Seoul Journal of Business
Volume 27, Number 1 (June 2021)
DOI: https://doi.org/10.35152/snusjb.2021.27.1.001

The Marginal Decomposition Approach Quantifying Direct and Indirect Effects in Causal Models*

SANGJUNE PARK**
Jeonbuk National University
Jeollabuk-do, Korea

YOUJAE YI***
Seoul National University
Seoul, Korea

ABSTRACT

Many researchers have analyzed causal mediation with the measures for direct and indirect effects in a system of regression models. The direct effect indicates the effect of a focal predictor not through a mediator, whereas the indirect effect indicates the effect of the focal predictor through a mediator. Various versions of two approaches (product approach and potential outcomes approach) have been used to find the measures indicating the quantified direct and indirect effects in a system of regression models. However, it may not be easy to identify the measures with the two approaches, because they do not provide a general formula for identifying the measures in various systems of regression models. Thus, this paper proposes a new approach providing a general formula for identifying the measures intuitively and clearly. The new approach decomposes the effect of a focal variable on a dependent variable into five additive components in view of moderation and mediation.

Keywords: causal mediation, direct effect, indirect effects

---

* This research was supported by the Institute of Management Research at Seoul National University.

** Professor of Marketing, Department of Business Administration, Jeonbuk National University. 567, Baekje-daero, Deokjin-gu, Jeonju-si, Jeollabuk-do, Republic of Korea. e-mail: psj@jbnu.ac.kr.

*** Professor of Marketing, College of Business Administration, Seoul National University. 1 Gwanak-ro, Gwanak-gu, Seoul, 08826, Republic of Korea. e-mail: youjae@snu.ac.kr.

INTRODUCTION

Many researchers have analyzed causal mediation with a system of regression models consisting of a dependent variable regression model and a mediator regression model (sometimes multiple mediator regression models). They have quantified the direct and indirect effects of a focal predictor in a system of regression models with various versions of two approaches (the product approach and the potential outcomes approach) and have diagnosed causal mediation with direct and indirect effects. The direct effect indicates the effect of a focal predictor on a final outcome not through a mediator, whereas the indirect effect indicates the effect of the focal predictor through a mediator. There are unique measures for direct and indirect effects corresponding to a system of regression models.

Previous research has identified the measures corresponding to a system of regression models with two approaches: the product approach and the potential outcomes approach. The product approach finds the measures with the impact functions of a focal predictor and a mediator expressed in a system of regression models (e.g., Baron and Kenny 1986; MacKinnon, Warsi, and Dwyer 1995; Preacher et al. 2007; Sobel 1982). The potential outcomes approach finds the measures with the impact functions recovered with the changes in potential outcomes by changes of a focal predictor and a mediator in a system of regression models (e.g., Imai, Keele, and Tingley 2010; Imai, Keele, and Yamamoto 2010; VanderWeele 2014). There are three versions of the product approach: the simple version, the extended version, and the generalized version. There are

Table 1. Previous Approaches and Representative Examples

Previous Approach		Representative Examples
The Product Approach	The Simple Version	Baron and Kenny (1986), Sobel (1982)
	The Extended Version	Hayes (2013)
	The Generalized Version	Hayes and Preacher (2010)
The Potential Outcomes Approach	The Basic Version	Imai et al. (2010), VanderWeele (2014)
	The Generalized Version	Knafl et al. (2017)

two versions of the potential outcomes approach: the basic version and the generalized version.

The simple version of the product approach is based on multiplying the coefficients corresponding to the variables that are causally linked in a system of linear regression models. The extended version of the product approach is used to identify the measures for conditional direct and indirect effects in such a system. The conditional direct and indirect effects imply that the direct and indirect effects of a focal predictor vary depending on moderators. In addition, the generalized version of the product approach is used to identify the measures in a system of nonlinear regression models. It is also possible to identify the measures with two versions of the potential outcomes approach. One is the basic version of the potential outcomes approach corresponding to the simple version and the extended version of the product approach, in view that it can identify the measures in a system of linear regression models. The other is the generalized version of the potential outcomes approach corresponding to the generalized version of the product approach, in view that it can identify the measures in a system of nonlinear regression models.

However, there may be a case in which it is not easy to find the measures for direct and indirect effects with the previous approaches, if one is not familiar with the various versions of the product approach or the potential outcomes approach. Previous research does not provide a general formula for identifying the measures in a system of regression models.

We propose the marginal decomposition approach decomposing the effect of a focal predictor on a dependent variable into various effects in view of moderation and mediation. We assume that a focal predictor is exogenous and the assumption can be satisfied for the data obtained from controlled experiments. Note that an independent variable is assumed to be exogenous in regression analysis examining the causal effect of the independent variable on the final outcome. The various effects are summarized as the marginal effects on two sources of the focal predictor (direct and indirect effects). The marginal decomposition approach provides a general formula with which one can intuitively and clearly find the measures for direct and indirect effects in any system of regression models. Thus, the measures identified by the marginal decomposition approach are consistent with the econometric interpretation on marginal effects in a regression model. The basic idea for the marginal decomposition approach is also adopted in previous research identifying the measures for direct and indirect effects in a system of nonlinear regression models such as the generalized version of the product approach and the generalized version of the potential outcomes approach.

