

$$ab = c - c'.$$

In particular, if $\epsilon_{j,i} \overset{i.i.d.}{\sim} N\left(0, \sigma_j^2\right)$ for $j = 1, 2$, then they satisfy the zero-mean condition, where $i.i.d.$ is an abbreviation of "independent and identically distributed". And if we further assume that $\epsilon_{1,i}$ and $\epsilon_{2,i}$ are independent, then $\epsilon_{0,i}$ is also normally distributed with variance $\sigma_0^2 = b^2\sigma_1^2 + \sigma_2^2$.

Least-Squares Regression

The standard mediation analysis uses the least-squares regression to estimate the regression parameters. Specifically, by minimizing the sum of squared errors, where arg min represents obtaining the optimal values of arguments by minimizing the the sum of squared errors, we have

$$(\hat\beta_0, \hat c)^T = (\tilde X^T \tilde X)^{-1} \tilde X^T Y = \arg\min_{\beta_0, c} \sum_{i=1}^n (Y_i - \beta_0 - cX_i)^2, \quad (5)$$

$$(\hat\beta_1, \hat a)^T = (\tilde X^T \tilde X)^{-1} \tilde X^T M = \arg\min_{\beta_1, a} \sum_{i=1}^n (M_i - \beta_1 - aX_i)^2, \quad (6)$$

$$(\hat\beta_2, \hat c', \hat b)^T = (\check X^T \check X)^{-1} \check X^T Y = \arg\min_{\beta_2, c', b} \sum_{i=1}^n (Y_i - \beta_2 - c'X_i - bM_i)^2, \quad (7)$$

where $X = (X_1, \ldots, X_n)^T$, $M = (M_1, \ldots, M_n)^T$, $Y = (Y_1, \ldots, Y_n)^T$, $I = (1, \ldots, 1)^T$, $\tilde X = (I, X)$, and $\check X = (I, X, M)$. Moreover, if we assume that $\epsilon_{1,i} \overset{i.i.d.}{\sim} N\left(0, \sigma_1^2\right)$, $\epsilon_{2,i} \overset{i.i.d.}{\sim} N\left(0, \sigma_2^2\right)$, and they are independent. Then the least-squares estimators in Eqs. 5- 7 follow the normal distributions as

$$\hat c \sim N\left(c, \sigma_0^2 e_{2,2}^T (\tilde X^T \tilde X)^{-1} e_{2,2}\right),$$
$$\hat a \sim N\left(a, \sigma_1^2 e_{2,2}^T (\tilde X^T \tilde X)^{-1} e_{2,2}\right),$$
$$\hat c' \sim N\left(c', \sigma_2^2 e_{2,3}^T (\check X^T \check X)^{-1} e_{2,3}\right),$$
$$\hat b \sim N\left(b, \sigma_2^2 e_{3,3}^T (\check X^T \check X)^{-1} e_{3,3}\right),$$

where $e_{2,2} = (0, 1)^T$, $e_{2,3} = (0, 1, 0)^T$, $e_{3,3} = (0, 0, 1)^T$, and $\sigma_0^2 = b^2\sigma_1^2 + \sigma_2^2$.

The above results are straightforward and hence the proof is omitted. When the random errors are normally distributed, it is known that the least-squares estimator is the most efficient estimator, the minimum variance unbiased estimator, and also the maximum likelihood estimator. In principle, the normality of the errors is too strong for model consistency, and the least- squares estimator does not need the normality assumption, but only requires that the expectations of the errors are zero. As an example, the error distribution that is symmetric about zero satisfies the zero-mean condition.

Equivalence Between the Difference and Product Estimators

The indirect effect of X on Y can be estimated by two methods: the difference of the estimated coefficients $\hat c - \hat c'$, and the product of the estimated coefficients $\hat a\hat b$. In this subsection, we show that the two methods provide the same estimate in mediation analysis.

Theorem 2 *In the simple mediation model, the difference estimator is equivalent to the product estimator, i.e.*

$$\hat a\hat b = \hat c - \hat c',$$

regardless of the error distribution.

The proof of Theorem 2 is provided in Appendix A. This theorem shows that, no matter what the error distribution is, the two estimators for the indirect effect are exactly the same in the closed expression of the least-squares estimators. The equivalence of the two estimators is attributed to three facts: complete data, the linear equation, and a least-squares regression. If there are missing data, or if the model is multilevel or logistic, or if we apply the least absolute deviation or the other loss functions, then the equivalence between the two estimators will no longer hold, i.e., in the logistic regression model (Eshima et al., 2001). From the viewpoint of the collapsibility, Guo and Geng (1995) and Guo et al. (2001) gave further discussions.

MacKinnon et al. (1995) provided the result in Theorem 1, by explicitly deriving the formulas of ab and $c - c'$; we provide an alternative proof without explicit derivation. MacKinnon et al. (1995) claimed the numerical equivalence of $\hat a\hat b$ and $\hat c - \hat c'$ by examining some samples, but no rigorous proof was provided; Theorem 2 fills this gap.

Inference

In mediation analysis, the main aim is to test whether the estimated indirect effect is significantly different from zero.

For the difference and product estimators, the test statistics can be constructed as

$$z_p = \frac{\hat{a}\hat{b}}{\hat{\sigma}_p} \tag{8}$$

and

$$z_d = \frac{\hat{c} - \hat{c}'}{\hat{\sigma}_d}, \tag{9}$$

where $\hat{\sigma}_p$ and $\hat{\sigma}_d$ are the standard errors of the two estimators, respectively. Since the two estimators are equivalent, their variances should satisfy $\sigma_p^2 = \sigma_d^2$, where σ^2 without a hat means the variance, and with a hat means an estimator of the variance. Under the assumption of normality, the variance estimation is a key step for inference. There are many works on the variance estimation.

For the product estimator, Sobel (1982) applied the multivariate delta method and proposed an approximate formula for the standard error as

$$\hat{\sigma}_p = \sqrt{\hat{a}^2\hat{\sigma}_b^2 + \hat{b}^2\hat{\sigma}_a^2},$$

For the difference estimator, McGuigan and Langholtz (1988) developed a method to study binary health measures and proposed to estimate the standard error of $\hat{c} - \hat{c'}$ by

$$\hat{\sigma}_d = \sqrt{\hat{\sigma}_c^2 + \hat{\sigma}_{c'}^2 - 2\hat{\rho}\hat{\sigma}_c\hat{\sigma}_{c'}},$$

where $\hat{\rho}$ is the sample correlation between c and c'. For more estimators of the standard error, one may refer to MacKinnon et al. (2002). Thus the $(1 - \alpha)$ confidence intervals of ab and $c - c'$ are

$$[\hat{a}\hat{b} - z_{1-\alpha/2}\hat{\sigma}_p, \hat{a}\hat{b} + z_{1-\alpha/2}\hat{\sigma}_p], \tag{10}$$

$$[\hat{c} - \hat{c}' - z_{1-\alpha/2}\hat{\sigma}_d, \hat{c} - \hat{c}' + z_{1-\alpha/2}\hat{\sigma}_d], \tag{11}$$

respectively, where α is the specified significance level, and $z_{1-\alpha/2}$ is the $(1 - \alpha/2)$ quantile of the standard normal distribution. Usually, the variance estimators are numerically different, and so are the inference results, even though the same data and the same estimate of the indirect effect are given. This motivates us to derive rigorous statistical theory for the inference of the indirect effect.

