



Published in final edited form as:

Ann Epidemiol. 2022 March ; 67: 73–80. doi:10.1016/j.annepidem.2021.12.010.

Estimating sibling spillover effects with unobserved confounding using gain-scores

David C. Mallinson^{*,1}, Felix Elwert^{2,3,4}

¹Population Research Center, University of Texas-Austin, United States

²Department of Sociology, College of Letters and Sciences, University of Wisconsin-Madison, United States

³Department of Biostatistics and Medical Informatics, School of Medicine and Public Health, University of Wisconsin-Madison, United States

⁴Center for Demography and Ecology, University of Wisconsin-Madison, United States

Abstract

Purpose—A growing area of research in epidemiology is the identification of health-related sibling spillover effects, or the effect of one individual’s exposure on their sibling’s outcome. The health within families may be inextricably confounded by unobserved factors, rendering identification of sibling spillovers challenging.

Methods—We demonstrate a gain-score regression method for identifying exposure-to-outcome spillover effects within sibling pairs in linear models. The method can identify the exposure-to-outcome spillover effect if only one sibling’s exposure affects the other’s outcome; and it identifies the difference between the spillover effects if both siblings’ exposures affect the others’ outcomes. The method fails with outcome-to-exposure spillover or outcome-to-outcome spillover. Analytic results, Monte Carlo simulations, and a brief application demonstrate the method and its limitations.

Results—We estimate the spillover effect of a child’s preterm birth on an older sibling’s literacy skills, measured by the Phonological Awareness Literacy Screening-Kindergarten test. We analyze 20,010 sibling pairs from a population-wide, Wisconsin-based (United States) birth cohort. Without covariate adjustment, we estimate that preterm birth modestly decreases an older sibling’s test score.

Conclusion—Gain-scores are a promising strategy for identifying exposure-to-outcome spillover effects in sibling pairs while controlling for sibling-invariant unobserved confounding.

*Corresponding Author 2.406B Patton Hall, 305 E 23rd Street, Austin, TX 78712, david.mallinson@austin.utexas.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

CONFLICT OF INTEREST

None declared.

DeclarationStatement

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests

Keywords

Causality; Epidemiologic Methods; Family; Siblings

INTRODUCTION

A sibling spillover effect (i.e., “interference,” “carryover effect”) is the effect of an individual’s exposure on their sibling’s outcome [1-3]. The past two decades of epidemiologic research witnessed a burgeoning interest in the role of family environments in childhood health, calling attention to the importance of spillovers within families [4-11]. Yet, sibling spillovers are largely unexamined in the epidemiologic literature, as most field-specific advancements in spillover identification have been restricted to infectious diseases [12-22]. With growing interest in the familial interdependence of health [5-11], the need for analytical tools to identify sibling spillovers is acute.

Unobserved confounding is particularly salient with sibling spillovers. Siblings often share family and social environments that affect their outcomes but remain unmeasured even in data-rich contexts [10]. Fixed effect (FE) designs that control for unobserved time-invariant confounding are immediately appealing [3, 23, 24], but there is little precedent for their use to identify sibling spillovers. Sjölander et al. (2016) investigated two-sibling FE models for identifying an individual’s exposure on their own outcome in the presence of spillover, noting that spillover may be identifiable [3]. Black et al.’s (2021) difference-in-differences method with three-sibling clusters measured a lower-bound estimate of the effect of a child’s disability on an older sibling’s academic performance [25].

In this paper, we demonstrate a method for identifying spillovers in sibling pairs with gain-scores (i.e., “change scores,” “difference scores”), a staple of FE estimation that removes shared confounding by differencing outcomes [26, 27]. Consistent with applied FE studies, we focus on linear models with homogenous effects [23, 24, 26, 27]. First, we briefly introduce causal directed acyclic graphs (DAGs), which illustrate our models. Second, we discuss various two-sibling models with one- or two-sided spillover and explain how and when gain-score methods identify spillover effects. Third, we illustrate our results with simulations. Fourth, we apply the method to identify the effect of a younger sibling’s preterm birth on an older sibling’s literacy test performance.

CAUSAL DIRECTED ACYCLIC GRAPHS

Causal DAGs are useful for explaining the identification of causal effects. We review necessary terminology for this exposition. Causal DAGs are diagrams consisting of *nodes* (variables) and *directed edges* (direct causal effects) that represent the assumed data-generating process (causal model) [28-33]. *Paths* are sequences of adjacent edges, regardless of the arrows’ directions. On *causal paths* between exposure and outcome, all arrows point from the exposure to the outcome. On *non-causal paths* between an exposure and an outcome, at least one arrow points away from the outcome. Causal paths “transmit” causal effects, whereas non-causal paths may transmit spurious associations. *Colliders* are variables that receive two inbound arrows on a path (a given variable may be a collider

on one path but not on another). Pearl's *d-separation* criterion determines which variables in data generated by the assumed DAG are conditionally or unconditionally independent: two variables are independent if all paths between them are closed; and a path is closed if it includes a conditioned noncollider variable, or if it includes an *unconditioned* collider variable [29, 30, 33]. Conversely, two variables may be associated if at least one path between them is open (*d-connected*); and a path is open if it is not closed.

Typically, health researchers attempt to identify causal effects by conditioning on observed variables through adjustment via regression analysis, matching, or inverse-probability weighting to close all non-causal paths between the exposure and outcome [29, 30, 33]. Unfortunately, open non-causal paths containing only unobserved intermediate variables cannot be closed by covariate adjustment. In multilevel analyses in which observations are clustered into groups (e.g., children in families), however, FE methods can sometimes identify causal effects by subtracting out certain types of group-level unobserved confounding [23-27]. Next, we describe several sibling spillover models with unobserved confounding and show when gain-score estimation can identify spillover effects.

METHOD FOR SIBLING SPILLOVER IDENTIFICATION

Model and assumptions

We present our baseline sibling spillover model and subsequently introduce variations on this model. For illustration, we discuss the spillover of a child's early health shock on their sibling's academic outcome. This example is purposefully generic but broadly applicable, drawing upon prior work of health-related spillover effects on academic performance [25] while motivating our empirical application.

Our baseline model is a linear two-sibling comparison design with one-sided spillover (Figure 1A). Subscript $i = 1, \dots, N$ indicates family and subscript $j = 1, 2$ indicates sibling. T_{ij} represents a binary or continuous exposure (e.g., the health shock), Y_{ij} represents a continuous outcome (e.g., academic performance), U_i represents unobserved family-level confounding (i.e., the FE), and D_i represents the gain-score, $D_i = Y_{i2} - Y_{i1}$. Parameters include the spillover effect, $\theta(T_{i1} \rightarrow Y_{i2})$, of sibling 1's exposure on sibling 2's outcome; the targeted effect, $\delta(T_{ij} \rightarrow Y_{ij})$, of an individual's exposure on their own outcome; and unobserved family-level confounding effects $\chi(U_i \rightarrow T_{i1})$, $\gamma(U_i \rightarrow T_{i2})$, and $\Psi(U_i \rightarrow Y_{ij})$.

