ISBN
978-3-031-55550-3

Authors
Kim, Hanna
Kim, Jee-Seon

Publication Date
2024

DOI
10.1007/978-3-031-55548-0_15

Assessing the Effects of a Yearly Renewable Education Program Through Causal Mediation Analysis

Hanna Kim[1[0000-0001-8311-1045]] and Jee-Seon Kim[1[0000-0002-3392-3675]]

[1] University of Wisconsin-Madison, Madison WI 537056, USA
hanna.kim@wisc.edu

Abstract. When education programs are renewed yearly, participation in such programs can vary over time, resulting in multiple patterns of participation. One such example is the national Head Start program administered for two consecutive years, serving children aged three and four. Even though there are four possible patterns of Head Start attendance, Head Start has often been examined as a one-time event in early childhood education literature, with only the effect of early Head Start attendance at age three being evaluated. In this study, we propose to apply causal mediation analysis to the study of yearly renewable education programs, separating the effect of initial program attendance into sequential effects of the programs over time and long-term effects of initial program attendance. We adopt a parametric closed-form estimation that combines regression models to examine the effect of Head Start on children's receptive vocabulary using data from the Head Start Impact Study as an illustration. Our analysis exemplifies how the effect of a yearly renewable education program can be attributed to different program attendance histories and invites further research on studying time-varying treatment effects as causal mediation effects.

Keywords: Time-varying treatments, Causal mediation, Regression-based estimation, Exposure-mediator interaction, Head Start Impact Study.

1. Evaluating Yearly Renewable Education Programs as Time-Varying Treatments

As longitudinal studies are becoming more popular, interest is on the rise regarding statistical methods for examining how a treatment reaches its effect over time. When treatment levels vary over time, such as education programs being renewable each year, subjects may join or leave the treatment over the course of a study with various patterns or *treatment histories*. For example, the national Head Start program in early childhood education is yearly renewable, where children can attend an early Head Start program at age three to four, followed by a regular Head Start program at age four to five (i.e., pre-Kindergarten year). Over two years, a total of four Head Start attendance histories are possible: children could attend Head Start for two years, attend only the early Head Start and switch to an alternative childcare, join the regular Head Start after receiving alternative childcare, or never attend any Head Start programs. Throughout this article, we illustrate the effect of Head Start on children's school readiness as a running example.

Analyzing the effects of such renewable education programs requires methods specifically designed for time-varying treatments. If instead methods for cross-sectional studies are applied, renewable education programs are treated as one-time events. Then, only the total effect of the initial program administration is analyzed. This is problematic in that treatment trajectories after the initial program administration are disregarded and the roles of subsequent program implementations cannot be evaluated (Vandecandelaere et al., 2016).

To examine the effects of renewable education programs considering treatment histories over time, we propose to adopt causal mediation analysis and divide the total effect of the initial treatment implementation into multiple paths that relate to different sequences of treatment participation. By doing so, we may explain how the first treatment leads to successive treatments and eventually to different outcomes, which cannot be obtained by applying cross-sectional data analysis methods to longitudinal studies.

2. Disentangling Time-Varying Treatment Effects as Causal Mediation Effects

Causal mediation analysis separates the total effect of an exposure into a direct effect and an indirect effect. Direct effects capture the effects that are only attributable to the exposure and do not relate to the changes in the mediator, whereas indirect effects occur as a response to changes in the mediator caused by the exposure. By conceptualizing the effects of time-varying treatments as causal mediation[1] effects, we can explain how the total effect of an initial treatment implementation is achieved through multiple paths depending on treatment participation behavior over time, which is not addressable with a cross-sectional approach. Specifically, we may consider the initial treatment implementation (e.g., early Head Start attendance) as an exposure and the subsequent treatment implementation (e.g., regular Head Start attendance) as a mediator, followed by an outcome of interest (e.g., children’s receptive vocabulary scores). Below, we lay out how the total effect of the initial implementation of a time-varying treatment can be decomposed into natural direct and indirect effects in the context of Head Start effect analysis.