Remark 1 Although the estimators $\hat{a}$ and $\hat{b}$ are normally distributed, the product $\hat{a}\hat{b}$ is not normally distributed, no matter whether or not the two estimators are independent (Cui et al., 2016; Nadarajah & Pogány, 2016). Due to the equivalence of the two estimators, the distribution of the difference between $\hat{c}$ and $\hat{c}'$ is not normally distributed either, which implies that $\hat{c}$ and $\hat{c}'$ are not jointly normally distributed.

Since the sampling distributions of z_p and z_d are not normally distributed, but are skewed and leptokurtic in most cases (Kisbu-Sakarya et al., 2014; MacKinnon et al., 2002; MacKinnon et al., 2004; Preacher & Hayes, 2004; 2008), the tests based on the statistics Eqs. 8 and 9 have low powers and are often criticized in the literature. Some remedies are proposed by MacKinnon et al. (1998), MacKinnon and Lockwood (2001), and Shrout and Bolger (2002). In the literature, there are many methods for inference of the indirect effect: the method based on the correct distribution of the product, the resampling method, the Monte Carlo method, the method based on Bayesian credible intervals, and the joint significance test. The distribution of the product strategy explores the correct distribution of $\hat{a}\hat{b}$ rather than assumes its normality. The distribution function of the product of two standardized normal variables are presented in Meeker et al. (1981). Preacher and Selig (2012) indicated:"This method performs well in simulation studies, but until recently required recourse to tables with limited availability and knowledge of the population values of either a or b". The bootstrap is a popular resampling method to conduct inference (MacKinnon et al., 2004; Preacher & Hayes, 2008). To improve the finite-sample performance, other bias-corrected and bias-adjusted versions are also provided (Hayes, 2013; MacKinnon, 2008; Preacher & Selig, 2012). The Monte Carlo simulation is an alternative method to the bootstrap, which directly generates sample statistics from their joint distribution, not resampling the original data (MacKinnon et al., 2004; Preacher & Selig, 2012). Drawbacks to the sampling method include slight inconsistency among replications of the same experiment with the same data due to resampling variability and no theoretical results to guarantee their asymptotic consistency.

R Packages for Mediation Analysis

In total, there are more than thirty R packages developed for all types of mediation models. Specifically for the basic mediation model, the R packages include: *JSmediation*, *medflex*, *mediation*, *powerMediation* and *RMediation*. In *JSmediation*, Batailler et al. (2020) suggested reporting and testing component paths. In *medflex*, Steen et al. (2020) suggested running flexible mediation analysis in presence of nonlinear relations. In *mediation*, Tingley et al. (2019) implemented parametric and nonparametric mediation analysis. In *powerMediation*, Qiu (2020) calculated the minimal detectable slope for a mediator given sample size and power, and calculated the power for testing mediation effect. Finally in *RMediation*, Tofighi and MacKinnon (2016) computed confidence intervals for a nonlinear function of the model parameters. In addition, we note that there are also R packages for high-dimensional data analysis, likelihood analysis, correlated-error analysis, and nonparametric analysis.

In practical applications, however, little attention has been paid to the application scope of these mediation models and the corresponding statistical theory. Beyond the qualified scope of these models, they are very likely to yield invalid inference or even wrong results. This motivates us to propose three useful suggestions for practitioners. First, one should correctly understand the mediation models and the corresponding statistical theory; second, one should choose the suitable mediation model and R packages for the real data; third, one should interpret the results obtained from the mediation model via software, and explain the reasons why these results are sound.

Beyond Simple Mediation Model

In the previous section, we have focused on the simple mediation model with only one mediator, in which the only mediator transmits the influence of the independent variable to the dependent variable. In applications, the mediation chain with more than two paths or one mediator is also popular (Allen & Griffeth, 2001; Kim & Cicchetti, 2010; Nübold et al., 2015; Tein et al., 2000; Tekleab et al., 2005). In this section, we consider two such mediation models: the three-path mediation model (Taylor et al., 2008) and the multiple mediators model (Preacher & Hayes, 2008), and prove the equivalence of the two least-squares estimators for indirect mediation effects.

Three-Path Mediation Model

In a three-path mediation (or serial) model, two mediators M_1 and M_2 intervene in a series between an independent variable and a dependent variable (Taylor et al., 2008), which is depicted as a path diagram in Fig. 2. It consists of four regression equations:

$$
\begin{aligned}
Y_i &= \beta_0 + cX_i + \epsilon_{0,i}, && (12)\\
M_{1,i} &= \beta_1 + a_1 X_i + \epsilon_{1,i}, && (13)\\
M_{2,i} &= \beta_2 + a_2 X_i + dM_{1,i} + \epsilon_{2,i}, && (14)\\
Y_i &= \beta_3 + c' X_i + b_1 M_{1,i} + b_2 M_{2,i} + \epsilon_{3,i}, && (15)
\end{aligned}
$$

where the coefficients can be similarly interpreted as in the simple mediation model.

The total indirect effect, the effect passing through all paths, is defined as the sum of the products of the coefficients:

$$a_1 b_1 + a_2 b_2 + a_1 d b_2.$$

Taylor et al. (2008) indicated that:"Although it may be possible to develop a three-path test of mediation based on differences in coefficients, this method would likely be cumbersome in comparison to the product-of-coefficients test." As a result, the difference method is not considered in Taylor et al. (2008). In this subsection, we consider the indirect effect based on both the product and difference methods.

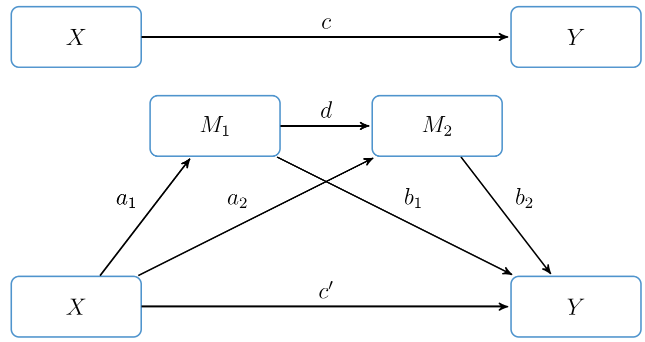

Fig. 2 Causal diagram of three-path (or serial) mediation model, where X is the independent variable, M_1 and M_2 are two mediators, and Y is the dependent variable. There are two pathways through one mediator ($X \to M_1 \to Y$; $X \to M_2 \to Y$), and one pathway through two mediators ($X \to M_1 \to M_2 \to Y$)

Following the discussion in Section "Simple Mediation Model", we have a similar equivalence result.

