

Marginal Structural Models in the Social Sciences: A Literature Review

1 Introduction

Marginal structural models (MSMs) are a class of causal models designed to estimate the effects of time-varying treatments in the presence of time-varying confounders that are themselves affected by prior treatment. Developed primarily within biostatistics and epidemiology, MSMs have become increasingly influential across the social sciences, offering researchers a principled approach to longitudinal causal inference that avoids biases inherent in standard regression-based methods. This review traces the origins of MSMs, describes their methodological foundations, and surveys their adoption and development within political science, sociology, criminology, demography, economics, and psychology.

2 Origins in Biostatistics and Epidemiology

2.1 The Problem of Time-Varying Confounding

The intellectual foundations of MSMs lie in a fundamental problem of longitudinal causal inference. In many observational studies, researchers seek to estimate the effect of a treatment or exposure that varies over time. A standard approach would adjust for confounders—variables that jointly influence treatment and outcome—by including them as covariates in

a regression model. However, when a time-varying confounder is itself affected by prior treatment, conditioning on it introduces post-treatment bias while failing to condition on it leaves confounding uncontrolled. Standard regression methods cannot resolve this dilemma (Robins, 1986; Robins et al., 2000).

Consider the canonical example from HIV epidemiology: a researcher wishes to estimate the effect of antiretroviral therapy on survival. CD4 cell count is a time-varying confounder—it predicts both future treatment decisions and survival—but it is also affected by prior treatment. Including CD4 count as a covariate in a regression model blocks part of the causal effect of treatment (mediated through CD4), while excluding it leaves the treatment–outcome relationship confounded. MSMs were developed precisely to navigate this methodological impasse.

2.2 Robins and the G-Methods

The development of MSMs is primarily the work of James M. Robins, a biostatistician and epidemiologist at Harvard University. Robins (1986) introduced the *g-computation algorithm*, a general method for estimating causal effects of time-varying treatments under what would later be called the sequential exchangeability assumption. This foundational paper established the counterfactual framework for longitudinal treatments and demonstrated that standard regression fails when time-varying confounders are affected by prior treatment, using the “healthy worker survivor effect” as a motivating example.

MSMs belong to a broader family of techniques known as Robins’ “g-methods,” which also includes g-computation and g-estimation for structural nested models (Naimi et al., 2017). Robins first formally introduced MSMs as a distinct class of models in Robins (1997), distinguishing them from the previously developed structural nested models. A detailed comparison of the two approaches appeared in Robins (1999).

2.3 The Canonical Exposition

The paper that brought MSMs to widespread use is Robins et al. (2000), published in *Epidemiology*. This article provided the first accessible exposition of MSMs for applied researchers, clearly describing the problem of time-varying confounding affected by prior treatment and introducing inverse probability of treatment weighting (IPTW) as the primary estimation strategy. The authors formalized the key identification assumptions: consistency (the observed outcome equals the counterfactual outcome under the observed treatment), sequential exchangeability (no unmeasured confounding conditional on measured covariate history), positivity (every unit has a positive probability of receiving each treatment level at every time point), and correct specification of the treatment model.

The term “marginal” in marginal structural models refers to the fact that these models specify the marginal distribution of counterfactual outcomes—that is, population-average causal effects—rather than conditional distributions. The term “structural” draws on the econometric and social science tradition of using “structural” to denote causal models, as opposed to merely associational ones.

Published alongside this methodological paper, Hernán et al. (2000) provided the first major empirical application. Using data from the Multicenter AIDS Cohort Study, they estimated the causal effect of zidovudine on survival among HIV-positive men. The results were striking: the crude mortality rate ratio suggested that zidovudine was harmful (rate ratio of 3.6), whereas the MSM-adjusted estimate indicated a protective effect (rate ratio of 0.7). The dramatic reversal illustrated how confounding by indication—where sicker patients are more likely to receive treatment—can produce severely misleading estimates. Hernán et al. (2001) subsequently extended the framework to estimate the joint causal effects of multiple time-varying treatments.

