Assessing longitudinal causal effects

A comparison of marginal structural models, and structural equation modeling approaches

Jeroen D. Mulder

Department of Methodology & Statistics, Faculty of Social Sciences
Utrecht University

Consortium on Individual Development (CID)
NWO grant number: 024.001.003

June 27, 2023

Current projects

Estimating longitudinal causal effects: The overlap and differences between marginal structural models and structural equation modeling

Jeroen D. Mulder, Kim Luijken, Bas B.L. Penning de Vries, Ellen L. Hamaker

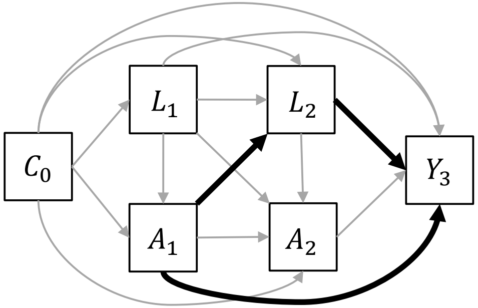
Different kinds of causal questions: Joint effects, reciprocal effects, and a comparison of three modeling approaches for estimating them

Jeroen D. Mulder, Satoshi Usami, and Ellen L. Hamaker

Empirical example: Effect of smoking cessation on weight

Variables:

- Time-varying binary exposure A: *Smoking cessation*
- Continuous end-of-study outcome Y_3: *Weight*
- Baseline covariates C_0: *Sex*, *age*, ...
- Time-varying covariates L: *Previous weight*, *hours of physical activity*, ...



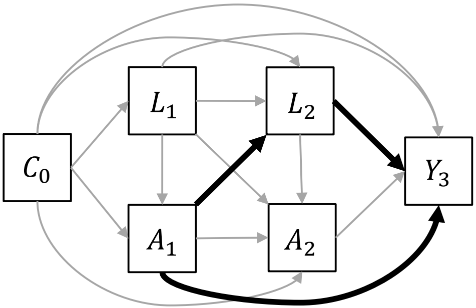
Targeted causal effect:
Controlled direct effect of A_1 on Y_3

Modeling approach 1: Structural equation modeling (SEM)

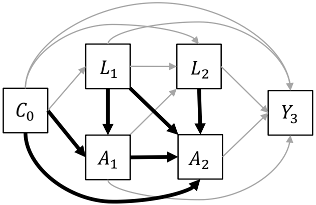
A SEM approach:
Modeling the entire data generating mechanism.

The CDE of A_1 linear combination of:

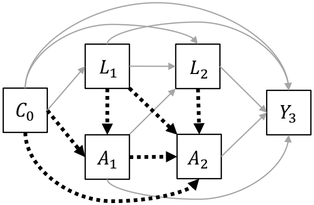
- $A_1 \rightarrow L_2$,
- $L_2 \rightarrow Y_3$, and
- $A_1 \rightarrow Y_3$.



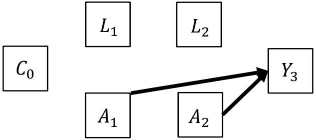
Modeling approach 2: Marginal structural models (MSM, using IPW)



Step 1:
Propensity score model



Step 2:
Balance sample using inverse probability weights (IPW)



Step 3:
Marginal structural model

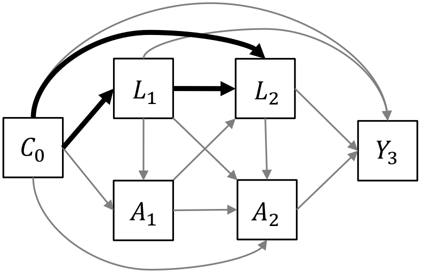
Simulation setup

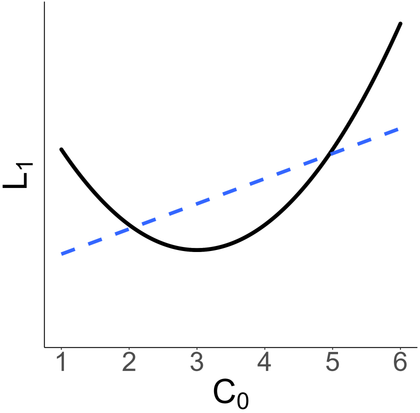
Goal:
To compare the performance of SEM and MSM approaches for estimating controlled direct effects under different scenario's of statistical misspecification.