Theorem 3 *In the three-path mediation model, assume that $\epsilon_{j,i}$ for $j = 1, 2, 3$ are zero-mean distributed. Then we have*

(1) the parameters of the regression model satisfy the equality

$$c - c' = a_1 b_1 + a_2 b_2 + a_1 d b_2;$$

(2) the least-squares estimates of parameters satisfy the equality

$$\hat{c} - \hat{c}' = \hat{a}_1 \hat{b}_1 + \hat{a}_2 \hat{b}_2 + \hat{a}_1 \hat{d} \hat{b}_2.$$

The proof of Theorem 3 is provided in Appendix B.

Multiple Mediators Model

In this subsection, we consider the general mediation model with multiple mediators (Preacher & Hayes, 2008). For simplicity, we consider a model with only two mediators, with mediators M_1 and M_2, which can be expressed in the form of four regression equations (Fig. 3):

$$
\begin{aligned}
Y_i &= \beta_0 + cX_i + \epsilon_{0,i}, && (16)\\
M_{1,i} &= \beta_1 + a_1 X_i + \epsilon_{1,i}, && (17)\\
M_{2,i} &= \beta_2 + a_2 X_i + \epsilon_{2,i}, && (18)\\
Y_i &= \beta_3 + c' X_i + b_1 M_{1,i} + b_2 M_{2,i} + \epsilon_{3,i}. && (19)
\end{aligned}
$$

This form of the model is a special case of the three-path model with d in Eq. 14 equal to zero. The total indirect effect, the effect passing through either mediator, is defined as the sum of the products of the coefficients:

$$a_1 b_1 + a_2 b_2.$$

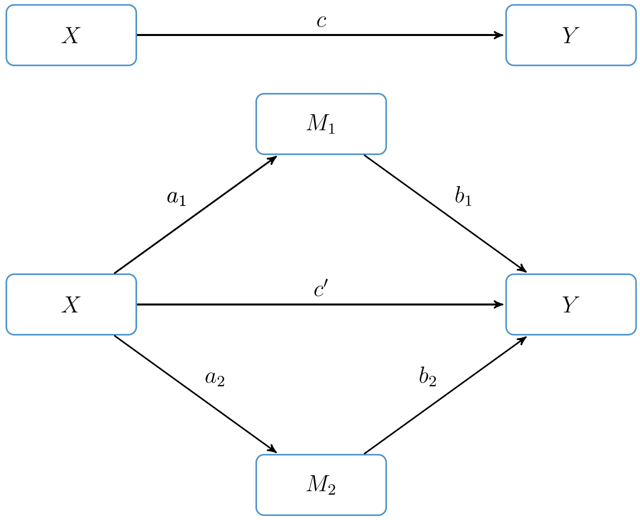

Fig. 3 Causal diagram of parallel mediation model with two mediators, where X is the independent variable, M_1 and M_2 are two mediators, and Y is the dependent variable. There are two pathways, each one through only one mediator ($X \rightarrow M_1 \rightarrow Y$; $X \rightarrow M_2 \rightarrow Y$)

A similar equivalence relationship between the product and difference estimators can be established.

Theorem 4 *In the two mediators model, assume that $\epsilon_{j,i}$ for $j = 1, 2, 3$ are zero-mean distributed. Then we have*

(1) the parameters of the regression model satisfy the equality

$$c - c' = a_1 b_1 + a_2 b_2;$$

(2) the least-squares estimates of parameters satisfy the equality

$$\hat{c} - \hat{c}' = \hat{a}_1 \hat{b}_1 + \hat{a}_2 \hat{b}_2.$$

The proof of Theorem 4 is simpler than that of Theorem 3, and thus is omitted.

We can extend Theorem 4 to the case with $k > 2$ mediators. Now, the model is expressed as

$$\begin{aligned} Y_i &= \beta_0 + cX_i + \epsilon_{0,i}, \\ M_{j,i} &= \beta_j + a_j X_i + \epsilon_{j,i}, \; j = 1, \cdots, k, \\ Y_i &= \beta_{k+1} + c' X_i + \sum\nolimits_{j=1}^{k} b_j M_{j,i} + \epsilon_{k+1,i}. \end{aligned}$$

Corollary 5 *In the multiple mediators model with $k > 2$ mediators, we have the estimation equality*

$$\hat{c} - \hat{c}' = \sum_{j=1}^{k} \hat{a}_j \hat{b}_j,$$

where $\hat{a}_j \hat{b}_j$ is the estimated indirect effect through mediator M_j.

This corollary is a special case of the general result in the next section.

General Linear Mediation Models

There are many linear mediation models with more than two mediators or more than two paths. A mediation graph consists of a set V of vertices and a set E of edges that connect some pairs of vertices (Pearl, 2009). The vertices in mediation graphs correspond to variables including the independent variable X, the dependent variable Y and the mediating variables M_j, and the edges denote a certain linear relationship between pairs of variables in linear mediation models. A path is defined as a sequence of edges (e.g. $((X, M_1), (M_1, M_2), (M_2, Y))$) that start from X and end at Y, and each edge starts with the vertex ending the preceding edge. In the general linear mediation models, we assume all paths start from X; in other words, M_j cannot start a path.

In this section, we consider the cases where each edge is directed, which means that each edge in a path is an arrow that points from the first to the second vertex of the pair. However, the mediation graph is restricted to be acyclic, i.e., contains no directed cycles (e.g., $X \rightarrow M$, $M \rightarrow X$) and no self-loops ($M \rightarrow M$). Now a specific group of linear regression equations is one-to-one to a mediation graph.

Based on the discussion in the previous sections, we provide a theorem on the equivalence between the difference and product estimators in the general linear mediation model.

Theorem 6 *In a linear mediation model, if*

(i) the mediation graph is acyclic;
(ii) the errors are zero-mean distributed;
(iii) each M_j equation contains X as a regressor;

then the least-squares estimates of parameters satisfy the equivalence relationship: the difference estimator equals the sum of the product of the estimated parameters in each path.

The assumption (iii) that each M_j equation contains X as a regressor cannot be dropped; see the simulation in the following Section "Empirical Distributions when X is Not a Regressor of an M Equation" for an illustration. The proof of Theorem 6 is provided in Appendix C.

Simulation Studies

In this section, we conduct simulations to illustrate the empirical distributions of the indirect effect estimators. The first subsection consider the simple mediation model and the second subsection considers the three-path mediation model in Fig. 2 where the edge from X to M_2 is deleted. We do not intend to evaluate test methods under the assumption of normal error distribution.

Empirical Distribution in a Simple Mediation Model

In this simulation study, the independent variable and error are generated from the standard normal distribution independently. The values of ab were chosen to be zero (0), medium (0.02, −0.02), and large (2), corresponding to the cases where (a, b) is equal to (0, 0)/ (0.2, 0.1) , (0.02, 1)/ (−0.2, 0.1), (−2, 0.01) and (2, 1), respectively. The sample size is 100, and the number of replications is set to be 10000 for each case.