2.1 Long-term effects of the initial treatment as natural direct effects

[1] It may be argued that traditional mediation analysis combining regression coefficients can achieve similar goals without relying on causal inference. However, such mediation analyses often do not model exposure-mediator interaction or systematically adjust for confounding, making implicit assumptions about the absence of exposure-mediator interaction and confounding.

Direct effects of a time-varying treatment evaluate the impact of the initial treatment implementation not due to any changes in the subsequent implementation by conditioning on the level of the subsequent treatment. Natural direct effects (NDE) capture the effect of the initial treatment on the outcome that would remain after excluding the pathway from the initial implementation through the second implementation. This is achieved by fixing the second phase not to actual levels (e.g., 0 or 1 for a binary treatment), but instead to the natural values that individuals would take after a given initial treatment level (Nguyen et al., 2020; Pearl, 2022; VanderWeele, 2015)[2].

Formally, let $Y_{A_1=a,A_2=A_2(a')} = Y_{a,A_2(a')}$ indicate the nested potential outcome of a person with initial treatment level $A_1 = a, a \in \{0, 1\}$ and the potential level of the second treatment A_2 equal to the level that would have occurred after $A_1 = a'$.[3] Then, for a binary treatment, the NDE with regards to $A_1 = a$ is defined as

$$NDE(a) = E\left[Y_{1,A_2(a)} - Y_{0,A_2(a)} \middle| C'\right], \tag{1}$$

given a vector of covariates C'. Note that the amounts of $NDE(0)$ and $NDE(1)$ may differ when the impact of the initial treatment differs by the level of the second treatment.

In the Head Start example, the NDEs examine the difference in children's school readiness after attending early Head Start but keeping regular Head Start attendance the same level as what would have happened had children (not) attended early Head Start. There can be two distinct NDEs depending on the hypothetically fixed level of the second treatment. Specifically, regular Head Start attendance can be fixed to the natural level that children would have taken after attending early Head Start (e.g., an average regular Head Start attendance probability of 60%; $NDE(1)$), versus to the level children would have taken after not attending early Head Start (e.g., an average probability of 40%; $NDE(0)$). Then, $NDE(1)$ would describe how much children's school readiness would change by attending early Head Start if everyone would proceed to regular Head Start with a 60% chance regardless of their early Head Start attendance. $NDE(0)$ on the other hand, would describe the change in school readiness by attending early Head Start if everyone, regardless of their early Head Start attendance, would proceed to regular Head Start with a 40% chance.

2.2 Sequential effects of time-varying treatments as natural indirect effects

---

[2] Controlled direct effects, on the other hand, fix the subsequent treatment to a specific level for everyone, making it able to compare the effects of specific treatment trajectories. However, this implies enforcing the second phase of a longitudinal treatment to a uniform level, which is often unattainable and less relevant to evaluating education programs.

[3] Note that a and a' may not necessarily be identical to allow for counterfactual outcomes such as $Y_{1,A_2(0)}$ or $Y_{0,A_2(1)}$.

Natural indirect effects (NIE) are the effects of A_1 on Y after taking out NDEs, expressing the effects of A_1 on Y that operate through A_2 (Nguyen et al., 2020; Pearl, 2022; VanderWeele, 2015). Technically, the NIE expresses how the outcome would change if the initial treatment remained fixed at a specific level, but the subsequent treatment levels would change from the natural level after the absence of the initial treatment to the level following the implementation of the initial treatment. Regarding the Head Start example, the NIE represents how attending early Head Start at age three would exert an effect on children's school readiness through their changed average regular Head Start attendance probability at age four (i.e., from 40% to 60% in our previous scenario). Studying such natural indirect effects would explain how the effects of Head Start vary as a function of treatment history. Then, the NIE with regards to $A_1 = a$ is defined as

$$NIE(a) = E\left[Y_{a,A_2(1)} - Y_{a,A_2(0)} \middle| C'\right], \tag{2}$$

measuring the impact that a change in A_2 due to changes from $A_1 = 0$ to $A_1 = 1$ has on Y when A_1 would remain at level a, conditional on a vector of covariates C'.