All models embed several simplifying assumptions. First, targeted and unobserved family-level confounding effects on outcomes are sibling-invariant [23, 24]. Second, all effects are linear and homogenous. Third, spillovers exist *within* families but not *between* families (i.e., partial interference) [15, 34]. Aside from permitting partial interference, these assumptions align with conventional FE models [3, 23, 24, 26, 27].

Notably, our presentation abstracts from sibling-specific baseline covariates, C_{ij} . Covariates may be added to our baseline and subsequent models as long as controlling for C_{ij} in the gain-score regression does not induce new biases [35-38].

Gain-score estimation and identification of spillover effects

This subsection analyzes whether the gain-score estimator point-identifies spillover effects (i.e., recovers the spillover effect precisely) for nine different sibling spillover models. that differ by whether spillover is one- or two-sided and by whether additional spillovers are present.

Gain-score estimation—We investigate the ability of a gain-score estimator to identify exposure-to-outcome spillover effects. First, we regress the gain-score on both siblings' exposures,

$$D_i = b_1 T_{i1} + b_2 T_{i2} + e_i, \quad (1)$$

where b_1 and b_2 are partial regression coefficients for T_{i1} and T_{i2} , respectively, and e_i is the residual. We then sum the partial regression coefficients to compute a “spillover coefficient” (SC),

$$SC = b_1 + b_2. \quad (2)$$

We will now interrogate whether the SC identifies causal spillover effects in each of several commonly assumed data generating processes in health research.

Settings with one-sided spillover—The object of interest (estimand) is the direct spillover effect, θ , of sibling 1's health shock, T_{i1} , on sibling 2's academic performance, Y_{i2} . Under the baseline model (Figure 1A), three open paths connect T_{i1} and Y_{i2} . The first path, $T_{i1} \rightarrow Y_{i2}$, is the causal spillover effect of interest. The other two paths are non-causal paths that may transmit spurious association. The first non-causal path, $T_{i1} \leftarrow U_i \rightarrow T_{i2} \rightarrow Y_{i2}$, can be closed by adjusting for T_{i2} . However, the second non-causal path, $T_{i1} \leftarrow U_i \rightarrow Y_{i2}$, cannot be closed by covariate adjustment because it only contains the unobserved variable U_i .

Nonetheless, we can identify θ through gain-score regression. Under the assumptions of Figure 1A, it can be shown that $b_1 = \theta - \delta$ and $b_2 = \delta$, using elementary regression algebra. Therefore, the spillover coefficient equals $SC = b_1 + b_2 = \theta$.

The intuition for this result is that first-differencing exactly offsets confounding biases involving U_i [26, 27], and that the SC corrects for the contamination of the spillover estimate in b_1 . Specifically, the coefficient b_1 on T_{i1} captures the association flowing along the two open paths from T_{i1} to D_i . There are five paths from T_{i1} to D_i (listed together with their corresponding path coefficients):

1. $T_{i1} \leftarrow U_i \rightarrow T_{i2} \rightarrow Y_{i2} \rightarrow D_i$ (non-causal): $\delta\chi\gamma$
2. $T_{i1} \leftarrow U_i \rightarrow Y_{i1} \rightarrow D_i$ (non-causal): $-\Psi\chi$
3. $T_{i1} \leftarrow U_i \rightarrow Y_{i2} \rightarrow D_i$ (non-causal): $\Psi\chi$
4. $T_{i1} \rightarrow Y_{i1} \rightarrow D_i$ (causal): $-\delta$

5. $T_{i1} \rightarrow Y_{i2} \rightarrow D_i$ (causal): θ

The first path is closed because the gain-score regression adjusts for T_{i2} . The second and third paths cancel each other out exactly. The fourth path transmits the negative of the targeted effect. The fifth path transmits the spillover effect. Hence, the regression coefficient $b_1 = \theta - \delta$ identifies the difference between the spillover and targeted effect.

The coefficient b_2 on T_{i2} captures the association flowing along the open paths from T_{i2} and D_i . There are four paths from T_{i2} to D_i :

1. $T_{i2} \leftarrow U_i \rightarrow Y_{i1} \rightarrow D_i$ (non-causal): $-\delta\chi\gamma$
2. $T_{i2} \leftarrow U_i \rightarrow Y_{i1} \rightarrow D_i$ (non-causal): $-\Psi\gamma$
3. $T_{i2} \leftarrow U_i \rightarrow Y_{i2} \rightarrow D_i$ (non-causal): $\Psi\gamma$
4. $T_{i2} \rightarrow Y_{i2} \rightarrow D_i$ (causal): δ

The first path is closed because the gain-score regression adjusts for T_{i1} ; the second and third paths cancel each other out; and the fourth path captures the targeted effect. Thus, $b_2 = \delta$ identifies the targeted effect, and $SC = b_1 + b_2 = \theta$ identifies the causal spillover effect.

Many statistical software packages have functions for summing regression coefficients and obtaining standard errors. Examples include Stata's `lincom` command, R's `contrast` package, and SAS's `SCORE` procedure [39-41].

The analysis is only slightly complicated in the presence of exposure-to-exposure spillover ($T_{ij} \rightarrow T_{ij'}$)—for example, when one child's serious illness increases their sibling's risk of illness. When $T_{i2} \rightarrow T_{i1}$ (Figure 1B), the interpretation of the SC does not change. However, if $T_{i1} \rightarrow T_{i2}$ (Figure 1C), then the interpretation of $SC = \theta$ changes from representing the entire spillover effect of T_{i1} on Y_{i2} to capturing only the direct spillover effect, since the indirect component of the spillover effect that operates via the causal path $T_{i1} \rightarrow T_{i2} \rightarrow Y_{i2}$ is closed because the regression controls for T_{i2} . See the Appendix for details.

Settings with two-sided spillover—Analysts may also encounter scenarios with two-sided spillover. In our example, each sibling's health shock could affect the other's academic performance ($T_{i1} \rightarrow Y_{i2}$ and $T_{i2} \rightarrow Y_{i1}$). Reflecting this possibility, Figure 2A modifies the baseline model of Figure 1A to allow spillover $T_{i2} \rightarrow Y_{i1}$ with effect κ . The partial regression coefficients in the gain-score approach identify $b_1 = \theta - \delta$ and $b_2 = \delta - \kappa$, so that $SC = b_1 + b_2 = \theta - \kappa$. Consequently, with two-sided exposure-to-outcome spillover, the SC does not identify the spillover effect of T_{i1} on Y_{i2} but instead the difference between the two exposure-to-outcome spillover effects. However, if the analyst can defend assumptions about one or more of the signs of the two spillover effects, then the SC remains informative even though it no longer point-identifies θ . Specifically, if $\kappa > 0$, the SC underestimates (i.e., gives a lower bound for) θ . By contrast, if $\kappa < 0$, then SC overestimates (gives an upper bound for) θ . One can make additional inferences about θ depending on the value of SC and the assumed sign of κ . For example, if $SC > 0$ and $\kappa > 0$, then $\theta > 0$. Of note, a finding that $SC = 0$ is uninformative, because it is compatible with the possibility that the two spillover effects are equal ($\theta = \kappa = 0$) or that exposure-to-outcome spillovers are absent ($\theta = \kappa = 0$).