Figure 4 shows the empirical distributions of the indirect effect estimators in red color. For comparison, we plot the corresponding normal distributions with the same mean and variance as the empirical distributions. None of the distributions are normal. However, some may approach or approximate normality as discussed in MacKinnon et al. (2007b) and Tofighi and MacKinnon (2011). For the zero indirect effect with $(a, b) = (0, 0)$, the empirical distribution is not normally distributed, because it has a sharper peak than the corresponding normal distribution. For the small indirect effects with $(a, b) = (0.2, 0.1)$ and $(-0.2, 0.1)$, the empirical distributions are skewed, right-skewed for the positive effect and left-skewed for the negative effect. For the small indirect effects with $(a, b) = (0.02, 1)$ and $(-2, 0.01)$, the empirical distributions are still normally distributed. For the large indirect effect with $(a, b) = (2, 1)$, the empirical distribution is also normally distributed. The empirical distribution is asymptotically normally distributed as long as one of the values of a and b is large in the simple mediation model.

Empirical Distributions when *X* is Not a Regressor of an *M* Equation

In the three-path mediation model of Fig. 2, suppose there is no edge from X to M_2. Other parameters are set as $\beta_1 = \beta_2 = \beta_3 = a_1 = b_1 = b_2 = c' = d = 1$ for simplicity, and the errors ϵ_1, ϵ_2 and ϵ_3 are independently normal distributed with the same variance 1. The sample sizes are set at $n = 100$ and 1000, and the number of replications is set to be 10000.

The upper two graphs in Fig. 5 show the histograms of the difference between the difference and product estimators when $n = 100$ and 1000. For comparison, we also report in the lower two graphs the histograms of the difference between the two estimators when the

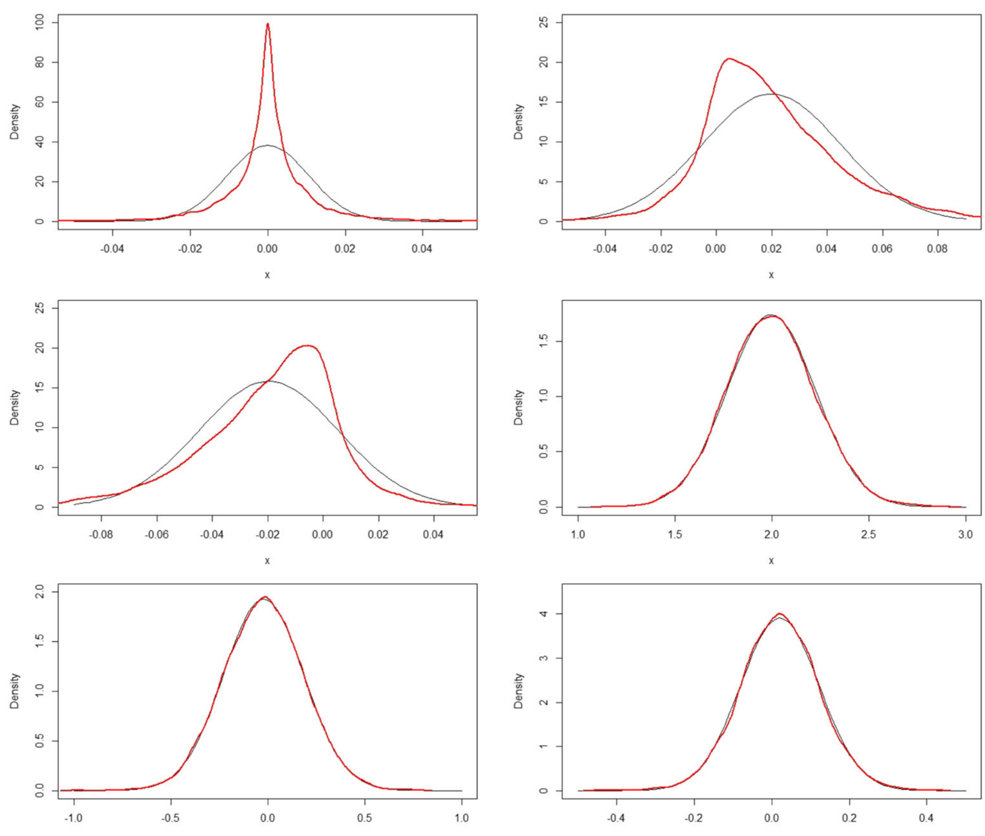

Fig. 4 The empirical distribution (red) vs the normal distribution (black) with the same mean and variance as the empirical distribution: the upper-left, upper-right, middle-left, middle-right, lower-left and lower-right figures correspond to the cases where (a, b) is equal to (0, 0), (0.2, 0.1), (−0.2, 0.1), (2, 1), (−2, 0.01) and (0.02, 1), respectively. It shows that the empirical distribution is asymptotically normal as long as either a or b is far away from zero

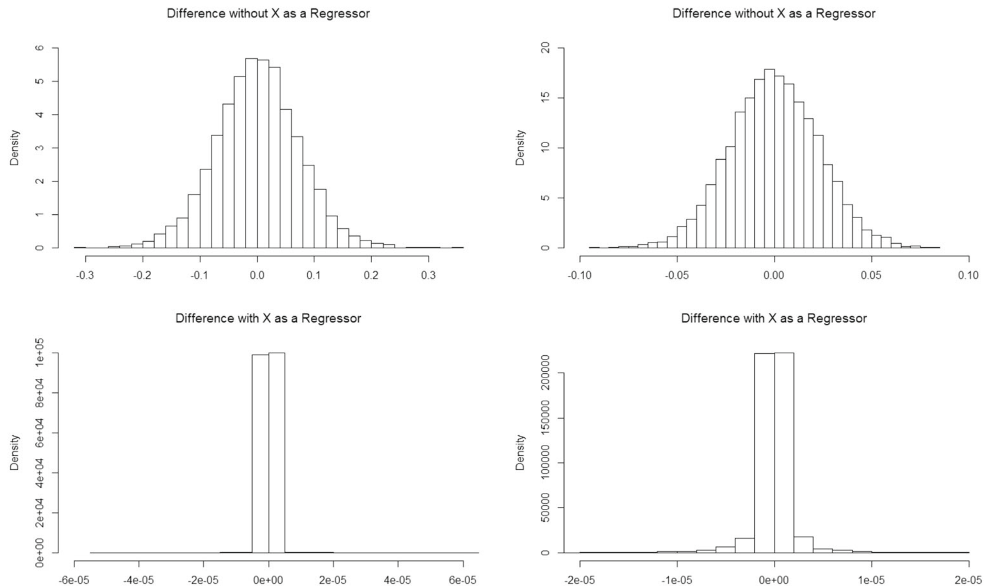

Fig. 5 Histograms of the difference between the difference and product estimators with and without X as a regressor: the left two graphs and right two graphs correspond to $n = 100$ and $n = 1000$, respectively; the upper two graphs and lower two graphs correspond to $a_2 = 0$ and $a_2 \neq 0$, respectively

edge from X to M_2 is added in and a_2 is set at zero. It shows that the two estimators with X as a regressor are numerically equivalent, while they are not without X as a regressor. Simulation results match the prediction of Theorem 6. Specifically, the mean and standard error for the difference between the two estimators are computed as $(-1.17 \times 10^{-7}, 5.29 \times 10^{-6})$, $(2.22 \times 10^{-3}, 7.24 \times 10^{-2})$, $(1.05 \times 10^{-9}, 3.00 \times 10^{-6})$ and $(1.45 \times 10^{-4}, 2.27 \times 10^{-2})$, which correspond to the cases: $n = 100$ without X, $n = 1000$ and without X, $n = 100$ and with X, $n = 1000$ and with X, respectively.