The remainder of the paper is organized as follows. In the next section, we explain basic concepts of moderation and mediation and review previous approaches identifying the measures for direct and indirect effects in a system of regression models. Then, we introduce the marginal decomposition approach and provide some implications for identifying the measures.

BACKGROUND

Causal mediation analysis refers to various analyses to investigate causal mediation in a variety of disciplines including epidemiology, psychology, sociology, and so on. For example, researchers are interested in whether a new form of therapy is more effective than existing methods for treating certain conditions, or whether people who have certain experiences such as psychological trauma are more likely to suffer later in life from certain symptoms (Hayes and Rockwood 2017). Causal mediation analysis aims at rigorously ensuring causal identification of effects using techniques like randomization, covariate adjustment, and instrumental variables for tackling confounding. Thus, there may be differences in experiment designs or data collection methods in causal mediation analysis.

If a researcher wants to investigate casual mediation with the direct and indirect effects of a focal predictor, the researcher has to find the measures for direct and indirect effects in a theoretical model representing the causal mediation, and statistically estimate and diagnose the measures. This paper aims to propose a new scheme for identifying the measures for direct and indirect effects in a system of regression models. Thus, it can be used regardless of experimental designs. It can be applied regardless of techniques estimating and diagnosing the identified measures for direct and indirect effects. In the following section, we briefly review two previous approaches identifying the measures for direct and indirect effects in a system of regression models.

Basic Concepts

Let us consider the mediation effect in a system of regression models consisting of a dependent variable regression model and a mediator regression model. The dependent variable regression model captures the effect of a focal predictor (X) on a dependent variable (Y), which is defined as a function of a focal predictor, a mediator, a moderator and covariates. In contrast, the mediator regression model captures the effect of X on a mediator (M), which is defined as a function of a focal predictor, a moderator and covariates. If the dependent variable is a function of multiple mediators, one can use multiple mediator regression models. In the case of multiple mediators, a mediator regression model may be formulated as a function of a focal predictor and other mediators.

A third variable is defined as a mediator (M) when X affects Y indirectly via its effect on the third variable (Baron and Kenny 1986). That is, a mediation effect of M (an indirect effect of X) occurs when the effect of X on Y is transmitted by M (Preacher, Rucker, and Hayes 2007). If the mediation effect of M is expressed as a function of other variables, it is referred to as the context-dependent mediation effect. The context-dependent mediation effect is also referred to as the conditional indirect effect (Hayes 2013). The moderation effect [i.e., the moderated effect of a focal predictor (X)] indicates the effect of X that varies depending on other variables. It is measured with the interaction effect of X and a moderator in the dependent variable regression model. Unlike a mediator, a moderator is not expressed as a function of X. If the indirect effect of X interacts with a third variable, the effect of X is referred to as the moderated mediation effect. In contrast, the effect of X only through the mediator can be referred to as the pure mediation effect. The effect of X resulting from its interaction with non-mediators can be referred to as the pure moderation effect. If the effect of X does not depend on any moderator or mediator, it is referred to as the perfectly pure effect of X (the effect of X due to neither moderation nor mediation).

In sum, the direct effect indicates the effect of the focal predictor not through mediator(s), whereas the indirect effect indicates the effect of the focal predictor through mediator(s). Thus, the direct effect and the indirect effect can be referred to as the non-mediated effect and the mediated effect, respectively. In this view, the direct effect of *X* includes the perfectly pure effect and the pure moderation effect, whereas the indirect effect includes the pure mediation effect. The moderated mediation effect is the direct and indirect effect at the same time.

The Product Approach

The simple version of the product approach is the most popular approach used to detect a mediation effect (e.g., Baron and Kenny 1986; MacKinnon, Warsi, and Dwyer 1995; Preacher et al. 2007; Sobel 1982), which is rooted in Baron and Kenney's (1986) test and Sobel's (1982) test. The simple version measures the indirect effect of a focal predictor as the product of two slopes in a system of regression model describing that a focal predictor affects a final outcome through a mediator. One slope indicates the coefficient representing the impact of the focal predictor in the mediator regression model in which the mediator is determined by the focal predictor. The other slope indicates the coefficient representing the impact of the mediator in the dependent variable regression model in which a dependent variable is determined by the mediator and the focal predictor. In other words, the indirect effect is determined by a product of the coefficients indicating the effects of causal paths from a focal predictor to a dependent variable through a mediator. In contrast, the direct effect is determined by the coefficient indicating the effects of causal paths not through a mediator. In a simple system of regression models with a focal predictor, a mediator, and a dependent variable, previous researchers assess the simple mediation (indicating the effect of *X* via *M*) with the simple version of the product approach (e.g., Baron and Kenny 1986; Judd and Kenny 1981; Sobel 1982).