If $T_{ij} \rightarrow T_{ij}'$ in addition to two-sided spillover (Figures 2B-C), then *SC* identifies the difference between the direct parts of the siblings' spillover effects that are not mediated by T_{i1} or T_{i2} , respectively, i.e., *SC* will not consider the indirect spillover effects, $T_{ij} \rightarrow T_{ij}' \rightarrow Y_{ij}'$. See the Appendix for details.

Settings with spillovers from outcomes—Analysts may also encounter settings with outcome-to-outcome spillover ($Y_{ij} \rightarrow Y_{ij}'$) or outcome-to-exposure spillover ($Y_{ij} \rightarrow T_{ij}'$). In our example, siblings' academic outcomes may affect each other via outcome-to-outcome spillover. In contrast, an academic outcome causing a health shock is implausible, but outcome-to-exposure spillovers may be relevant elsewhere.

If one sibling's outcome causes the other sibling's exposure or outcome (Figure 3), then our gain-score approach does not identify spillovers or simple functions of spillovers. See the Appendix for details.

SIMULATION

Nine Monte Carlo simulations [42]—one simulation for each model in Figures 1-3—demonstrated when the method identifies exposure-to-outcome spillover. Our simulation equations are¹:

$$U_i, v_{i1}, v_{i2} \sim N(0, 1)$$

$$T_{i1} = \begin{cases} 0 & \text{if } \tau T_{i2} + \chi U_i \leq 0.5 \\ 1 & \text{if } \tau T_{i2} + \chi U_i > 0.5 \end{cases}$$

$$T_{i2} = \begin{cases} 0 & \text{if } \phi T_{i1} + \omega Y_{i1} + \gamma U_i \leq 0.2 \\ 1 & \text{if } \phi T_{i1} + \omega Y_{i1} + \gamma U_i > 0.2 \end{cases}$$

$$Y_{i1} = \delta T_{i1} + \kappa T_{i2} + \lambda Y_{i2} + \psi U_i + v_{i1}$$

$$Y_{i2} = \delta T_{i2} + \theta T_{i1} + \eta Y_{i1} + \psi U_i + v_{i2}$$

$$D_i = Y_{i2} - Y_{i1}$$

We simulated each model with 1000 runs of 5000 observations each, where each observation represented a sibling pair (i.e., family). We set the following parameters at fixed values: $\theta =$

¹When exposures T_{ij} are binary, the equations for T_{ij} are nonlinear; the causal effects on T_{ij} in the models of Figures 1-3 should then be read as the coefficients of linear projections onto T_{ij} . This does not affect the identification results of our gain-score approach [26].

0.5, $\delta = 1$, $\psi = 1$, $\chi = 2$, and $\gamma = 3$. Parameters that distinguish models— κ , τ , ϕ , ω , η , and λ —were set to zero in models where absent and were set to 0.3 in models where present. To avoid simultaneity, at least one parameter in each pair (τ , ϕ), (η , λ), and (κ , ω) was always set to zero. In each sample, we regressed the gain-score on siblings' outcomes and computed the spillover coefficient according to equations (1) and (2). We conducted simulations in Stata Statistical Software: Release 16 [43]. Simulation code is in the Appendix.

Figure 4 displays the simulation results. The first three rows confirm that the spillover coefficient is unbiased in the three settings with one-sided exposure-to-outcome spillover of Figure 1, as the average of estimated spillover coefficient equals the known spillover effect, $\widehat{SC}_{Figure1} = 0.5$ (empirical 95% CI: 0.42, 0.58). The subsequent three rows demonstrate that the spillover coefficient in the three models of Figure 2 with two-sided exposure-to-outcome spillover identifies the difference between the two spillovers, $\widehat{SC}_{Figure2} = 0.5 - 0.3 = 0.2$ (empirical 95% CI: 0.12, 0.28). Since $\kappa > 0$, $\widehat{SC}_{Figure2}$ underestimates the spillover effect, θ . The final three simulations show that $\widehat{SC}_{Figure3}$ is biased in all models of Figure 3 with spillovers from outcomes. Size and direction of the biases are complicated functions of the coefficients in the data-generating model and can be large. The estimated ordinary least squares standard errors closely resemble the empirical standard errors for each model, indicating that the built-in standard errors in Stata's `lincom` command are accurate [39].

EMPIRICAL APPLICATION

We applied the method to estimating the spillover effect of a child's preterm birth (gestational age <37 weeks) on their older sibling's literacy skills. This analysis builds upon evidence that short gestational age and other health shocks may harm children's academic outcomes [25, 44, 45]. If a child is born preterm, parents may reallocate investments (time, financial, or otherwise) from older siblings to support the younger sibling's health, thereby inhibiting the older siblings' development, including early literacy.

For this application, we analysed Big Data for Little Kids (BD4LK), a longitudinal cohort of birth records for all live in-state resident deliveries in Wisconsin during 2007-2016 ($N > 660,000$ deliveries) that links to multiple administrative data sources, including Medicaid data (2007-2016) and children's Phonological Awareness Literacy Screening-Kindergarten (PALS-K) test scores from Wisconsin public schools (2012-2016 school years). BD4LK's linking process is described elsewhere [45, 46]. PALS-K evaluates readiness for kindergarten-level literacy instruction on six domains (rhyme awareness; beginning sound awareness; alphabet knowledge; letter sounds; spelling; word concept) [46]. In Wisconsin, children must be five years-old at kindergarten enrolment to qualify for PALS-K testing [48]. Our analysis includes 20,010 sibling pairs (40,020 children) that were sequentially-born from different deliveries to the same biological mother and had non-missing English-language PALS-K test scores and covariates. The Appendix contains the full sampling description.

We estimate the following gain-score regression,

$$D_i = b_1 PTB_{i1} + b_2 PTB_{i2} + \beta_3 C_{i2} + v_i$$

where $D_i = PALS_{K_{i2}} - PALS_{K_{i1}}$ is the gain-score. Subscripts $i = 1, \dots, N$ and $j = 1, 2$ indicate family and sibling, respectively, where $j = 1$ is the *younger* sibling. PTB_{ij} is a binary preterm birth indicator (1 if preterm; 0 otherwise), $PALS_{K_{ij}}$ is the continuous PALS-K score (0-102 points), and C_{i2} is a vector of covariates measured at the older sibling's delivery. Covariates include maternal age (years), maternal education (no high school diploma; high school diploma/equivalent; 1-3 years college; 4+ years college), and Medicaid delivery payment (no; yes).