Real Data Analysis

In this section, we apply the least-squares methods to a real data set to estimate the indirect effects (IE) of socioeconomic status (SES) on body mass index (BMI) that might be mediated by DNA methylation CpG sites on chromosome 17, where SES is quantified by a scalar index ranging from 0 to 100, and BMI is a body mass index of an individual as in Loucks et al. (2016). To compare the two estimators, we choose three possible continuous mediators from DNA methylation: cg05157340, cg05156120 and cg05157970, take SES as the exposure X, and BMI as the outcome Y.

For the simple mediation analysis, the IEs are estimated using the least- squares method in Section "Simple Mediation Model", the standard errors (SE) are estimated by the formulas $\hat{\sigma}_p$ and $\hat{\sigma}_d$ for product and difference estimators, and the 95% confidence intervals (CI) based on variance estimation (which are denoted by CI_1) are constructed using formulas Eqs. 10 and 11, and the bootstrap CIs (which are denoted by CI_2) are constructed through 1000 bootstrap samples. Table 1 summarizes the estimated IE values, SEs and the two 95% CIs.

It can be seen that the difference and product estimators are numerically equivalent, while their SEs are different

Table 1 The indirect effects in the SES-BMI data: IE_1 and IE_2 are the difference and product estimators, CI_1 and CI_2 are the confidence intervals based on variance estimation and the bootstrap, respectively

	Estimate	SE	95% CI_1	95% CI_2
		cg05157340		
IE_1	–0.0056	0.0056	[-0.0166, 0.0054]	[-0.0337, 0.0220]
IE_2	–0.0056	0.0205	[-0.0457, 0.0345]	[-0.0381, 0.0222]
		cg05156120		
IE_1	–0.0227	0.0104	[-0.0373, -0.0080]	[-0.0465, 0.0106]
IE_2	–0.0227	0.0156	[-0.0532, 0.0078]	[-0.0462, 0.0110]
		cg05157970		
IE_1	–0.0922	0.0174	[-0.1264, -0.0580]	[-0.1431, -0.0265]
IE_2	–0.0922	0.0272	[-0.1455, -0.0388]	[-0.1410, -0.0303]

because different formulae are employed, which results in different CIs and sometimes even different inference conclusions, e.g., when the mediator is cg05156120, we can reject the null that the indirect effect is zero based on CI_1 of IE_1 but cannot reject based on CI_1 of IE_2. In summary, the same inference method (e.g., CI_1) applied to the two estimators can generate different inference conclusions, which makes the choice of estimator (beyond the choice of inference method) a topic deserving further exploration.

Discussions

In the literature, there are two estimation methods for the indirect mediated effect: the difference method and the product method. Most researchers recommend using the product form as the measure of the indirect effect, because it is in line with the causal interpretation of the mediation effect (MacKinnon et al., 2007a; Pearl, 2009; Yuan & MacKinnon, 2014).

In this paper, we provided an identification condition for expectation consistency in the simple mediation model, shown that the difference estimator is numerically equivalent to the product estimator in the least- squares regression, and summarized the statistical theories. One interesting finding is that the equivalence of the two estimators depends only on the least-squares estimation method, not on the error distribution. Furthermore, the equivalence can be generalized to the three-path mediation model and the multiple mediators model.

Since the two estimators are equivalent, they should have the same distribution. However, inference based on the two estimators may be different, and our real data analysis indicates this phenomenon. MacKinnon et al. (2002), MacKinnon et al. (2004), and Preacher and Hayes (2004) made extensive simulations based on normal errors to assess their Type I error rate and the power, and recommended to use the product estimator. These empirical results depend on the assumption of normal error distribution. As far as we know, there are no asymptotic results for either estimator, and thus the performance of empirical studies is not well understood. The mathematical expressions for the indirect effect estimators pave a way to develop the asymptotic theory. We are currently developing the asymptotic theory for all possible values of the model parameters, which would be helpful to make uniformly valid statistical inferences for the indirect effect.

In practical applications, violations of normality commonly encountered include heavy tails, skewness, outliers, contamination, and multimodality. Micceri (1989) examined 440 data sets from the psychological and educational literature, including 125 psychometric measures such as scales measuring personality, anxiety, and satisfaction. None of these data sets are normally distributed at the significance level of $\alpha = 0.01$; rather, the distributions were often heavy-tailed or skewed. For non-normal data, Yuan and MacKinnon (2014) proposed the least absolute deviation (LAD) estimator for the mediation effect; (Wang & Yu, 2020) further established the asymptotic theory for two LAD estimators based on the difference and product methods, and showed that they are not asymptotically equivalent; and Alfons (2020) provided an R package via a fast and robust bootstrap test. In order to further improve the estimation efficiency and the power for the test, we can apply the weighted quantile average regression (Zhao & Xiao, 2014), and the difference method by Wang et al. (2019) to analyze the mediation model. These works deserve further investigation.

Appendix A: Proof of Theorem 2

Proof We first consider the simple case where $\beta_0 = \beta_1 = \beta_2 = 0$. The least-squares estimators for the simplified models are

$$\hat{c} = \arg\min_{c} \sum_{i=1}^{n} (Y_i - cX_i)^2 = \frac{X^T Y}{X^T X},$$

$$\hat{a} = \arg\min_{a} \sum_{i=1}^{n} (M_i - aX_i)^2 = \frac{X^T M}{X^T X},$$

$$(\hat{c}', \hat{b})^T = \arg\min_{c',b} \sum_{i=1}^{n} (Y_i - c'X_i - bM_i)^2 = \frac{Y \begin{pmatrix} M^T M X^T - X^T M M^T \\ X^T X M^T - X^T M X^T \end{pmatrix}}{X^T X M^T M - X^T M X^T M}.$$

where $X = (X_1, \ldots, X_n)^T$, $M = (M_1, \ldots, M_n)^T$, and $Y = (Y_1, \ldots, Y_n)^T$. By the above least-squares estimators, the difference estimator is

$$\hat{c} - \hat{c}' = \frac{X^T M (X^T X M^T - X^T M X^T) Y}{X^T X (X^T X M^T M - X^T M X^T M)},$$

and the product estimator is

$$\hat{a}\hat{b} = \frac{X^T M (X^T X M^T - X^T M X^T) Y}{X^T X (X^T X M^T M - X^T M X^T M)}.$$

This shows that $\hat{c} - \hat{c}' = \hat{a}\hat{b}$. That is, the difference estimator is equivalent to the product estimator for the linear regression models with zero intercept.