Two things are notable with regards to the NIEs. First, for the NIEs to be nonzero for a binary treatment, $A_2(0)$ and $A_2(1)$ should be different, such that participating in the initial treatment has a nonzero effect on subsequent treatment participation, eventually impacting the outcome. NIEs do not compare exact levels of A_2 but instead their potential levels that naturally occur after different A_1 levels. Therefore, they address the impact of A_2 not by exact contrasts but by how they would behave due to A_1. Consequently, the NIE captures the idea of sequential treatment effects, or mediation.

Second, indirect effects through A_2 may differ depending on the level of A_1. For example, $NIE(1)$ illustrates the average change in children's school readiness that occurs through different natural probabilities of participation in regular Head Start if children would have *attended* early Head Start. This may differ from $NIE(0)$ that captures the change in children's school readiness due to the same level of changes in the regular Head Start attendance probability after children *do not* attend early Head Start.

From equations (1) and (2), it follows that the total effect (TE) of A_1 on Y can be decomposed in two combinations of natural direct and indirect effects. Namely,

$$\begin{aligned} TE &= E\left[Y_{1,A_2(1)} - Y_{0,A_2(0)} \middle| C'\right] \\ &= E\left[Y_{1,A_{2(1)}} - Y_{0,A_2(1)} \middle| C'\right] + E\left[Y_{0,A_2(1)} - Y_{0,A_2(0)} \middle| C'\right] \\ &= NDE(1) + NIE(0) \end{aligned} \tag{3}$$

$$\begin{aligned} TE &= E\left[Y_{1,A_2(1)} - Y_{0,A_2(0)} \middle| C'\right] \\ &= E\left[Y_{1,A_{2(1)}} - Y_{1,A_2(0)} \middle| C'\right] + E\left[Y_{1,A_2(0)} - Y_{0,A_2(0)} \middle| C'\right] \\ &= NIE(1) + NDE(0) \end{aligned} \tag{4}$$

When there exists an interaction between the early and regular Head Start attendance and their effects on children's vocabulary skills, the two decompositions will result in different indirect and direct effects depending on hypothetical conditioning.

3. Regression-Based Estimation of Causal Mediation Effects

Causal mediation analysis is characterized by clarifying the set of assumptions that are needed to identify causal estimands from observed data. In order to identify natural direct and indirect effects conditional on covariates, there should be no unmeasured confounding between A_1 and $Y_{a,A_2(a')}$, A_2 and $Y_{a,A_2(a')}$, A_1 and $A_2(a)$, and $Y_{a,A_2(a')}$ and $A_2(1-a)$[4] given measured covariates C'. Assuming our measured set of covariates cover all such confounders, a closed-form parametric estimation is made available based on regression models for estimating natural direct and indirect effects (Valeri & VanderWeele, 2013).

In the Head Start example, regular Head Start constitutes a binary mediator and children's school readiness is evaluated by measuring receptive vocabulary scores on a continuous scale. Therefore, we estimate the following regression models.

$$logit\{P(A_2 = 1 | A_1 = a, C' = c)\} = \beta_0 + \beta_1 a + \beta_2' c, \qquad (5)$$

$$E[Y | A_1 = a, A_2 = a', C' = c] = \theta_0 + \theta_1 a + \theta_2 a' + \theta_3 a * a' + \theta_4' c. \qquad (6)$$

A_1 and A_2 denote early and regular Head Start participation, respectively (measured in Spring 2003 and Spring 2004; 1 = Yes, 0 = No), Y children's receptive vocabulary score measured with a shortened version of the Peabody Picture Vocabulary Test (PPVT) towards the end of Spring 2004, and C' a vector of possible confounders, including children's demographic variables (gender, race/ethnicity, primary language, special needs, urbanicity), environmental conditions (primary caregiver's age, living with both biological parents, primary caregiver's level of depression, recent immigration of biological mother, mothers' marital status, education level, household risk level), and children's receptive vocabulary levels at the baseline in Fall 2002 (Puma et al., 2010b).