We ran the regression twice, once with and once without covariates, and then computed the spillover coefficient, $SC = b_1 + b_2$. Assuming the one-sided spillover model of Figure 1A, SC from the regression without covariates identifies the effect of a younger sibling's preterm birth on the older sibling's PALS-K score. Additionally, b_2 identifies the effect of each sibling's preterm birth on their own PALS-K score. We performed all analyses in Stata Statistical Software: Release 16 [43]. The University of Wisconsin-Madison minimal risk institutional review board approved our project.

Tables A.1 and A.2 summarize baseline characteristics of our sample (Appendix). Preterm birth incidence was slightly greater among older siblings relative to younger siblings (6.78% vs. 6.65%). On average, older siblings received slightly lower PALS-K scores (mean 63.58 points; SD 24.12 points) relative to younger siblings (mean 64.22 points; SD 23.83 points). Approximately 10% of observed families had discordant preterm birth exposures. In the regression without covariate adjustment, the older sibling's preterm birth coefficient was $\hat{b}_2 = -2.49$ points (95% CI $-3.83, -1.15$ points), the younger sibling's preterm birth coefficient was $\hat{b}_1 = 0.38$ points (95% CI: $-0.97, 1.73$ points), and the resulting \widehat{SC} was -2.11 points (95% CI: $-3.82, -0.40$ points) (Table 1). This indicates that a younger sibling's preterm birth modestly harmed their older sibling's PALS-K performance. Figure 5 displays these results graphically relative to the assumed data-generating model. However, covariate adjustment attenuated the \widehat{SC} to -1.49 points (95% CI $-3.21, 0.22$ points).

DISCUSSION

We described a simple gain-score approach for identifying spillovers in linear models of sibling pairs. This method can point-identify spillovers if only one sibling's exposure affects the other's outcome, and it can identify the difference in siblings' spillovers in the presence of two-sided spillover. The method leverages the primary benefit of FE estimation: controlling for family-level, sibling-invariant, unobserved confounding. Whereas preceding epidemiologic research on spillover identification primarily considered infectious diseases, our work contributes to the growing literature on spillovers within families.

We acknowledge some limitations. First, we restricted our attention to linear models. This method does not necessarily apply to other settings, such as models with binary outcomes (see Sjölander et al. [2016] for binary outcomes in our Figure 1A [3]). Second, we accepted conventional FE assumptions, which include equal effects of the unobservables on siblings'

outcomes and a constant spillover effect that does not vary by characteristics, such as siblings' age difference. Deviation from these assumptions may induce bias [3, 37, 38]. Third, we did not consider families of three or more siblings. Spillovers that originate from larger sibships may pose unique challenges that are unaddressed here—for example, whether one can identify the effect of a middle child's exposure on the youngest sibling's outcome if an eldest sibling's exposure affects all siblings' outcomes. Lastly, we did not investigate spillover in the presence of shared mediator or collider variables. Sjölander and Zetterqvist (2017) interrogated sibling comparison models with shared mediators and colliders, finding that such factors may induce bias [49].

Nonetheless, our paper lays groundwork for subsequent research. Specific avenues that advance this method include testing in nonlinear settings or settings with shared mediator variables, expanding models to allow three or more siblings, and developing tests for assessing bias from outcome-induced spillovers.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

We thank the Wisconsin Department of Children and Families, Department of Health Services, and Department of Public Instruction for the use of data. We also thank Steven T. Cook, Dan Ross, Jane A. Smith, Kristen Voskuil, and Lynn Wimer for data access and programming assistance. We thank Michael Sobel for methodological discussions, Sneha Kumar for verifying simulation results, and Deborah B. Ehrental, John Mullahy, and Paul E. Peppard for feedback on this manuscript.

FUNDING

This work was supported by the following sources: the Eunice Kennedy Shriver National Institute for Child Health and Human Development through the Population Research Center at the University of Texas-Austin (P2C HD042849) and through the Center for Demography and Ecology at the University of Wisconsin-Madison (T32 HD007014-42; P2C HD047873); the University of Wisconsin-Madison Clinical and Translational Science Award program through the National Institutes of Health National Center for Advancing Translational Sciences (UL1TR00427); the National Institute on Aging through the Center for Demography of Health and Aging at the University of Wisconsin-Madison (P30 AG017266); the H.I. Romnes Faculty Fellowship through the University of Wisconsin-Madison; the University of Wisconsin-Madison School of Medicine and Public Health's Wisconsin Partnership Program; and the University of Wisconsin-Madison Institute for Research on Poverty. The content is solely the responsibility of the authors and does not necessarily represent the official views of supporting agencies. Supporting agencies do not certify the accuracy of the analyses presented.

ABBREVIATIONS

| | |
|---------------|--|
| BD4LK | Big Data for Little Kids |
| DAG | directed acyclic graph |
| FE | fixed effects |
| PALS-K | Phonological Literacy Awareness Screening-Kindergarten |