The proof can readily be generalized to the models with non-zero intercept by replacing X and M by their demeaned couterparts and so is omitted. $\square$

Appendix B: Proof of Theorems 3

Proof By substituting Eqs. 13 and 14 into Eq. 15, it follows that

$$\begin{aligned}Y_i &= \beta_3 + c' X_i + b_1 M_{1,i} \\ &\quad + b_2(\beta_2 + d M_{1,i} + a_2 X_i + \epsilon_{2,i}) + \epsilon_{3,i} \\ &= (\beta_3 + b_2\beta_2) + (c' + a_2 b_2) X_i + (b_1 + b_2 d) M_{1,i} \\ &\quad + (b_2 \epsilon_{2,i} + \epsilon_{3,i}) \\ &= (\beta_3 + b_2\beta_2) + (c' + a_2 b_2) X_i \\ &\quad + (b_1 + b_2 d)(\beta_1 + a_1 X_i + \epsilon_{1,i}) + (b_2\epsilon_{2,i} + \epsilon_{3,i}) \\ &= (\beta_3 + b_2\beta_2 + (b_1 + b_2 d)\beta_1) \\ &\quad + (c' + a_2 b_2 + a_1 b_1 + a_1 d b_2) X_i + \epsilon_i, \end{aligned} \tag{20}$$

where $\epsilon_i = (b_1 + b_2 d)\epsilon_{1,i} + b_2\epsilon_{2,i} + \epsilon_{3,i}$. Since $\epsilon_{j,i}$ for $j = 1, 2, 3$ are zero-mean distributed, ϵ_i is also zero-mean distributed with $\mathrm{E}[\epsilon_i|X_i] = 0$. Taking expectation of Eqs. 12 and 20, we have

$$\begin{aligned}E[Y_i|X_i] &= \beta_0 + c X_i, \\ E[Y_i|X_i] &= (\beta_3 + b_2\beta_2 + (b_1 + b_2 d)\beta_1) \\ &\quad + (c' + a_1 b_1 + a_2 b_2 + a_1 d b_2) X_i.\end{aligned}$$

This leads to $\beta_0 = \beta_3 + b_2\beta_2 + (b_1 + b_2 d)\beta_1$ and $c - c' = a_1 b_1 + a_2 b_2 + a_1 d b_2$.

For the simplified models with $\beta_0 = \beta_1 = \beta_2 = \beta_3 = 0$, the least-squares estimators are

$$\begin{aligned}\tilde{c} &= \arg\min_c (Y_i - c X_i)^2 = \frac{A_7}{A_1}, \\ \tilde{a}_1 &= \arg\min_a (M_{1,i} - a_1 X_i)^2 = \frac{A_2}{A_1}, \\ (\tilde{a}_2, \tilde{d})^T &= \arg\min_{a_2,d} (M_{2,i} - a_2 X_i - d M_{1,i})^2 = \frac{\begin{pmatrix} A_3 A_4 - A_2 A_{10} \\ A_1 A_{10} - A_2 A_3 \end{pmatrix}}{A_1 A_4 - A_2 A_2}, \\ (\tilde{c}', \tilde{b}_1, \tilde{b}_2)^T &= \arg\min_{c',b_1,b_2} (Y_i - c' X_i - b_1 M_{1,i} - b_2 M_{2,i})^2 \\ &= \frac{\begin{pmatrix} (A_4 A_6 - A_5 A_5) A_7 + (A_3 A_5 - A_2 A_6) A_8 + (A_2 A_5 - A_3 A_4) A_9 \\ (A_3 A_5 - A_2 A_6) A_7 + (A_1 A_6 - A_3 A_3) A_8 + (A_2 A_3 - A_1 A_5) A_9 \\ (A_2 A_5 - A_3 A_4) A_7 + (A_2 A_3 - A_1 A_5) A_8 + (A_1 A_4 - A_2 A_2) A_9 \end{pmatrix}}{A_1(A_4 A_6 - A_5 A_5) + A_2(A_3 A_5 - A_2 A_6) + A_3(A_2 A_5 - A_3 A_4)}.\end{aligned}$$

where $A_1 = X^T X$, $A_2 = X^T M_1$, $A_3 = X^T M_2$, $A_4 = M_1^T M_1$, $A_5 = M_1^T M_2$, $A_6 = M_2^T M_2$, $A_7 = X^T Y$, $A_8 = M_1^T Y$, $A_9 = M_2^T Y$, and $A_{10} = M_1^T M_2$. By the above least-squares estimators, it is easy to verify that $\tilde{c} - \tilde{c}' = \tilde{a}_1\tilde{b}_1 + \tilde{a}_2\tilde{b}_2 + \tilde{a}_1\tilde{d}\tilde{b}_2$. That is, the difference estimator is equivalent to the product estimator for the linear regression models with zero intercept.

The proof can readily be generalized to the models with non-zero intercept and so is omitted. □

Appendix C: Proof of Theorem 6

Proof Suppose there are k mediators:

$$\begin{aligned}M_j &= \beta_j + R_j^T a_j + \epsilon_j, \quad j = 1, \ldots, k, \\ Y &= \beta_{k+1} + X c' + R_Y^T x b + \epsilon_{k+1},\end{aligned}$$

where R_j contains the non-constant regressors in the equation for M_j, which may include X and/or other M_j's, and R_Y contains M_j's that appear in the equation for Y. After substituting the equations for M_j, $j = 1, \ldots, k$, into the equation for Y, suppose we have

$$Y = \alpha_0 + (c' + \alpha_1) X + \epsilon_0 \equiv \alpha_0 + c X + \epsilon_0,$$

where (α_0, α_1) are functions of the coefficients in the equations for $\{M_j\}_{j=1}^k$ and Y, i.e.,

$$\begin{aligned}\alpha_0 &= f(\beta_1, \ldots, \beta_{k+1}; a_1, \ldots, a_k, b), \\ \alpha_1 &= f(a_1, \ldots, a_k, b),\end{aligned}$$

α_1 does not depend on $\beta_1, \ldots, \beta_{k+1}$ because it measures the sensitivity of Y to X while $\beta_1, \ldots, \beta_{k+1}$ does not contain such information, c' is the coefficient of X in the equation for Y, and ϵ_0 is a linear combination of the error terms in these $(k+1)$ equations, so it satisfies $\mathrm{E}[\epsilon_0|X] = 0$.