Simulation strategy:

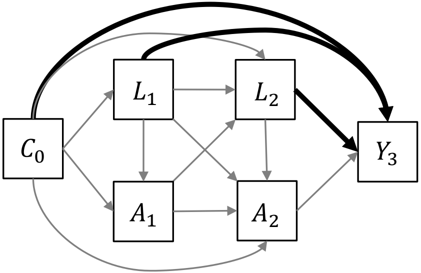
1. Create experimental conditions based on sample size $(n = 300, 1000)$, proportion exposed $(0.1,\ 0.5,\ 0.9)$, and **5 different kinds of statistical misspecification**.
2. For each experimental condition, generate 1000 datasets and estimate the CDEs using the different modeling approaches.
3. Compare the performance of the point estimates.

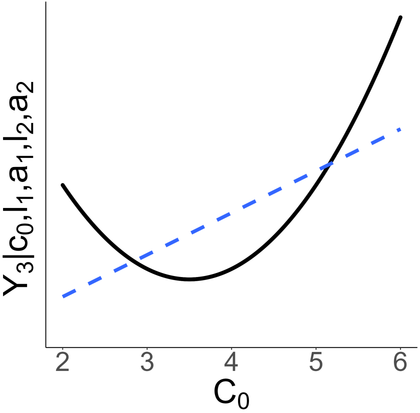
Simulation: DGM 2



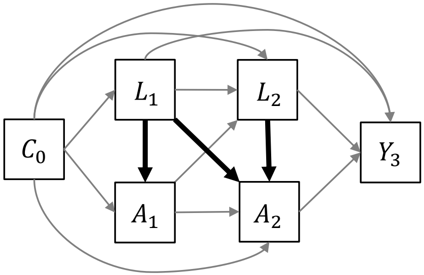


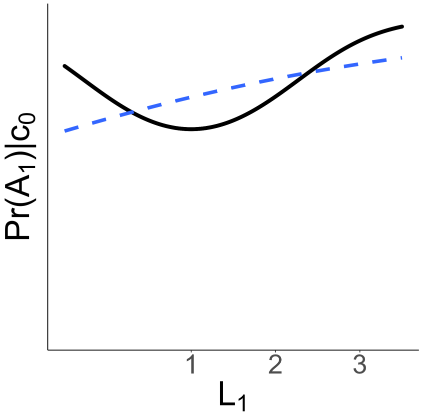
Simulation: DGM 3