REFERENCES

1. Ogburn EL, VanderWeele TJ. Causal diagrams for interference. *Stat Sci.* 2014;29:559–578.

2. VanderWeele TJ. *Explanation in Causal Inference: Methods for Mediation and Interaction*. Oxford, UK: Oxford University Press, 2015.
3. Sjölander A, Frisell T, Kuja-Halkola R, Öberg S, Zetterqvist J. Carryover effects in sibling comparison designs. *Epidemiol.* 2016;27:852–858.
4. Kuh D, Ben-Shlomo Y, Lynch J, Hallqvist J, Power C. Life course epidemiology. *J Epidemiol Community Health.* 2003;57:778–783. [PubMed: 14573579]
5. Lawlor DA, Mishra GD. *Family Matters: Designing, Analysing and Understanding Family Based Studies in Life Course Epidemiology*. 1st edn. Oxford, UK: Oxford University Press, 2009.
6. Liu S, Jones RN, Glymour MM. Implications of lifecourse epidemiology for research on determinants of adult disease. *Public Health Rev.* 2010;32:489–511. [PubMed: 24639598]
7. Feinberg ME, Solmeyer AR, McHale SM. The third rail of family systems: sibling relationships, mental and behavioral health, and preventive intervention in childhood and adolescence. *Clin Child Fam Psychol Rev.* 2012;15:43–57. [PubMed: 22105663]
8. Viner RM, Ross D, Hardy R, et al. Life course epidemiology: recognising the importance of adolescence. *J Epidemiol Community Health.* 2015;69:719–720. [PubMed: 25646208]
9. Ben-Shlomo Y, Cooper R, Kuh D. The last two decades of life course epidemiology, and its relevance for research on ageing. *Int J Epidemiol.* 2016;45:973–988. [PubMed: 27880685]
10. De Neve JW, Kawachi I. Spillovers between siblings and from offspring to parents are understudied: a review and future directions for research. *Soc Sci Med.* 2017;183:56–61. [PubMed: 28478353]
11. Morris AS, Robinson LR, Hays-Grudo J, Claussen AH, Hartwig SA, Treat AE. Targeting parenting in early childhood: a public health approach to improve outcomes for children living in poverty. *Child Dev.* 2017;88:388–397. [PubMed: 28138978]
12. Halloran ME, Struchiner CJ. Study designs for dependent happenings. *Epidemiol.* 1991;2:331–338.
13. Halloran ME, Struchiner CJ. Causal inference in infectious diseases. *Epidemiol.* 1995;6:142–151.
14. Longini IM, Sagatelian K, Rida WN, Halloran ME. Optimal vaccine trial design when estimating vaccine efficacy for susceptibility and infectiousness from multiple populations. *Stat Med.* 1998;17:1121–1136. [PubMed: 9618773]
15. Hudgens MG, Halloran ME. Toward causal inference with interference. *J Am Stat Assoc.* 2008;103:832–842. [PubMed: 19081744]
16. VanderWeele TJ, Tchetgen Tchetgen EJ. Effect partitioning under interference in two-stage randomized vaccine trials. *Stat Probabil Lett.* 2011;81:861–869.
17. Clemens J, Shin S, Ali M. New approaches to the assessment of vaccine herd protection in clinical trials. *Lancet Infect Dis.* 2011;11:482–487. [PubMed: 21616458]
18. Halloran ME. The minicommunity design to assess indirect effects of vaccination. *Epidemiol Methods.* 2012;1:83–105. [PubMed: 23599908]
19. Tchetgen Tchetgen EJ, VanderWeele TJ. On causal inference in the presence of interference. *Stat Methods Med Res.* 2012;21:55–75. [PubMed: 21068053]
20. VanderWeele TJ, Tchetgen Tchetgen EJ, Halloran M. Components of the indirect effect in vaccine trials: identification of contagion and infectiousness effects. *Epidemiol.* 2012;23:751–761.
21. Halloran ME, Hudgens MG. Dependents happenings: a recent methodological review. *Curr Epidemiol Rep.* 2016;3:297–305. [PubMed: 28133589]
22. Benjamin-Chung J, Arnold BF, Berger D, et al. Spillover effects in epidemiology: parameters, study designs and methodological considerations. *Int J Epidemiol.* 2018;47:332–347. [PubMed: 29106568]
23. Gunasekara FI, Richardson K, Carter K, Blakely T. Fixed effects analysis of repeated measures data. *Int J Epidemiol.* 2014;43:264–269. [PubMed: 24366487]
24. Imai K, Kim IS. When should we use unit fixed effects regression models for causal inference with longitudinal data? *Am J Pol Sci.* 2019;63:467–490.
25. Black SE, Breining S, Figlio DN, et al. Sibling spillovers. *Econ J.* 2021;131:101–128.
26. Kim Y, Steiner PM. Causal graphical views of fixed effects and random effects models. *Br J Math Stat Psychol.* 2021;74:165–183. [PubMed: 33063334]

27. Kim Y, Steiner PM. Gain scores revisited: a graphical models perspective. *Sociol Methods Res.* 2021;50:1353–1375.
28. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. *Epidemiol.* 1999;10:37–48.
29. Pearl J. *Causality: Models, Reasoning, and Inference*. 2nd edn. Cambridge, UK: Cambridge University Press, 2009.
30. Shpitser I, VanderWeele TJ, Robins JM. On the validity of covariate adjustment for estimating causal effects. In: *Proceedings of the 26th conference on uncertainty and artificial intelligence*. Corvallis, OR: AUAI Press; (2010), p. 527–536.
31. Elwert F. Graphical causal models. In: Morgan SL (ed). *Handbook of Causal Analysis for Social Research*. Dordrecht, NL: Springer Netherlands, 2013, pp. 245–273.
32. Pearl J. Linear models: a useful "microscope" for causal analysis. *J Causal Inference.* 2013;1:155–170.
33. Morgan SL, Winship C. *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. 2nd edn. Cambridge, UK: Cambridge University Press, 2014.
34. Sobel ME. What do randomized studies of housing mobility demonstrate? Causal inference in the face of interference. *J Am Stat Assoc.* 2006;101:1398–1407.
35. Elwert F, Winship C. Endogenous selection bias: the problem of conditioning on a collider variable. *Annu Rev Soc.* 2014;40:31–53.
36. Steiner PM, Kim Y. The mechanics of omitted variable bias: bias amplification and cancellation of offsetting biases. *J Causal Inference.* 2016;4.
37. Frisell T, Öberg S, Kuja-Halkola R, Sjölander A. Sibling comparison designs: bias from non-shared confounders and measurement error. *Epidemiol.* 2012;23:713–720.
38. Sjölander A, Frisell R, Öberg S. Causal interpretation of between-within models for twin research. *Epidemiol Methods.* 2012;1:217–237.
39. StataCorp. *Stata 16 Base Reference Manual*. College Station, TX: Stata Press, 2019. Available from: <https://www.stata.com/manuals/r.pdf>
40. Kuhn M, Weston S, Wing J, Forester J. The contrast Package [Internet]. 2016. Available from: <https://cran.r-project.org/web/packages/contrast/vignettes/contrast.pdf>
41. SAS Institute Inc. *SAS/STAT 13.1 User's Guide: The SCORE Procedure*. Cary, NC: SAS Institute Inc.; 2013. Available from: <https://support.sas.com/documentation/onlinedoc/stat/131/score.pdf>
42. Adkins LC, Gade MN. Monte Carlo experiments using Stata: a primer with examples. *Adv Econ.* 2012;30:429–77.
43. StataCorp. *Stata Statistical Software: Release 16*. College Station, TX: StataCorp LLC; 2019.
44. Mathiasen R, Hansen BM, Andersen AM, Forman JL, Greisen G. Gestational age and basic school achievements: a national follow-up study in Denmark. *Pediatrics.* 2010;126:e1553–e1561. [PubMed: 21059721]
45. Mallinson DC, Grodsky E, Ehrental DB. Gestational age, kindergarten-level literacy, and effect modification by maternal socio-economic and demographic factors. *Paediatr Perinat Epidemiol.* 2019;33:467–479. [PubMed: 31503367]
46. Larson A, Berger LM, Mallinson DC, Grodsky E, Ehrental DB. Variable uptake of Medicaid-covered Prenatal Care Coordination: the relevance of treatment level and service context. *J Community Health.* 2019;44:32–43. [PubMed: 30022418]
47. Invernizzi M, Juel C, Swank L, Meier J. *PALS-K Technical Reference*. Charlottesville, VA: University of Virginia Curry School of Education, 2015.
48. Wisconsin Department of Public Instruction. Admissions and early entrance to four- and five-year-old kindergarten. <https://dpi.wi.gov/early-childhood/kind/admission> (1 September 2020, date last accessed).
49. Sjölander A, Zetterqvist J. Confounders, mediators, or colliders. *Epidemiol.* 2017;28:540–547.