Because all the coefficients are estimated by least-squares regression, they employ the moment conditions

$$\mathrm{E}\left[\begin{pmatrix} 1 \\ R_j \end{pmatrix} (M_j - \beta_j - R_j^T a_j)\right] = 0,$$

$$\mathrm{E}\left[\begin{pmatrix} 1 \\ X \\ R_Y \end{pmatrix} (Y - \beta_{k+1} - X c' - R_Y^T b)\right] = 0.$$

If these moment conditions imply

$$\mathrm{E}\left[\begin{pmatrix}1\\X\end{pmatrix}(Y-\alpha_0-cX)\right]=0,$$

then our result follows since we just replace $\mathrm{E}[\cdot]$ by $1/n\sum_{i=1}^{n}$ in the least-squares estimation. However, this indeed holds because ϵ_0 is a linear function of $\{\epsilon_j\}_{j=1}^{k+1}$ so that

$$\mathrm{E}\left[\begin{pmatrix}1\\X\end{pmatrix}\epsilon_j\right]=0$$

implies

$$\mathrm{E}\left[\begin{pmatrix}1\\X\end{pmatrix}\epsilon_0\right]=0.$$

Here, note that X must be a regressor in the M_j equation, otherwise $\mathrm{E}\left[\begin{pmatrix}1\\X\end{pmatrix}\epsilon_j\right] = 0$ cannot hold such that $\mathrm{E}\left[\begin{pmatrix}1\\X\end{pmatrix}\epsilon_0\right] = 0$ cannot hold and the equivalence result fails. □

Acknowledgements Wang's work is supported by National Natural Science Foundation of China (No.12071248), and National Statistical Science Research Project of China (No.2020LZ26).

Declarations

Conflict of Interests The authors declare that they have no conflict of interest.

This is a theory paper with an application to DNA methylation open data. Our paper does not involve research on human participants or animals. Thus, we depended on those who collected the primary data about DNA methylation (Loucks et al., 2016), for compliance with the 1964 Declaration of Helsinki and its later addenda. The datasets generated during and/or analysed during the current study are available in an open software "JT-Comp", http://www.stat.sinica.edu.tw/ythuang/JT-Comp.zip.

The work has not been submitted elsewhere for publication, in whole or in part. All authors have seen the manuscript, approved to submit to Current Psychology, and consented to its review by Current Psychology.

References

Alfons, A. (2020). probmed: (robust) mediation analysis. R package, Version 0.7.0.

Allen, D. G., & Griffeth, R. W. (2001). Test of a mediated performance-turnover relationship highlighting the moderating roles of visibility and reward contingency. *Journal of Applied Psychology*, *86*(5), 1014–1021. https://doi.org/10.1037/0021-9010.86.5.1014.

Allison, P. D. (1995). The impact of random predictors on comparison of coefficients between models: Comment on Clogg, Petkova, and Haritou. *American Journal of Sociology*, *100*(5), 1294–1305. https://doi.org/10.1086/230639.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*(6), 1173–1182. https://doi.org/10.1037/0022-3514.51.6.1173.

Batailler, C., Muller, D., Yzerbyt, V., Judd, C., Ho, A., Kteily, N., Chen, J., Dohle, S., & Siegrist, M. (2020). Jsmediation: Mediation analysis using joint significance. R package Version 0.1.

Cui, G., Yu, X., Iommelli, S., & Kong, L. (2016). Exact distribution for the product of two correlated Gaussian random variables. *IEEE Signal Processing Letters*, *23*(11), 1662–1666. https://doi.org/10.1109/LSP.2016.2614539.

Daniel, R. M., De Stavla, B. L., Cousens, S. N., & Vansteelandt, S. (2015). Causal mediation analysis with multiple mediators. *Biometrics*, *71*(1), 1–14. https://doi.org/10.1111/biom.12248.

Eshima, N., Tabata, M., & Geng, Z. (2001). Path analysis with logistic regression models: Effect analysis of fully recursive causal systems of categorical variables. *Journal of the Japan Statistical Society*, *31*(1), 1–14. https://doi.org/10.14490/jjss1995.31.1.

Frölich, M., & Huber, M. (2017). Direct and indirect treatment effects-causal chains and mediation analysis with instrumental variables. *Journal of the Royal Statistical Society Series B*, *79*(5), 1645–1666. https://doi.org/10.1111/rssb.12232.

Guo, J., & Geng, Z. (1995). Collapsibility of logistic regression coefficients. *Journal of the Royal Statistical Society Series B*, *57*(1), 263–267. https://doi.org/10.1111/j.2517-6161.1995.tb02029.x.

Guo, J., Geng, Z., & Fung, W.K. (2001). Consecutive collapsibility of odds ratios over an ordinal background variable. *Journal of Multivariate Analysis*, *79*(1), 89–98. https://doi.org/10.1006/jmva.2000.1957.

Hayes, A. F. (2013). Introduction to mediation, moderation, and conditional process analysis: A regression-based approach. Guilford Press.

Kenny, D. A., Kashy, D. A., & Bolger, N. (1998). Data analysis in social psychology. In D. Gilbert, S. Fiske, & G. Lindzey (Eds.) *The handbook of social psychology, pp 233–265, McGraw-Hill.*

Kim, J., & Cicchetti, D. (2010). Longitudinal pathways linking child maltreatment, emotion regulation, peer relations, and psychopathology. *Journal of Child Psychology and Psychiatry*, *51*(6), 706–716. https://doi.org/10.1111/j.1469-7610.2009.02202.x.

Kisbu-Sakarya, Y., MacKinnon, D. P., & Miočević, M. (2014). The distribution of the product explains normal theory mediation confidnce interval estimation. *Multivariate Behavioral Research*, *49*(3), 261–268. https://doi.org/10.1080/00273171.2014.903162.

Lachowicz, M. J., Preacher, K. J., & Kelley, K. (2018). A novel measure of effect size for mediation analysis. *Psychological Methods*, *23*(2), 244–261. https://doi.org/10.1037/met0000165.

Lockhart, G., MacKinnon, D. P., & Ohlrich, V. (2011). Mediation analysis in psychosomatic medicine research. *Psychosomatic Medicine*, *73*(1), 29–43. https://doi.org/10.1097/PSY.0b013e318200a54b.

Loucks, E. B., Huang, Y. T., Agha, G., Chu, S., Eaton, C. B., Gilman, S. E., Buka, S. L., & Kelsey, K.T. (2016). Epigenetic mediators between childhood socioeconomic disadvantage and mid-life body mass index: The New England family study. *Psychosomatic Medicine*, *78*(9), 1053–1065. https://doi.org/10.1097/PSY.0000000000000411.

MacKinnon, D. P. (2008). Introduction to statistical mediation analysis. Taylor & Francis Group.

MacKinnon, D. P., Fairchild, A. J., & Fritz, M.S. (2007a). Mediation analysis. *Annual Review of Psychology*, *58*, 593–614. https://doi.org/10.1146/annurev.psych.58.110405.085542.