2.4 Practical and Methodological Developments

Following these foundational contributions, several papers addressed practical challenges in implementing MSMs. Cole and Hernán (2008) provided detailed guidance on constructing inverse probability weights, including the use of stabilized weights (where the numerator of the weight is the marginal probability of treatment) to reduce variance while preserving consistency. This paper became a standard reference for applied researchers.

Shinozaki and Suzuki (2020) identified common pitfalls in applied MSM work, including the conflation of MSMs (the causal model for counterfactual outcomes) with IPTW (the estimation method) and the failure to distinguish the structural model from the nuisance model for treatment probability. Bijlsma et al. (2022) connected MSMs to life-course theories in social epidemiology, describing how they map onto common theoretical frameworks such as critical period, accumulation, and sensitive period models.

The definitive textbook treatment of MSMs appears in Hernán and Robins (2020), which devotes substantial attention to IP weighting, MSMs for point treatments and time-varying treatments, and the broader family of g-methods.

3 Marginal Structural Models in the Social Sciences

The adoption of MSMs in the social sciences has been motivated by a recognition that many core questions in these disciplines involve time-varying treatments and confounders with precisely the structure that MSMs were designed to address. Panel data—the workhorse of empirical social science—routinely exhibit feedback between treatments, confounders, and outcomes across time periods. The sections below review key applications and methodological contributions by discipline.

3.1 Criminology: An Early Adoption

One of the earliest and most influential applications of MSMs outside epidemiology appeared in criminology. Sampson et al. (2006) used IPTW to estimate the within-individual causal effect of marriage on criminal offending, using longitudinal data on 500 high-risk boys followed from adolescence to age 32. They found that being married reduces the odds of crime by approximately 35 percent. This study was pathbreaking in applying modern causal inference methods to the longstanding criminological question of desistance from crime, demonstrating that the “marriage effect” persists after accounting for the dynamic selection processes through which individuals enter and exit marriage. More recently, Van Der Wal (2022) applied MSMs to U.S. state-level panel data to estimate the causal effect of right-to-carry gun laws on crime rates, correcting for a range of time-varying economic, demographic, and political confounders.

3.2 Sociology: Neighborhood Effects and Cumulative Exposure

MSMs have found particularly fertile ground in sociology, where researchers have long been interested in cumulative exposure effects—the consequences of sustained exposure to social conditions over time. Wodtke et al. (2011) produced a landmark application by using MSMs to estimate the causal effect of sustained neighborhood disadvantage on high school graduation. Following 4,154 children in the Panel Study of Income Dynamics from age 1 to 17, with annual neighborhood measurements, they found that the cumulative effect of sustained neighborhood disadvantage is substantially larger than cross-sectional studies had suggested. This study powerfully illustrated that standard regression approaches, which condition on time-varying neighborhood characteristics, understate cumulative effects by simultaneously adjusting away part of the treatment effect.

Sharkey and Elwert (2011) extended the MSM framework to multigenerational settings, estimating the joint effect of parent and child exposure to neighborhood poverty on child cognitive ability. They found that two generations of continuous exposure to high-poverty

neighborhoods reduces cognitive ability by more than half a standard deviation—a finding with profound implications for understanding the intergenerational transmission of disadvantage. Wodtke (2020) subsequently provided methodological guidance for social scientists, comparing IPTW estimation of MSMs with regression-based alternatives for structural nested mean models.

3.3 Political Science: Dynamic Causal Inference

Blackwell (2013) provided the foundational treatment of MSMs for political science. Adapting the framework from biostatistics, he demonstrated that many questions in political science concern dynamic treatment strategies—sequences of decisions made over time in response to evolving conditions—and that standard approaches to panel data force an impossible choice between omitted variable bias and post-treatment bias. He applied MSMs to estimate the effectiveness of negative campaign advertising, showing that campaign strategy decisions are shaped by prior performance in ways that confound naive estimates.