To account for the possibility that the effect of attending regular Head Start might differ depending on early Head Start experience and vice versa (Jenkins et al., 2018), interaction between the early and regular Head Start attendance (θ_3) is modeled. Then, $NDE(a)$ and $NIE(a)$ are estimated using the following closed forms:

---

[4] The last assumption rules out the possibility of multiple mediators that are causally related to each other (Tingley et al., 2014). In other words, nothing should be on the pathway from A_1 to A_2 that also affects Y.

$$NDE(a) = \{\theta_1(a - a^*)\} + \{\theta_3(a - a^*)\}\frac{\exp[\beta_0 + \beta_1 a + \beta_2' c]}{1 + \exp[\beta_0 + \beta_1 a + \beta_2' c]}, \quad (7)$$

$$NIE(a) = (\theta_2 + \theta_3 a)\left\{\frac{\exp[\beta_0 + \beta_1 a + \beta_2' c]}{1 + \exp[\beta_0 + \beta_1 a + \beta_2' c]} - \frac{\exp[\beta_0 + \beta_1 a^* + \beta_2' c]}{1 + \exp[\beta_0 + \beta_1 a^* + \beta_2' c]}\right\}, (8)$$

where $a^* = 1 - a$. It is the nonzero θ_3 (i.e., exposure-mediator interaction in the causal mediation literature) that determines the difference between $NDE(0)$ and $NDE(1)$, or between $NIE(0)$ and $NIE(1)$. Standard errors of these estimators are obtained using the delta method or bootstrapping. For a comparison of estimation methods other than regression-based estimators, we refer interested readers to Park et al. (2023).

4. Empirical Example: The Effects of Head Start on Children's School Readiness by Attendance History

4.1 Data and analysis

To illustrate how the causal mediation analysis discussed in the previous section can be applied to investigate the effects of early and regular Head Start on children's receptive vocabulary as one aspect of school readiness, we use data from the 3-year-old cohort of the Head Start Impact Study (HSIS). The HSIS was conducted to evaluate the effectiveness of Head Start as a nation-wide early childhood education program by collecting a nationally representative data set of children and families who were eligible to enroll in Head Start programs with limited capacity (Puma et al., 2010a).[5] Then, natural direct and indirect effects were estimated within R (R Core Team, 2022) using the package `regmedint` (Li et al., 2023) to synthesize estimates and standard errors across 20 imputed data sets.[6]

4.2 Results

Estimated coefficients are presented in Tables 1 and 2 for the mediator and outcome models, respectively, followed by derived total effect, natural direct and indirect effects in Table 3.

---

[5] Restricted data access was granted to the authors for secondary research on the HSIS data, which received an exemption from the University of Wisconsin-Madison Institutional Review Board. Description of the variables in the HSIS used for imputing missing values with multiple imputation (Little & Rubin, 2019; van Buuren & Groothuis-Oudshoorn, 2011), as well as description of the imputed variables are available upon request.

[6] Codes for preparing the HSIS data and applying analyses are available upon request.

Table 1. Coefficient estimates of the regular Head Start model ($n = 2{,}449$)

Variable	Coefficient	Odds	SE	p
Intercept	-0.11	0.90	0.42	0.80
Early Head Start	1.23	3.42	0.09	<0.01
Female	-0.08	0.92	0.09	0.34
Black	-0.39	0.68	0.12	<0.01
Hispanic	-0.04	0.96	0.14	0.74
Receptive vocabulary score at baseline	0.00	1.00	0.00	0.24
Special educational needs	0.21	1.24	0.14	0.13
Speaks English at home	0.19	1.21	0.16	0.24
Caregiver age	0.00	1.00	0.01	0.73
Living with both biological parents	-0.01	0.99	0.12	0.91
Caregiver depression index	0.10	1.10	0.05	0.03
Biological mother is a recent immigrant	-0.23	0.80	0.16	0.15
Mother's education: less than High School	0.27	1.31	0.12	0.02
Mother's education: High School	0.01	1.02	0.11	0.89
Mother never married	0.29	1.33	0.14	0.04
Mother is currently married	0.08	1.09	0.16	0.59
Household risk index	-0.09	0.91	0.08	0.28
urban	0.12	1.13	0.12	0.29

Table 2. Coefficient estimates of the receptive vocabulary score model ($n = 2{,}449$)