MacKinnon, D. P., Fritz, M. S., Williams, J., & Lockwood, C.M. (2007b). Distribution of the product confidence limits for the indirect effect: Program PRODCLIN. *Behavior Research Methods*, *39*, 384–389. https://doi.org/10.3758/BF03193007.

MacKinnon, D. P., Lockwood, C., & Hoffman, J. (1998). A new method to test for mediation. In *The annual meeting of the society for prevention research, Park City, USA*.

MacKinnon, D. P., & Lockwood, C. M. (2001). *Distribution of products tests for the mediated effect. Technical Report*. USA: Arizona State University.

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., Lockwood, C. M., & Williams, J. (2004). Confidence limits for the indirect effect: Distribution of the product and resampling methods. *Multivariate Behavioral Research*, *39*(1), 99–128. https://doi.org/10.1207/s15327906mbr3901_4.

MacKinnon, D. P., Warsi, G., & Dwyer, J.H. (1995). A simulation study of mediated effect measures. *Multivatiate Behavioral Research*, *30*(1), 41–62. https://doi.org/10.1207/s15327906mbr3001_3.

McGuigan, K., & Langholtz, B. (1988). A note on testing mediation paths using ordinary least-squares regression. Unpublished Note.

Meeker, W., Cornwell, L., & Aroian, L (1981). Selected table in mathematical statistics (Volume VII). The product of two normally distributed random variables. American Mathematical Society.

Micceri, T. (1989). The unicorn, the normal curve, and other improbable creatures. *Psychological Bulletin*, *105*(1), 156–166. https://doi.org/10.1037/0033-2909.105.1.156.

Nadarajah, S., & Pogány, T. K. (2016). On the distribution of the product of correlated normal random variables. *Comptes Rendus de l'académie des Sciences, Series I*, *354*(2), 201–204. https://doi.org/10.1016/j.crma.2015.10.019.

Newland, R. P., Crnic, K. A., Cox, M. J., Mills-Koonce, W. R., & Investigators, Family Life Project Key (2013). The family model stress and maternal psychological symptoms: Mediated pathways form economic hardship to parenting. *Journal of Family Psychology*, *27*(1), 96–105. https://doi.org/10.1037/a0031112.

Nübold, A., Dörr, S. L., & Maier, G.W. (2015). Considering the orphan: Personal identification and its relations with transformational leadership, trust, and performance in a three-path mediation model. *Leadership*, *11*(2), 230–254. https://doi.org/10.1177/1742715014522679.

Ogden, C. L., Carroll, M. D., Curtin, L. R., Lamb, M. M., & Flegal, K.M. (2010). Prevalence of high body mass index in U.S. children and adolescents, 2007-2008. *Journal of the American Medical Association*, *303*(3), 242–249. https://doi.org/10.1001/jama.2009.2012.

Pearl, J. (2009). *Causality: Models, reasoning and inference*, 2nd. Cambridge: Cambridge University Press.

Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods Instruments, & Computers*, *36*(4), 717–731. https://doi.org/10.3758/BF03206553.

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, *40*(3), 879–891. https://doi.org/10.3758/BRM.40.3.879.

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.

Qiu, W. (2020). powermediation: Power/sample size calculation for mediation analysis. R package, Version 0.3.2.

Richiardi, L., Bellocco, R., & Zugna, D. (2013). Mediation analysis in epidemiology: Methods interpretation and bias. *International Journal of Epidemiology*, *42*(5), 1511–1519. https://doi.org/10.1093/ije/dyt127.

Rucker, D. D., Preacher, K. J., Tormala, Z. L., & Petty, R.E. (2011). Mediation analysis in social psychology: Current practices and new recommendations. *Social and Personality Psychology Compass*, *5*(6), 359–371. https://doi.org/10.1111/j.1751-9004.2011.00355.x.

Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods*, *7*(4), 422–445. https://doi.org/10.1037/1082-989X.7.4.422.

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.

Steen, J., Loeys, T., Moerkerke, B., Vansteelandt, S., Meys, J., Lange, T., Legewie, J., & Fink, P. (2020). medflex: Flexible mediation analysis using natural effect models. R package, Version 0.6-7.

Taguri, M., Featherstone, J., & Cheng, J. (2018). Causal mediation analysis with multiple causally non-ordered mediator. *Statistical Methods in Medical Research*, *27*(1), 3–19. https://doi.org/10.1177/0962280215615899.

Taylor, A. B., MacKinnon, D. P., & Tein, J.Y. (2008). Tests of the three-path mediated effect. *Organizational Research Methods*, *11*(2), 241–269. https://doi.org/10.1177/1094428107300344.

Tein, J. Y., Sandler, I. N., & Zautra, A.J. (2000). Stressful life events, psychological distress, coping, and parenting of divorced mothers: A longitudinal study. *Journal of Family Psychology*, *14*(1), 27–41. https://doi.org/10.1037/0893-3200.14.1.27.

Tekleab, A. G., Bartol, K. M., & Liu, W. (2005). Is it pay level or pay raises that matter to fairness and turnover? *Journal of Organizational Behavior*, *26*, 899–921. https://doi.org/10.1002/job.352.

Tingley, D., Yamamoto, T., Hirose, K., Keele, L., Imai, K., Trinh, M., & Wong, W. (2019). mediation: Causal mediation analysis. R package, Version 4.5.0.

Tofighi, D., & MacKinnon, D. P. (2011). Rmediation: An R package for mediation analysis confidence intervals. *Behavior Research Methods*, *43*(3), 692–700. https://doi.org/10.3758/s13428-011-0076-x.

Tofighi, D., & MacKinnon, D. P. (2016). Rmediation: Mediation analysis confidence intervals. R package Version 1.1.4.

VanderWeele, T. J. (2015). *Explanation in causal inference: Methods for mediation and interaction*. Oxford: Oxford University Press.

VanderWeele, T. J., & Tchetgen Tchetgen, E. J. (2017). Mediation analysis with time varying exposures and mediators. *Journal of the Royal Statistical Society: Series B*, *79*(3), 917–938. https://doi.org/10.1111/rssb.12194.

Wang, W. W., & Yu, P. (2020). Nonequivalence of two least-absolute-deviation estimators for mediation effect. Complete Manuscript.

Wang, W. W., Yu, P., Lin, L., & Tong, T. (2019). Robust estimation of derivatives using locally weighted least absolute deviation regression. *Journal of Machine Learning Research*, *20*(60), 1–49.

White, H. (2000). A reality check for data snooping. *Econometrica*, *68*(5), 1097–1126. https://doi.org/10.1111/1468-0262.00152.

Yuan, Y., & MacKinnon, D. P. (2014). Robust mediation analysis based on median regression. *Psychological Methods*, *19*(1), 1–20. https://doi.org/10.1037/a0033820.

Zhao, Z., & Xiao, Z. (2014). Efficient regression via optimally combining quantile information. *Econometric Theory*, *30*(6), 1272–1314. https://doi.org/10.1017/S0266466614000176.