Recently, some researchers (Hayes 2013; Preacher et al. 2007) assess the mediation effects that vary depending on contexts (groups or a moderator variable) in a more complex system of regression models (e.g., Preacher et al. 2007). They find the measures for conditional direct and indirect effects with the extended version of the product approach. The conditional direct effect (or the conditional indirect effect) refers to the direct effect (or indirect effect) of a focal predictor that varies depending on moderators.

One can see various systems of regression models with which one can analyze different mediation effects in Hayes (2013). Hayes (2013) presents the measures for the conditional direct and indirect effects in the 76 systems identified with the extended version of the product approach. However, one's research model may not be represented by any of the 76 models. In such a case, it may not be easy to find the measures for the conditional effects.

The simple version and the extended version of the product approach can be applied to a system of linear regression models. In a system of nonlinear regression models, previous research has used the generalized version of the product approach to identify the measures for direct and indirect effects (e.g., Hayes and Preacher 2010). The generalized version of the product approach uses instantaneous rates of changes of a focal predictor and a mediator to identify the measures for direct and indirect effects. For example, in a system of nonlinear regression models with a focal predictor, a mediator and a final outcome, the generalized version of the product approach quantifies the indirect effect of the focal predictor with the product of the first partial derivative of the focal predictor in the mediator regression model and that of the mediator in the dependent variable regression model. The direct effect is quantified with the first partial derivative of the focal predictor in the dependent variable regression model. Note that the generalized version assumes that the focal predictor and the mediator are continuous and differentiable in a system of regression models under the assumption that the final outcome variable is continuous and differentiable.

However, it may be valuable to confirm whether the product approach can really identify the measures for direct and indirect effects with an alternative approach, because the product approach (including various versions) was not developed based on a strong theory. Besides, one may confront a case in which it is not easy to identify the measures with the product approach, because previous research does not provide a general formula with which one can identify the measures in any system of regression models with the product approach. If a researcher is not familiar with the various versions of the product approach, it may not be easy to identify the measures in a system of nonlinear regression models as well as linear regression models or a complex system in which the direct and indirect effects vary depending on moderators.

The Potential Outcomes Approach

There is a causal mediation analysis designed to diagnose causal mediation with a system of regression models. There is also a causal mediation analysis designed to diagnose causal mediation without specifying a model representing the mechanism for causal mediation. In the causal mediation analysis without relying on a specific model, it is common to collect data based on an experiment, quantify and diagnose the direct and indirect effects based on the differences between groups in the experimental design.

However, even in the causal mediation analysis designed not relying on a specific model, it is possible to find the measures for direct and indirect effects based on potential outcomes in a system of regression models, if it is possible to assume that causal mediation is theoretically specified by a system of regression models where certain key assumptions are satisfied. The key assumptions include “sequential ignorability,” “stable unit treatment value,” and others (see Preacher 2015 for more details).

Some researchers have proposed the basic version of the potential outcomes approach identifying the measures based on potential outcomes in a system of linear regression models under the key assumptions (e.g., Imai, Keele, and Tingley 2010; Imai, Keele, and Yamamoto 2010; VanderWeele 2014). The basic version of the potential outcomes approach identifies the measures based on the changes in potential outcomes by discrete changes of a focal predictor and a mediator in a system of linear regression models. It can also be used to identify the measures in a causal mediation analysis designed to use a system of regression models. The previous researchers have shown that the basic version of the potential outcomes approach can be used as an alternative of the simple version and the extended version of the product approach.

If a system of nonlinear regression models is used as a theoretical model specifying causal mediation, one can use the generalized version of the potential outcomes approach identifying the measures based on partial derivatives like the generalized version of the product approach (e.g., Knafl et al. 2017). Like the generalized version of the product approach, the generalized version of the potential outcomes approach assumes that the focal predictor and the mediator are continuous and differentiable under the assumption that the final outcome variable is continuous and differentiable.

However, like the product approach there may be a case that it is not easy to find the measures for direct and indirect effects with the potential outcomes approach, if one is not familiar with the potential outcomes approach. Previous research does not provide a general formula for identifying the measures in any system of regression models with the potential outcome approach. Furthermore, previous research does not provide the measures in various systems identified by the potential outcomes approach. In contrast, one can see the measures for the direct and indirect effects in various systems identified by the extended version of the product approach in the literature (see Hayes 2013). Thus, in the following section, we propose a new approach identifying the measures for direct and indirect effects and then compare the measures identified by the proposed approach and those identified by the product approach.

THE MARGINAL DECOMPOSITION APPROACH

In this section, we propose the marginal decomposition approach finding the measures for direct and indirect effects of a focal predictor in a system of regression models. It is consistent with the marginal effect in a regression model. In a dependent variable regression model, the effect of a focal predictor (X) on a dependent variable (Y) indicates the marginal effect of X on Y ($\partial Y / \partial X$). The marginal effect of a focal predictor on the dependent variable can be decomposed into two marginal effects when the effect of a focal predictor is mediated by a mediator. One is the marginal effect of the focal predictor not through the mediator (i.e., the direct effect) and the other is the marginal effect of the focal predictor through the mediator (i.e., the indirect effect). As discussed before, this idea has been adopted to quantify indirect effects when the constituent paths are nonlinear in previous research (e.g., Hayes and Preacher 2010; Knafl et al. 2017). In the following section, we present the marginal decomposition approach providing a general formula with which one can intuitively and clearly quantify direct and indirect effects in any system of regression models.