Variable	Coefficient	SE	p
Intercept	226.87	5.50	<0.01
Early Head Start	6.64	1.82	<0.01
Regular Head Start	-0.75	1.76	0.67
Early × Regular Head Start	-4.97	2.42	0.04
Female	2.68	1.14	0.02
Black	-21.81	1.54	<0.01
Hispanic	-14.84	1.76	<0.01
Receptive vocabulary score at baseline	0.38	0.02	<0.01
Special educational needs	-4.57	1.78	0.01
Speaks English at home	-19.81	2.11	<0.01
Caregiver age	0.22	0.08	<0.01
Living with both biological parents	-0.65	1.56	0.68
Caregiver depression index	1.27	0.60	0.04
Biological mother is a recent immigrant	-12.14	2.03	<0.01
Mother's education: less than High School	-10.68	1.52	<0.01
Mother's education: High School	-4.96	1.42	<0.01
Mother never married	-0.48	1.82	0.79
Mother is currently married	-0.31	2.06	0.88
Household risk index	-2.84	1.09	0.01
urban	2.80	1.55	0.07

Table 3. Causal mediation estimates of the Head Start example ($n = 2{,}449$)

Estimand	Estimate	SE	95% CI	
TE	3.27	1.34	[0.63,	5.91]
$NDE(0)$	4.86	1.43	[2.05,	7.66]
$NIE(1)$	-1.59	0.59	[-2.75,	-0.42]
$NDE(1)$	3.45	1.52	[0.45,	6.44]
$NIE(0)$	-0.18	0.62	[-1.40,	1.04]

The total effect estimate indicates that attending early Head Start has an overall positive effect on boosting children's receptive vocabulary as they near the end of their pre-K year ($\widehat{TE} = 3.27, 95\%\ CI = [0.63, 5.91]$). This is what we can learn from analyzing renewable education programs by applying methods for studying cross-sectional data. Decomposing this total effect into causal mediation effects reveals the following information in addition to an overall benefit of attending early Head Start.

First, most of the positive total effect comes from direct effects, indicating the effectiveness of early Head Start regardless of regular Head Start. Even though the point estimate was greater for $\widehat{NDE}(0)$ $(4.86, 95\%\ CI = [2.05, 7.66])$ than $\widehat{NDE}(1)$ $(3.45,\ 95\%\ CI = [0.45, 6.44])$, both natural direct effects were comparable with 95% confidence, implying that the amount of natural direct effects would not be meaningfully different depending on regular Head Start attendance across US children eligible for Head Start.

Considering that children who attended early Head Start were about 3.42 times ($\exp(\hat{\beta}_1) = \exp(1.23) = 3.42, p < .05$; See Table 1) more likely to attend regular Head Start than children who did not attend early Head Start but otherwise shared similar characteristics, negative $\widehat{NIE}(1)(-1.59, 95\%\ CI = [-2.75, -0.42])$ suggests that an increased likelihood of regular Head Start attendance after having participated in early Head Start worsens children's performance in receptive vocabulary. Even though the estimate of $NIE(0)$ was not statistically significant, the negative trend implies that children may not benefit from attending regular Head Start, even if they may not have attended any Head Start at age three. The difference in the $NIE(1)$ and $NIE(0)$ estimates coincide with a statistically significant interaction between early and regular Head Start attendance in terms of their effects on children's receptive vocabulary.

Inferring from the size and direction of the causal mediation effect estimates, it would be advised to administer early Head Start instead of regular Head Start, because children benefit from attending early Head Start and do not experience any positive effects through increased participation in regular Head Start. This partly agrees with findings that suggest no significant enough benefit to continuing Head Start for two years (Jenkins et al., 2018).

5. Discussion

This article focused on the variability in treatment histories of subjects being administered time-varying treatments, and demonstrated how the effects of yearly renewable education programs can be intuitively addressed with causal mediation analysis. The decomposition of the total effect of an initial treatment into natural direct and indirect effects has been discussed in relation to long-term effects and sequential effects, and illustrated with a real data analysis on the early and regular Head Start program's effect on boosting children's school readiness.