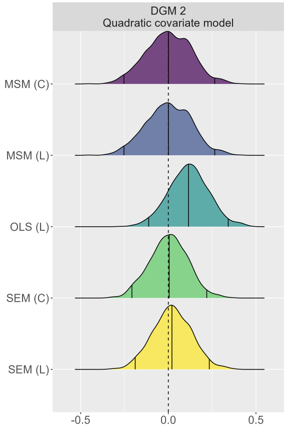


Simulation: DGM 4

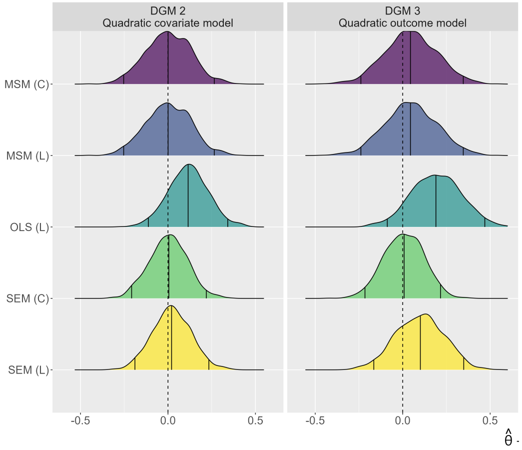




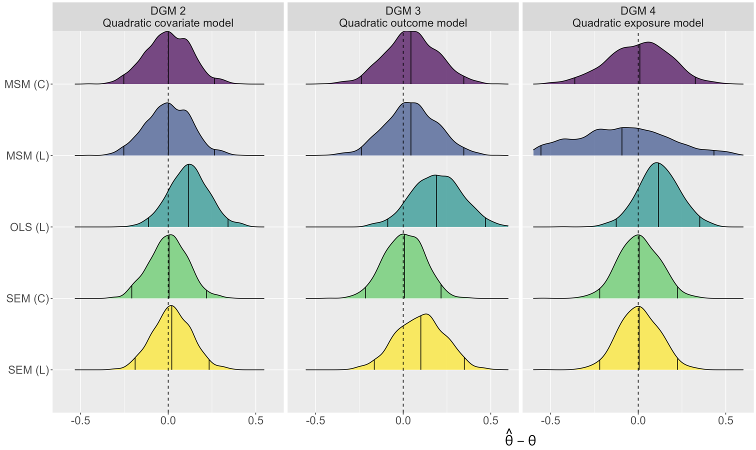
Results: $\hat{\theta} - \theta$ for CDE $A_1 \to Y_3$ $(n = 1000,\ \theta = 0.32,\ P[A_1 = 1] = .1)$



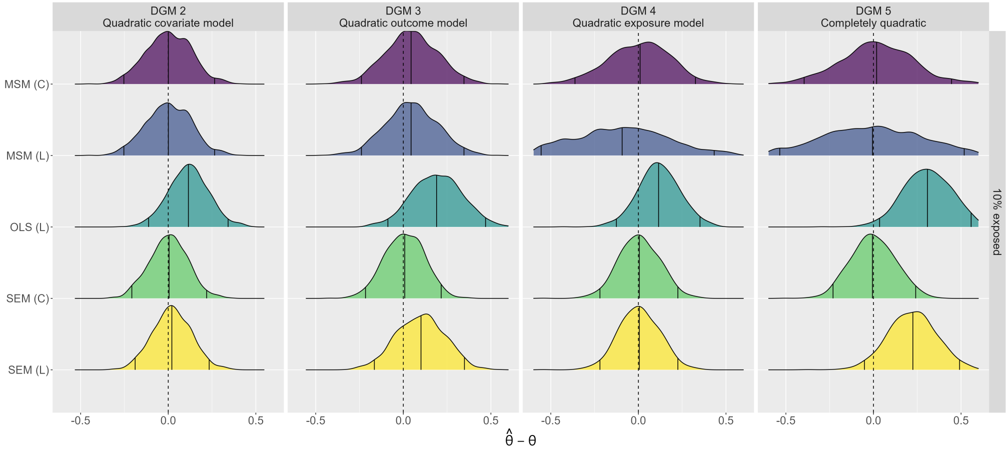
Results: $\hat{\theta} - \theta$ for CDE $A_1 \rightarrow Y_3$ ($n = 1000$, $\theta = 0.32$, $P[A_1 = 1] = .1$)



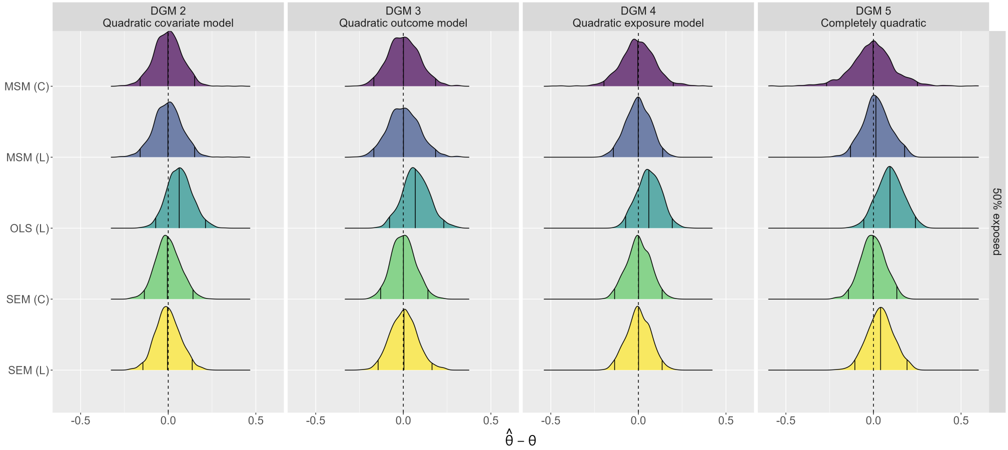
Results: $\hat{\theta} - \theta$ for CDE $A_1 \rightarrow Y_3$ ($n = 1000$, $\theta = 0.32$, $P[A_1 = 1] = .1$)