Decomposition Procedure

In a dependent variable regression model, it is possible to decompose the effects of a focal predictor according to the characteristics of effects: i.e., the effects of a focal predictor due to neither moderation nor mediation, due to moderation only, due to mediation only, and due to moderation and mediation. The decomposed model can be represented as:

$$Y = Y_O + Y_{PP} + Y_{PO} + Y_{PE} + Y_{MM} + e_Y. \tag{1}$$

where $Y_{PP} \equiv Y_{PP(X)}$, $Y_{PO} \equiv Y_{PO(X)}$, $Y_{PE} \equiv Y_{PE(X)}$, $Y_{MM} \equiv Y_{MM(X)}$, and Y_O indicate the perfectly pure effect of X (due to neither moderation nor mediation), the pure moderated effect of X (due to moderation only), the pure mediated effect of X (due to mediation only), the moderated mediation effect of X (due to both moderation and mediation), and the others.

Accordingly, the effect of X can be decomposed into four additive components:

$$TE(X) = PP(X) + PO(X) + PE(X) + MM(X), \tag{2}$$

where $TE(X) = \dfrac{\partial Y}{\partial X}$, $\text{PP(X)}= \dfrac{\partial Y_{PP}}{\partial X}$, $PE\left(X\right) = \dfrac{\partial Y_{PE}}{\partial X}$,

$PO(X) = \dfrac{\partial Y_{PO}}{\partial X}$, and $MM(X) = \dfrac{\partial Y_{MM}}{\partial X}$.

In equation 2,

$TE(X)$ is the total effect of X indicating the sum of decomposed effects,
$PP(X)$ is the perfectly pure effect of X (due to neither moderation nor mediation),
$PO(X)$ is the pure moderated effect of X (due to moderation only),
$PE(X)$ is the pure mediated effect of X (due to mediation only),
$MM(X)$ is the moderated mediation effect of X (due to both moderation and mediation).

If a mediator (M) is also a moderator in the relationship between a focal predictor (X) and the outcome (Y), the marginal effect of X

on the interaction term between X and M (capturing the moderated mediation effect) can be decomposed into two effects. One is the direct moderation effect (not through the marginal effect of X on the mediator) and the other is the indirect moderation effect (through the marginal effect of X on the mediator). Thus, it is possible to represent that:

$$MM(X) = MM_O(X) + MM_E(X), \tag{3}$$

where $MM_O(X)$ indicates the moderation effect by mediator(s) (direct moderation effect) and $MM_E(X)$ indicates the mediation effect by mediator(s) (indirect moderation effect).

More specifically, the moderation effect is captured by the interaction effect of M and X on Y, i.e., $Y_{MM} = \gamma XM$ for $M = M(X)$. The moderation effect is expressed as:

$$MM(X) \equiv \frac{\partial Y_{MM}}{\partial X} = \frac{\partial rXM}{\partial X} = MM_O(X) + MM_E(X), \tag{4}$$

where $MM_O(X) = rM(X)$ and $MM_E(X) = rX\dfrac{\partial M(X)}{\partial X}$.

In equation 4, the first term, $rM(X)$, can be interpreted as the direct moderation effect in the interaction term because the effect is not determined by the change in M (resulting from the marginal change of X) on Y. The effect captures the marginal effect of X in the interaction term on Y, i.e., $rM(X) = rM(X)\ (\partial X/\partial X)$. In contrast, the second term, $rX\{\partial M(X)/\partial X\}$, can be interpreted as the indirect moderation effect in the interaction term because it is determined by the change in M (resulting from the marginal change of X) on Y.

Finally, the effect of X on Y can be decomposed into five additive components:

$$TE(X) = PP(X) + PO(X) + PE(X) + MM_O(X) + MM_E(X). \tag{5}$$

In view of mediation effects, $PP(X)$ and $MM_O(X)$ are the direct effects (non-mediated effects) of X whereas $PE(X)$ and $MM_E(X)$ are the indirect effects (mediated effects) of X. We summarize this implication in Theorem 1.

Theorem 1: The conditional effects (total effect, direct effect, and indirect effect) of a focal predictor are represented as:

$$TE(X) = DE(X) + IE(X), \tag{6}$$

where $DE(X) = PE(X) + PO(X) + MM_O(X)$ and $IE(X) = PE(X) + MM_E(X)$.

Proof. See equations 1 to 5.

Like the generalized version of the potential outcomes approach and that of the product approach, the marginal decomposition approach also assumes that the focal predictor and the mediator are continuous and differentiable in a system of regression models under the assumption that the final outcome variable is continuous and differentiable. However, it is possible to find the measures for direct and indirect effect with a slightly modified version of the marginal decomposition approach which identifies the measures based on discrete changes of the focal predictor and the mediator, even if the assumption is not satisfied. We will discuss this for more details in the final section.