The approach of connecting sequential treatment participation and causal mediation analysis may be further developed in the following directions. First, we may compare how different groups of participants perform in terms of long-term or sequential treatment effects. This can add to the study of treatment effect heterogeneity in that the variation in individual treatment effects can be explained longitudinally, and that equitable program effects can be ensured by designing targeted solutions for distinct groups of participants. To this end, direct and indirect effects, in addition to the total effect, need to be investigated in terms of groupwise effect modification.

Second, further estimation methods need to be developed for flexible and efficient causal mediation effect estimation with time-varying treatments. In addition to the parametric closed forms introduced in this article, we may explore more flexible estimation models such as modeling any nonlinearity or effect moderation by key covariates regarding the effects of the first and second treatment administrations.

Finally, sensitivity analysis procedures need to be studied in combination with the study design presented in this article. Even though the two-phase treatment design is relatively simple and comes with initial randomization, there are already multiple assumptions attached to making causal inference from such studies. Thus, providing sensitivity analysis routines is an area deserving more attention to complete any development in causal inference methods.

References

Jenkins, J. M., Sabol, T. J., & Farkas, G. (2018). Double down or switch it up: Should low-income children stay in Head Start for 2 years or switch programs? *Evaluation Review*, *42*(3), 283–317. https://doi.org/10.1177/0193841X18786591

Li, Y., Mathur, M. B., Solomon, D. H., Ridker, P. M., Glynn, R. J., & Yoshida, K. (2023). Effect measure modification by covariates in mediation: Extending regression-based causal mediation analysis. *Epidemiology*, *34*(5), 661–672. https://doi.org/10.1097/EDE.0000000000001643

Little, R., & Rubin, D. B. (2019). *Statistical Analysis with Missing Data*. Wiley.

Nguyen, T. Q., Schmid, I., & Stuart, E. A. (2020). Clarifying causal mediation analysis for the applied researcher: Defining effects based on what we want to learn. *Psychological Methods*, *26*(2), 255–271. https://doi.org/10.1037/met0000299

Park, S., Kang, S., & Lee, C. (2023). Choosing an optimal method for causal decomposition analysis with continuous outcomes: A review and simulation study. *Sociological Methodology*. https://doi.org/10.1177/00811750231183711

Pearl, J. (2022). Direct and Indirect Effects. In *Probabilistic and causal inference: the works of Judea Pearl* (pp. 373–392). https://doi.org/10.1145/3501714.3501736

Puma, M., Bell, S., Cook, R., Heid, C., Shapiro, G., Broene, P., Jenkins, F., Fletcher, P., Quinn, L., Friedman, J., Ciarico, J., Rohacek, M., Adams, G., & Spier, E. (2010a). *Head Start Impact Study. Final Report*.

Puma, M., Bell, S., Cook, R., Heid, C., Shapiro, G., Broene, P., Jenkins, F., Fletcher, P., Quinn, L., Friedman, J., Ciarico, J., Rohacek, M., Adams, G., & Spier, E. (2010b). *Head Start Impact Study. Technical Report*.

R Core Team. (2022). *R: A Language and Environment for Statistical Computing*. https://www.r-project.org/

Tingley, D., Yamamoto, T., Hirose, K., Keele, L., & Imai, K. (2014). mediation: R Package for Causal Mediation Analysis. *Journal of Statistical Software*, *59*(5), 1–38. http://www.jstatsoft.org/

Valeri, L., & VanderWeele, T. J. (2013). Mediation analysis allowing for exposure-mediator interactions and causal interpretation: Theoretical assumptions and implementation with SAS and SPSS macros. *Psychological Methods*, *18*(2), 137–150. https://doi.org/10.1037/a0031034

van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, *45*(3), 1–67. https://doi.org/10.18637/jss.v045.i03

Vandecandelaere, M., Vansteelandt, S., De Fraine, B., & Van Damme, J. (2016). Time-Varying treatments in observational studies: Marginal structural models of the effects of early grade retention on math achievement. *Multivariate Behavioral Research*, *51*(6), 843–864. https://doi.org/10.1080/00273171.2016.1155146

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