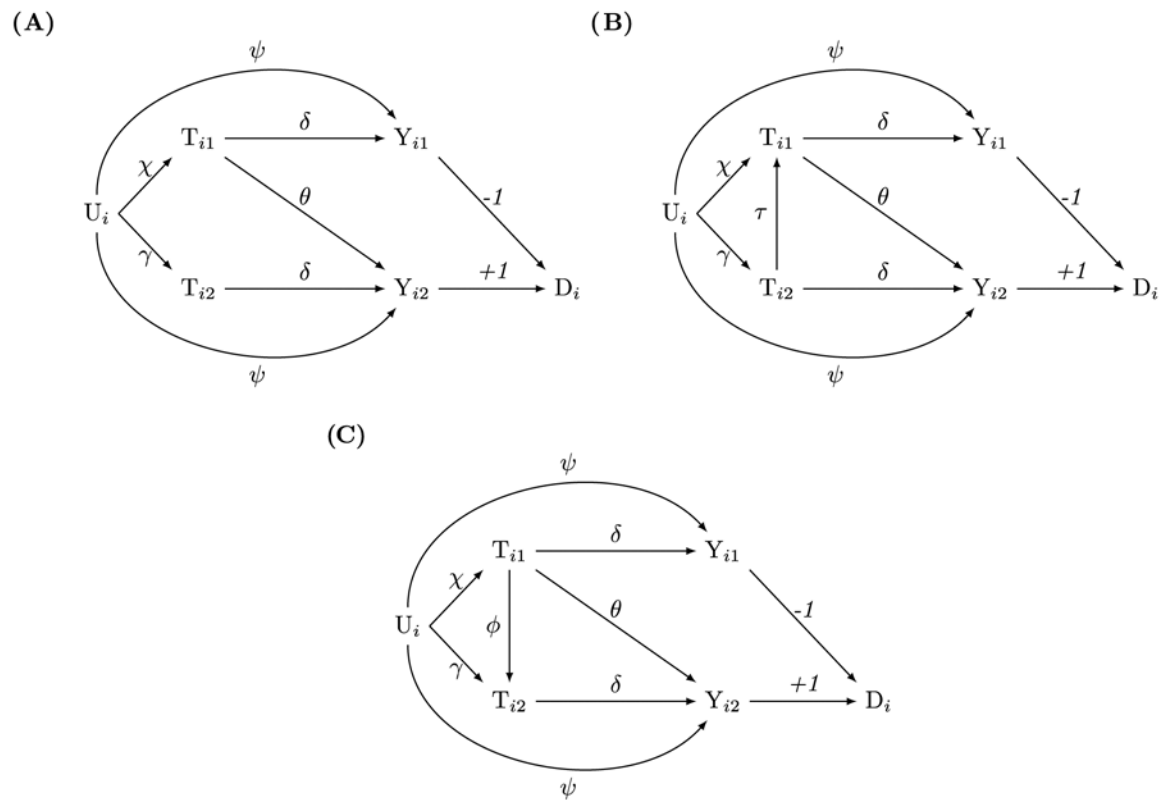


Figure 1.

Causal directed acyclic graphs for linear data-generating models with one-sided exposure-to-outcome sibling spillover. Subscripts i and j denote family and sibling, respectively. T_{ij} is the exposure, Y_{ij} is the outcome, D_j is the gain-score, and U_i is an unobserved family-level confounder. Greek letters denote effects. (1A) does not have exposure-to-exposure spillover, whereas (1B) and (1C) have exposure-to-exposure spillovers. The gain-score method precisely identifies the spillover effect $\theta(T_{i1} \rightarrow Y_{i2})$ in all three models.