Results: $\hat{\theta} - \theta$ for CDE $A_1 \to Y_3$ ($n = 1000$, $\theta = 0.32$, $P[A_1 = 1] = .1$)



Results: $\hat{\theta} - \theta$ for CDE $A_1 \rightarrow Y_3$ ($n = 1000,\ \theta = 0.32,\ P[A_1 = 1] = .5$)



Conclusions (preliminary)

Results are in line with the literature:

- Statistical misspecification in **SEM**s negatively affect validity of estimates (in other parts of the model, and to varying degrees).
- **MSM (IPW)** approach inefficient under misspecification of PS model and exposure imbalance (Vansteelandt & Sjolander, 2016).
- Bias and inefficiency less pronounced when balanced exposed/non-exposed, for both **MSM (IPW)** and **SEM**.
- Robustness: The maximum amount of "wrongness" is less for **MSM (IPW)** than for **SEM**.

Limitations

1. Limited amount of simulation scenarios and targeted causal effects.
2. Our conclusion are, in principle, not new. However, these projects:
   - bridge the disciplinary disconnected between the SEM and formal causal inference literature, and
   - evaluate the formal causal inference framework for a psychological context (using popular psychological data such as *LISS data*, and *self-esteem, rumination, and depression data*, Scherpenzeel, 2018; Kuster et al., 2012).
3. Alternative estimation methods for MSMs that offer additional benefits, such as doubly robustness (e.g. Tompsett et al., 2022).

Challenges: The formal causal roadmap in (psychological) practice

- Consistency likely to be compromised for psychological exposures/interventions (Eronen, 2020)
- Time zero: When are participants "eligible for the intervention" (Hernán et al., 2016)? *LISS data*: Newly at risk individuals for cardio-vascular disease: "Rolling enrollment" 2007-2020 (excl. 2014) results in 84 individuals.
- Missing data handling
- Readability of causal inference literature for applied researchers, and implementation of techniques in software (and documentation thereof).

Questions and discussion

Thank you!

Questions?

References I

Eronen, M. I. (2020). Causal discovery and the problem of psychological interventions. *New Ideas in Psychology*, *59*, 100785. https://doi.org/10.1016/j.newideapsych.2020.100785

Hernán, M. A., Sauer, B. C., Hernández-Díaz, S., Platt, R., & Shrier, I. (2016). Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses. *Journal of Clinical Epidemiology*, *79*, 70–75. https://doi.org/10.1016/j.jclinepi.2016.04.014

Kuster, F., Orth, U., & Meier, L. L. (2012). Rumination mediates the prospective effect of low self-esteem on depression: A five-wave longitudinal study. *Personality and Social Psychology Bulletin*, *38*(6), 747–759. https://doi.org/10.1177/0146167212437250

Scherpenzeel, A. C. (2018). "True" Longitudinal and Probability-Based Internet Panels: Evidence From the Netherlands. In M. Das, P. Ester, & L. Kaczmirek (Eds.), *Social and Behavioral Research and the Internet* (1st ed., pp. 77–104). Routledge. https://doi.org/10.4324/9780203844922-4

References II

Tompsett, D., Vansteelandt, S., Dukes, O., & De Stavola, B. (2022). Gesttools: General purpose g-estimation in R. *Observational Studies*, *8*(1), 1–28. https://doi.org/10.1353/obs.2022.0003

van der Laan, M. J., & Rose, S. (2011). *Targeted Learning*. Springer New York. https://doi.org/10.1007/978-1-4419-9782-1

VanderWeele, T. J. (2012). Invited commentary: Structural equation models and epidemiologic analysis. *American Journal of Epidemiology*, *176*(7), 608–612. https://doi.org/10.1093/aje/kws213

Vansteelandt, S., & Sjolander, A. (2016). Revisiting g-estimation of the Effect of a Time-varying Exposure Subject to Time-varying Confounding. *Epidemiologic Methods*, *5*(1), 37–56. https://doi.org/10.1515/em-2015-0005

Critique on SEM techniques

"So when should SEMs be used in epidemiology? I would argue that, in light of the strong assumptions made, they should be used only when (...) 2) **we are using them principally for exploratory and hypothesis-generating purposes***."*

- VanderWeele (2012)

"When (...) the statistical model is misspecified, **the estimate of the probability distribution can be extremely biased, and it is not even clear what the parameter estimates are even estimating.***"*

- van der Laan and Rose (2011)