Illustration of the Marginal Decomposition Approach

For example, it is possible to consider a system of regression models expressed as:

$$M = \alpha_0 + \alpha X + e_M, \tag{7}$$

$$\begin{aligned} Y &= \beta_0 + \beta_X X + \beta_M M + \beta_{XM} XM + \beta_{XZ} XZ + \beta_{MW} MW + e_Y \\ &= \beta_0 + \beta_X X + \beta_{XZ} XZ + (\beta_M + \beta_{MW} W)M + \beta_{XM} MX + e_Y, \end{aligned} \tag{8}$$

where X, M, and Y indicate the focal predictor, the mediator, and the final outcome. The system of regression models shows that Z has a moderating role on the effect of X on Y and W has a moderating role on the effect of M on Y. Equation 8 can be written as:

$$Y = Y_O + Y_{PP} + Y_{PO} + Y_{PE} + Y_{MM} + e_Y, \tag{9}$$

where $Y_O = \beta_0$, $Y_{PP} = \beta_X X$, $Y_{PE} = (\beta_M + \beta_{MW} W)M$,
$Y_{PO} = \beta_{XZ} ZX$, $Y_{MM} = \beta_{XM} MX$, and $M = M(X)$.

Accordingly,

$$TE(X) = DE(X) + IE(X), \tag{10}$$

where $DE(X) = PE(X) + PO(X) + MM_O\,(X)$ and $IE(X) = PE(X) + MM_E(X)$,

$$PP(X) = \frac{\partial Y_{PP}}{\partial X} = \beta_X,\ PO(X) = \frac{\partial Y_{PO}}{\partial X} = \beta_{XZ} Z,$$

$$PE(X) = \frac{\partial Y_{PE}}{\partial X} = (\beta_M + \beta_{MW} W)\frac{\partial M(X)}{\partial X} = \alpha(\beta_M + \beta_{MW} W),$$

$MM_O(X) = \gamma M$, and $MM_E(X) = \gamma X \dfrac{\partial M(X)}{X} = \alpha\gamma X.$

Note that it is easy to represent the dependent variable regression model like equation 9, according to the definitions of the additive components. Consequently, it is possible to intuitively and clearly identify the measures for the direct and indirect effects in any system of regression models based on Theorem 1.

APPLICATIONS

We compared the measures for direct and indirect effects identified by the product approach and those by the marginal decomposition approach in the cases selected from the systems of regression models presented in the literature. With the comparisons, we have confirmed that the two alternative approaches lead to the identical measures for the direct and indirect effects in various systems of linear regression models. In appendix, one can see the measures in some models that are identified with the marginal decomposition approach.

Next, let us compare the two alternative approaches in a system of nonlinear regression models. One may consider a system of nonlinear mediator regression models: e.g.,

$$M = \alpha_0 + \alpha X^2 + e_M, \tag{11}$$

$$Y = \beta_0 + \beta_M M + \beta_X X + e_Y. \tag{12}$$

The measures at $X = x$ in the marginal decomposition approach are written as:

$$DE(x) = \beta_X,\ IE(x) = 2\alpha\beta_M x, \text{ and } TE(x) = \beta_X + 2\alpha\beta_M x, \tag{13}$$

because:

$$DE(x) = PP(x) + PO(x) + MM_O(x),\ IE(x) = ME(x) + MM_E(x),$$
$$TE(x) = DE(x) + IE(x),$$

where $PP(X) = \beta_X \dfrac{\partial X}{\partial X} = \beta_X,\ PO(X) = 0,\ ME(X) = \beta \dfrac{\partial M(X)}{\partial X} = 2\alpha\beta_M X,$

$MM_O(X) = 0$, and $MM_E(X) = 0$

Note that the simple and extended versions of the product approach as well as the basic version of the potential outcomes approach may lead to incorrect measures in a system of nonlinear regression models. Thus, in such a system one should use the generalized version of the product approach (e.g., Hayes and Preacher 2010) or the generalized version of the potential outcomes approach (e.g., Knafl et al. 2017). The two generalized versions identify the measures based on partial derivatives like the marginal decomposition approach.

For example, the simple version of the product approach identifies the measures for direct and indirect effects with products of path coefficients in a diagram representing causal relationships, even when the causal relationships are represented as a system of nonlinear regression models. Thus, the simple version of the product approach leads to the measures at $X = x$ in the system of nonlinear regression models (equations 11 and 12) written as:

$$DE_C(x) = \beta_X,\ IE_C(x) = \alpha\beta_M, \text{ and } TE_C(x) = \beta_X + \alpha\beta_M, \tag{14}$$

where $DE_C(x)$, $IE_C(x)$, and $TE_C(x)$ indicates the direct effect, the indirect effect, and total effect measured by the product approach. In contrast, the generalized version of the product approach leads to the measures at $X = x$ in the system represented as:

$$DE_C(x) = \beta_X,\ IE_C(x) = 2\alpha\beta_M x, \text{ and } TE_C(x) = \beta_X + 2\alpha\beta_M x, \tag{15}$$

because $IE_C(X) = \dfrac{\partial M(X)}{\partial x} \dfrac{\partial Y(x, M(X))}{\partial M(X)} = 2\alpha\beta_M x.$

One can see that the indirect effect is calculated with a product of the marginal effects (instead of the coefficients) that are linked to the path representing the mediated relationship in the generalized version of the product approach. With the general formula provided by the marginal decomposition approach in Theorem 1, one can intuitively and clearly quantify the direct and indirect effects even in a system of nonlinear regression models (see equations 11 to 13).