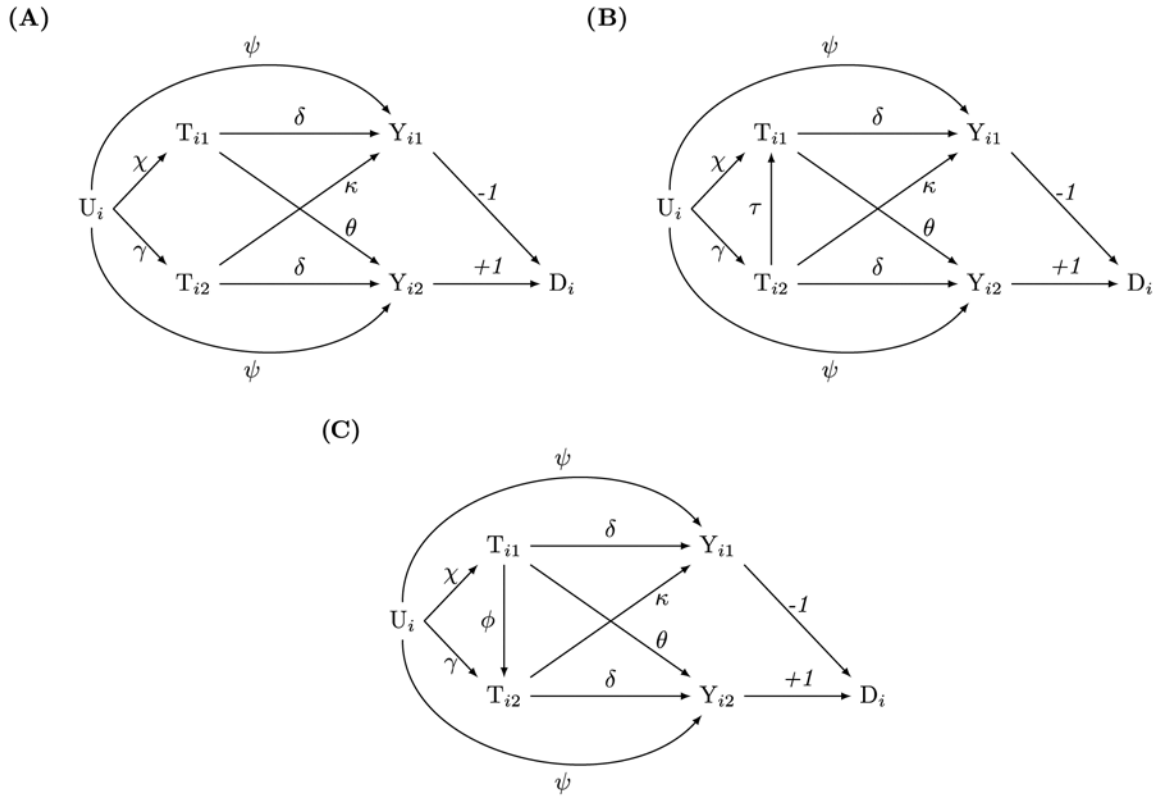
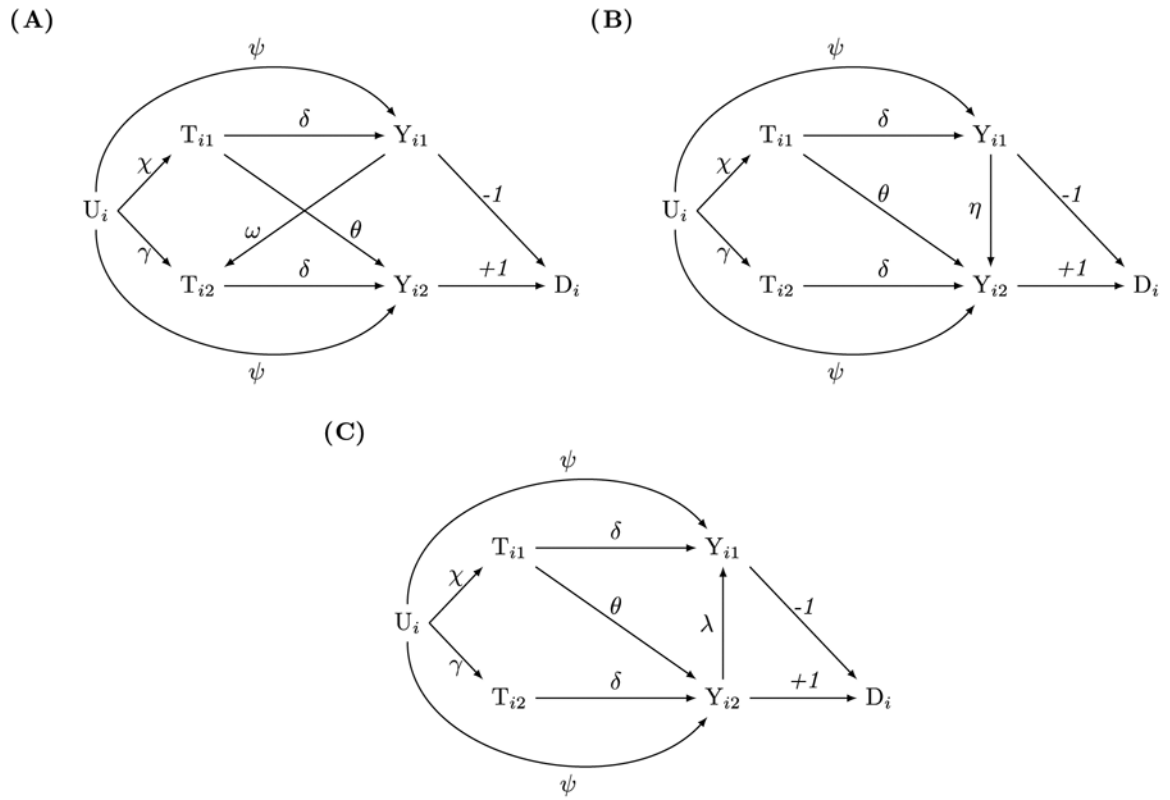


Figure 2. Causal directed acyclic graphs for linear data-generating models with two-sided exposure-to-outcome sibling spillover. Subscripts i and j denote family and sibling, respectively. T_{ij} is the exposure, Y_{ij} is the outcome, D_j is the gain-score, and U_i is an unobserved family-level confounder. Greek letters denote effects. (2A) does not have exposure-to-exposure spillover, whereas (2B) and (2C) have exposure-to-exposure spillover. The gain-score method identifies the differences between the spillover effects $\theta(T_{i1} \rightarrow Y_{i2})$ and $\kappa(T_{i2} \rightarrow Y_{i1})$ in all three models.

**Figure 3.**

Causal directed acyclic graphs for linear data-generating models with one-sided exposure-to-outcome sibling spillover and spillover from outcomes. Subscripts i and j denote family and sibling, respectively. T_{ij} is the exposure, Y_{ij} is the outcome, D_j is the gain-score, and U_i is an unobserved family-level confounder. Greek letters denote effects. (3A) has outcome-to-exposure spillover, and (3B) and (3C) have outcome-to-outcome spillover. The gain-score method does not identify the spillover effect $\theta(T_{i1} \rightarrow Y_{i2})$ in any of the three models.

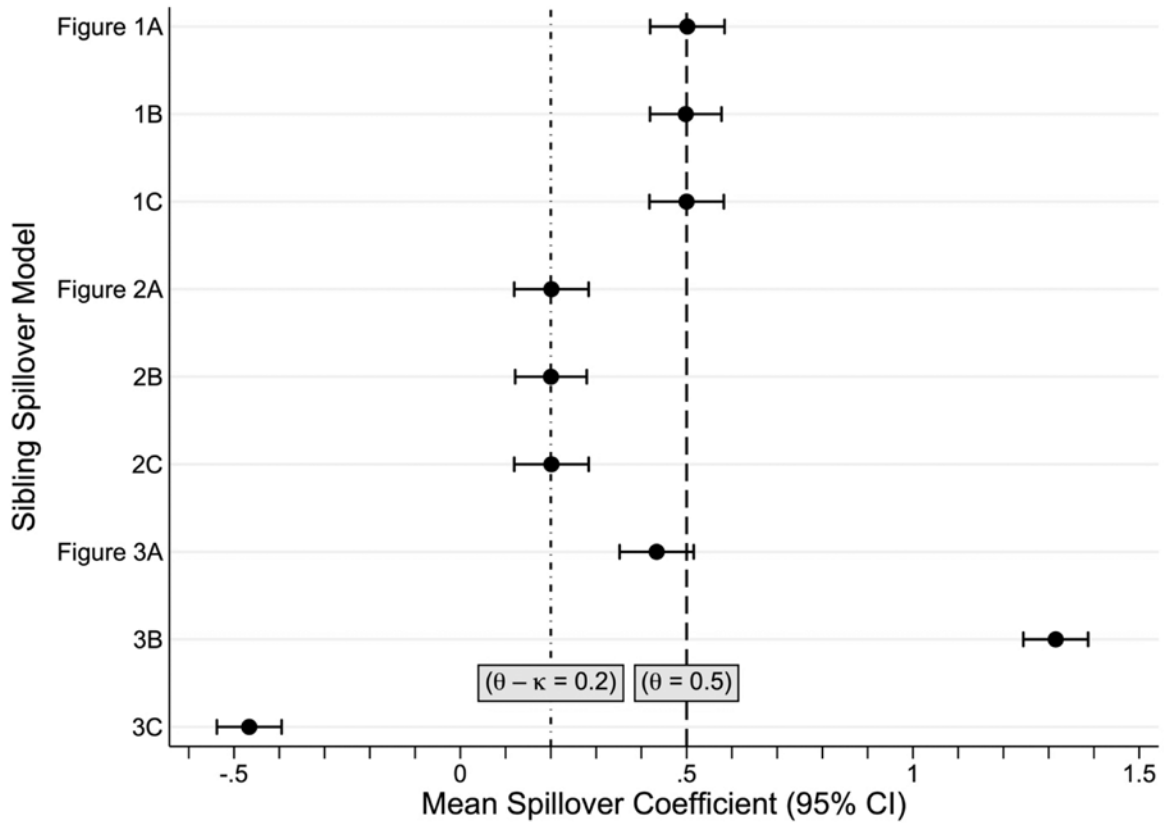


Figure 4.