CONCLUSION

Many researchers have identified the measures for direct and indirect effects with the product approach (including the simple, extended and generalized versions) and the potential outcomes approach (including the basic version and its generalized version). More specifically, one can use the extended version of the product approach (instead of the simple version) in a complex system of linear regression models in which the direct and indirect effects vary depending on moderators. In a system of nonlinear regression models, one may use the generalized version of the product approach (instead of the simple and extended versions). It is also possible to use the potential outcomes approach and its generalized version corresponding to the extended version and the generalized version of the product approach.

However, it is not clear how one can find the measures with the previous approaches in a system of nonlinear regression models as well as linear regression model or a complex system in which the direct and indirect effects vary depending on moderators, because previous research does not provide a general formula identifying the measures for direct and indirect effects in a system of regression models. Thus, we have proposed the marginal decomposition approach that decomposes the effect of a focal predictor into various effects in view of moderation and mediation. Unlike the various versions of the product approach and the potential outcomes approach, the marginal decomposition approach provides a general formula with which one can intuitively and clearly find the measures for direct and indirect effects in any system of regression models. In addition, the marginal decomposition approach decomposes the effect of a focal predictor into various effects: i.e., the perfectly pure effect (due to neither moderation nor mediation), the pure moderated effect (due to moderation only), the pure mediated effect (due to mediation only), and the moderated mediation effect (due to both moderation and mediation). Then it classifies the decomposed effects of the focal predictor as the direct effect (not through the mediator) and the indirect effect (through the mediator). Thus, this approach lets us know the links between the conditional direct and indirect effects of a focal predictor and the differently classified effects of the focal predictor.

Some limitations of the present study and ideas for future research should be noted. The marginal decomposition approach assumes that a focal predictor and a mediator are continuous and differentiable in a system of regression models under the assumption that a final outcome variable is continuous and differentiable. However, one may confront a case that explanatory variables are discrete and/or binary in a system of regression models. If the assumption is not satisfied in the marginal decomposition approach, one can use the basic version of the potential outcomes approach identifying the measures based on the changes in potential outcomes by discrete changes of a focal predictor and a mediator. It is notable that the three alternative approaches (the product approach, the potential outcomes approach, and the marginal decomposition approach) have advantages and disadvantages in identifying the measures for direct and indirect effects. It would be valuable to develop a framework that unifies the three approaches.

REFERENCES

Baron, R. M. and D. A. Kenny (1986), "Moderator-mediator Variables Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations." *Journal of Personality and Social Psychology* 51(6): 1173–1182.

Hayes, A. F. (2013), *Introduction to Mediation, Moderation, and Conditional Process Analysis*: *A Regression Based Approach*. New York: The Guilford Press.

Hayes, A. F. and K. J. Preacher (2010), "Quantifying and Testing Indirect

Effects in Simple Mediation Models When the Constituent Paths Are Nonlinear." *Multivariate Behavioral Research* 45: 627-660.

Hayes, A. F. and N. J. Rockwood (2017), "Regression-based Statistical Mediation and Moderation Analysis in Clinical Research: Observations, Recommendations, and Implementation." *Behaviour Research and Therapy* 98: 39-57.

Imai, K., L. Keele, and D. Tingley (2010), "A General Approach to Causal Mediation analysis." *Psychological Methods* 15: 309-334.

Imai, K., L. Keele, and T. Yamamoto (2010), "Identification, Inference and Sensitivity Analysis for Causal Mediation Effects." *Statistical Science* 25: 51-71.

Judd, C. M. and D. A. Kenny (1981), "Process Analysis: Estimating Mediation in Treatment Evaluations." *Evaluation Review* 5: 602–619.

Knafl, G. J., K. A., Knafl, M. Grey, J. Dixon, J. A. Deatrick, and A. M. Gallo (2017), "Incorporating Nonlinearity into Mediation Analyses," *BMC Medical Research Methodology*, 17: 45.

MacKinnon, D. P., G. Warsi, and J. H. Dwyer (1995), "A Simulation Study of Mediated Effect Measures." *Multivariate Behavioral Research* 30: 41–62.

Preacher, K. J. (2015), "Advances in Mediation Analysis: A survey and Synthesis of New Developments." *Annual Review of Psychology* 66: 825-852.

Preacher, K. J., D. D. Rucker, and A. F. Hayes (2007), "Addressing Moderated Mediation Hypotheses: Theory, Methods, and Prescriptions." *Multivariate Behavioral Research* 42(1): 185-227.

Sobel, M. E. (1982), "Asymptotic Confidence Intervals for Indirect Effects in Structural Equation Models." *Sociological Methodology* 13: 290-312.