Results (average spillover coefficients and empirical 95% confidence intervals [CI]) from simulations of the nine sibling spillover models in Figures 1-3. Each simulation consisted of 1000 runs of 5000 observations, where each observation represented a sibling pair (i.e., family). Subscripts i and j indicate family and sibling, respectively. T_{ij} is the exposure and Y_{ij} is the outcome. The target quantity is the spillover effect ($T_{i1} \rightarrow Y_{i2}$), set to $\theta = 0.5$ in all models. Other spillover effects include $\kappa (T_{i2} \rightarrow Y_{i1})$, $\tau (T_{i2} \rightarrow T_{i1})$, $\phi (T_{i1} \rightarrow T_{i2})$, $\psi (Y_{i1} \rightarrow T_{i2})$, $\eta (Y_{i1} \rightarrow Y_{i2})$, and $\lambda (Y_{i2} \rightarrow Y_{i1})$. Except for θ , all spillover parameters were set to zero except in the following cases: $\kappa = 0.3$ in Figures 2A-C; $\tau = 0.3$ in Figures 1B and 2B; $\phi = 0.3$ in Figures 1C and 2C; $\omega = 0.3$ in Figure 3A; $\eta = 0.3$ in Figure 3B; and $\lambda = 0.3$ in Figure 3C. The spillover coefficient identifies (i.e., is unbiased) for θ in all models of one-sided spillover (Figure 1); identifies the difference between the two exposure-to-outcome spillover effects with two-sided spillover (Figure 2); and is biased in the presence of spillovers originating from outcomes (Figure 3).

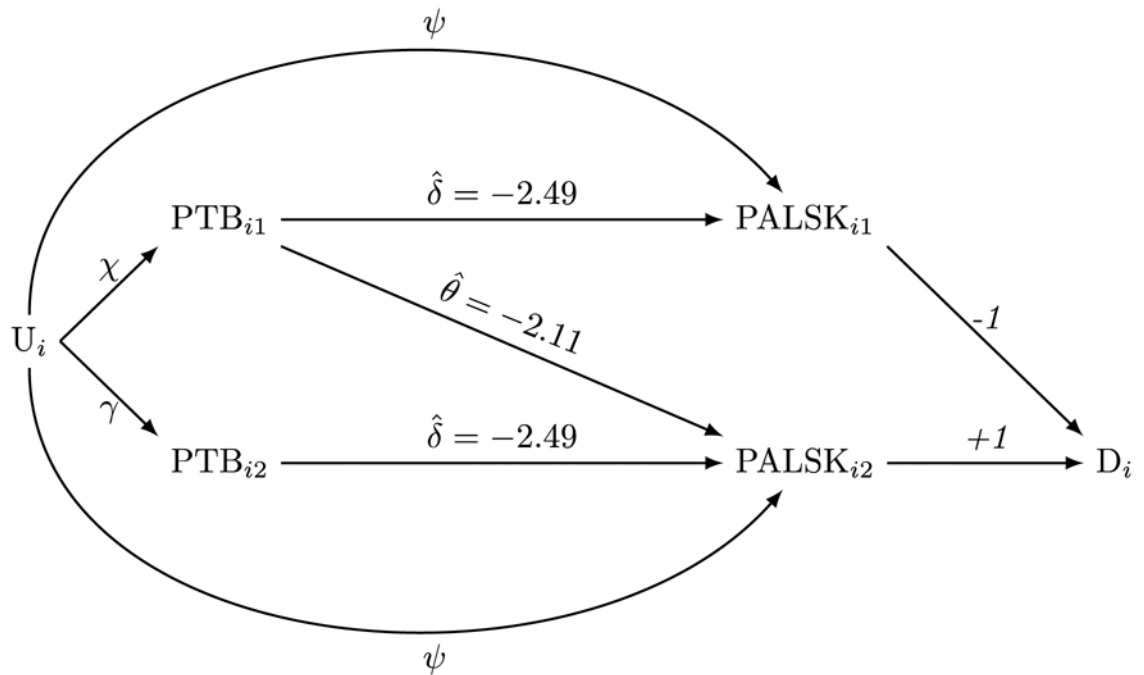


Figure 5.

A directed acyclic graph of the relationship between siblings' preterm birth (gestational age <37 weeks) and their score on the Phonological Awareness Literacy Assessment-Kindergarten test with overlaid estimates. Subscripts i and j denote family and sibling, respectively, where $j = 1$ is the younger sibling and $j = 2$ is the older sibling. PTB_{ij} is a preterm birth indicator, $PALS_{ij}$ the test score, D_i is a gain-score, and U_i is an unobserved confounder. Greek letters denote effects, and θ and δ are estimated using gain-score regression.

Table 1.

Ordinary least squares regression of the difference in siblings' PALS-K scores^a (points) on their preterm birth statuses (N = 20,010 sibling pairs)

| | Unadjusted Regression Coefficient (95% CI) | Adjusted Regression ^b Coefficient (95% CI) |
|---|---|--|
| Preterm birth (gestational age <37 weeks) | | |
| <i>Older sibling</i> | -2.49 (-3.83, -1.15) | -2.28 (-3.62, -0.94) |
| <i>Younger sibling</i> | 0.38 (-0.97, 1.73) | 0.79 (-0.57, 2.14) |
| Spillover coefficient ^c | -2.11 (-3.82, -0.40) | -1.49 (-3.21, 0.22) |

^aThe difference in PALS-K scores equals the older sibling's PALS-K Score minus the younger sibling's PALS-K score.

^bCovariates include maternal age at delivery (years), maternal education at delivery (no high school diploma; high school diploma/equivalent; 1-3 years college; 4+ years college) and Medicaid delivery payment (no; yes), all of which were measured at the time of the older sibling's delivery.

^cThe spillover coefficient is the sum of the partial regression coefficients for the older sibling's preterm birth indicator and the younger sibling's preterm birth indicator. Assuming one-sided spillover as in Figure 1A, this identifies the effect of a younger sibling's preterm birth on the older sibling's PALS-K score.

Abbreviations: "CI" confidence interval; "PALS-K" Phonological Awareness Literacy Screening-Kindergarten.