VanderWeele, T. J. (2014), "A Unification of Mediation and Interaction: A 4-way Decomposition." *Epidemiology* 25(5): 749-761.

Received December 14, 2020
Revised January 15, 2021
Accepted January 29, 2021

APPENDIX

Case I: A Case of One Mediator

Let us consider two regression models indicating the relationships among three variables, which are expressed as:

$$M = a_0 + aX + e_M, \tag{A1}$$

$$Y = b_0 + bM + c' X + e_Y. \tag{A2}$$

Then let us see how the product approach identifies the measure for the indirect effect of a focal variable in a case of three variables: a focal predictor (X), a mediator (M), and a dependent variable (Y). This system has been analyzed by many researchers (e.g., Baron and Kenny 1986; MacKinnon et al. 1995; Preacher et al. 2007; Sobel 1982).

Theorem 1 leads to the measures for direct and indirect effects at $X = x$ and $M = m$ corresponding to the three-variable system of equations A1 and A2 written as:

$$DE(x) = c',\ IE(x) = ab, \text{ and } TE(x) = c' + ab, \tag{A3}$$

because:

$$DE(x) = PP(x) + PO(x) + MM_O(x),\ IE(x) = PE(x) + MM_E(x),$$
$$TE(x) = DE(x) + IE(x),$$

where $PP(X) = c' \dfrac{\partial X}{\partial X} = c',\ PO(X) = 0,\ PE(X) = b \dfrac{\partial M(X)}{\partial X} = ab,$

$MM_O(X) = 0$, and $MM_E(X) = 0$.

One can see that the measures for direct and indirect effects are identical to those derived by the product approach.

Case II: A Case for Two Mediators

When the effect of X on Y is sequentially mediated by two mediators M_1 and M_2, the system of equations can be represented as:

$$M_1 = a_{01} + a_1 X + e_{M_1}, \tag{A4}$$

$$M_2 = a_{02} + a_2 X + d_{21} M_1 + e_{M_2}, \tag{A5}$$

$$Y = b_0 + c' X + b_1 M_1 + b_2 M_2 + e_Y. \tag{A6}$$

The system is represented by Figure 1.

Theorem 1 leads to the measures for direct and indirect effects

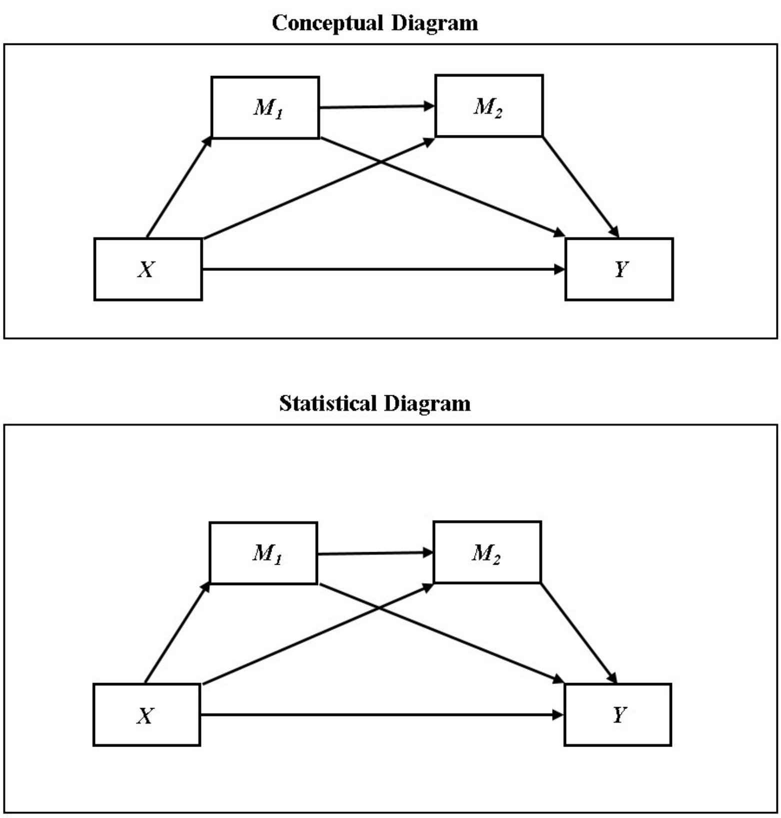

Figure 1. A Case of Two Mediators

at $X = x$ corresponding to the system of equations A4, A5, and A6 written as:

$$DE(x) = c',\ IE(x) = b_1a_1 + b_2(a_2 + d_{21}a_1),$$
$$\text{and } TE(x) = c' + a_1b_1 + a_2b_2 + d_{21}a_1b_2, \tag{A7}$$

because:

$$DE(x) = PP(x) + PO(x) + MM_O(x),\ IE(x) = PE(x) + MM_E(x),$$
$$TE(x) = DE(x) + IE(x),$$

where $PP(X) = c'\dfrac{\partial X}{\partial X} = c',\ PO(X) = 0,\ PE(X) = b_1\dfrac{\partial M_1(X)}{\partial X} + b_2\dfrac{\partial M_2(X, M_1(X))}{\partial X}$,

$$\frac{\partial M_1(X)}{\partial X} = a_1,\ \frac{\partial M_2(v)}{\partial X} = \frac{\partial M_2(X, M_1(X))}{\partial X}$$
$$+ \frac{\partial M_2(X, M_1(X))}{\partial M_1(X)}\frac{\partial M_1(X)}{\partial X} = a_2 + d_{21}a_1$$

$$MM_O(X) = 0, \text{ and } MM_E(X) = 0.$$

One can see that the product approach leads to the identical measures for direct and indirect effects from Model 6 in Hayes (2013).

Case III: A Case for One Mediator and Three Moderators

As a complex system, let us consider a system of regression models written as:

$$\begin{aligned} M_i &= a_{0i} + a_{1i}X + a_{2i}W + a_{3i}Z + a_{4i}WX + a_{5i}XZ + a_{6i}WZ + a_{7i}XWZ + e_{Mi} \\ &= a_{0i} + (a_{1i} + a_{4i}W + a_{5i}Z + a_{7i}WZ)X + a_{2i}W + a_{3i}Z + a_{6i}WZ + e_{Mi}, \end{aligned} \tag{A8}$$

$$\begin{aligned} Y &= b_0 + c_1'X + c_2'W + c_3'Z + c_4'XW + c_5'XZ + c_6'WZ + c_7'XWZ \\ &\quad + b_{1i}M_i + b_2V + b_{3i}WM_i + e_Y \\ &= b_0 + (c_1' + c_4'W + c_5'Z + c_7'WZ)X + (b_{1i} + b_{3i}W)M \\ &\quad + c_2'W + c_3'Z + c_6'WZ + b_2V + e_Y. \end{aligned} \tag{A9}$$

The system is represented by Figure 2.

Theorem 1 leads to the measures for direct and indirect effects at

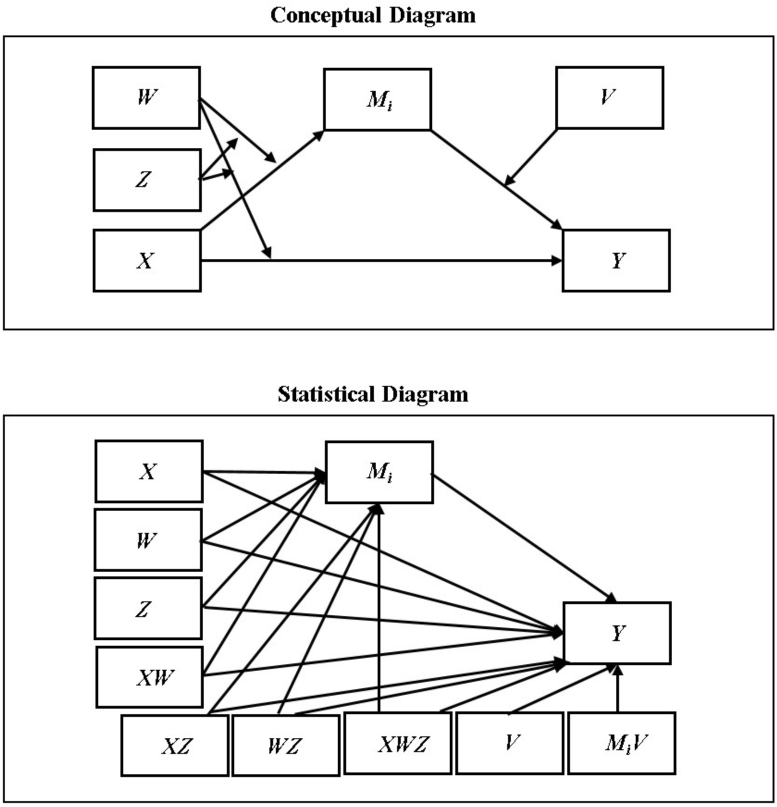

Figure 2. A Case of One Mediator and Three Moderators

$X = x$, $Z = z$, and $V = v$ corresponding to the system of equations A8 and A9 written as:

$$DE(x) = c_1' + c_4'w + c_5'z + c_7'wz,$$
$$IE(x) = (b_{1i} + b_{2i}v)(a_{1i} + a_{4i}w + a_{5i}z + a_{7i}wz),$$
$$TE(x) = DE(x) + IE(x), \tag{A10}$$

because:

$$DE(x) = PP(x) + PO(x) + MM_O(x),$$
$$IE(x) = PE(x) + MM_E(x),\ TE(x) = DE(x) + IE(x),$$

where $PP(X) = c_1'$, $PO(X) = c_4'W + c_5'Z + c_7'WZ$,

$$PE(X) = (b_{1i} + b_{2i}V)\frac{\partial M_i(X)}{\partial X},\ \frac{\partial M_i(X)}{\partial X} = a_{1i} + a_{4i}W + a_{5i}Z + a_{7i}WZ,$$

$MM_O(X) = 0$, and $MM_E(X) = 0$.

One can see that the product approach leads to the identical measures for direct and indirect effects from Model 26 in Hayes (2013).