Blackwell and Glynn (2018) extended this line of work to time-series cross-sectional (TSCS) data, the dominant data structure in comparative politics and international relations. They showed that two of the most common specifications for TSCS data—the lagged dependent variable model and the unit fixed effects model—both produce biased estimates under plausible data-generating processes involving time-varying confounding, and presented IPTW for MSMs as a viable alternative. Blackwell (2020) further applied MSMs to estimate controlled direct effects in political science, decomposing total effects into direct and indirect components.

Several methodological innovations in the estimation of MSMs have emerged from political science. Imai and Ratkovic (2015) addressed the practical sensitivity of MSMs to misspecification of the treatment model by generalizing the covariate balancing propensity score to longitudinal settings, producing weights that directly optimize covariate balance. Zhou and Wodtke (2020) proposed “residual balancing” as an alternative to IPTW, modeling

conditional means of post-treatment confounders rather than conditional treatment distributions. This approach is particularly attractive for continuous treatments and reduces sensitivity to model misspecification.

3.4 Economics: Program Evaluation

In economics, Lechner (2009) brought inverse selection probability weighting for dynamic causal effects into econometric program evaluation, applying the approach to Swiss active labor market programs. He proposed distribution theory based on sequential GMM estimation, connecting the MSM framework to the econometric tradition. While economists have more commonly relied on alternative identification strategies such as instrumental variables, regression discontinuity, and difference-in-differences, the MSM framework has gained recognition as a valuable tool when treatment varies over time and the selection-on-observables assumption is credible.

3.5 Demography and Family Sociology

Bacak and Kennedy (2015) introduced MSMs to the family demography literature, applying IPTW to estimate the causal effect of incarceration on marriage during young adulthood using panel data from the National Longitudinal Survey of Youth. They found that incarceration significantly reduces the likelihood of entering marriage, contributing to the growing literature on the collateral consequences of mass incarceration for family formation.

3.6 Psychology: Developmental and Trauma Research

Thoemmes and Ong (2016) provided an accessible tutorial introducing MSMs and IPTW to developmental psychologists, noting that the non-randomized transitions of emerging adulthood—leaving home, entering college, starting careers—create precisely the kind of time-varying confounding that MSMs are designed to address. This paper served as a key

entry point for psychologists learning these methods and has facilitated subsequent applications in developmental and clinical psychology.

4 Key Textbooks and Pedagogical Resources

Several textbooks have been instrumental in making MSMs accessible to social scientists. Hernán and Robins (2020) remains the definitive reference, offering a comprehensive and freely available treatment of causal inference methods including MSMs. Morgan (2013) introduced modern causal inference methods—including the directed acyclic graph (DAG) framework underlying MSMs—to a broad social science audience, with Elwert (2013) providing a particularly important chapter on graphical causal models. VanderWeele (2015) offered an extensive treatment of mediation analysis from a causal inference perspective, including the use of MSMs for estimating controlled direct effects. Hong (2015) introduced weighting-based causal inference methods to social science audiences with applications centered on social settings such as families, classrooms, and neighborhoods.

5 Conclusion

Marginal structural models have traveled a remarkable path from their origins in biostatistics to become an increasingly important tool across the social sciences. The core insight—that standard regression methods fail when time-varying confounders are affected by prior treatment, and that inverse probability weighting offers a principled alternative—has proven relevant to a wide range of substantive questions, from the effects of neighborhoods on educational attainment to the consequences of campaign strategy for election outcomes.

The diffusion of MSMs into the social sciences has not been merely a matter of applying existing methods to new data. Social scientists have contributed important methodological innovations, including new weight construction methods that improve robustness (Imai and Ratkovic, 2015; Zhou and Wodtke, 2020), extensions to controlled direct effects (Blackwell,

2020), and careful examinations of how MSMs relate to the standard econometric toolkit for panel data (Blackwell and Glynn, 2018; Lechner, 2009). At the same time, substantive applications have demonstrated the empirical importance of properly accounting for time-varying confounding: estimates from MSMs frequently differ substantially from those produced by conventional methods, sometimes reversing the sign of estimated effects.

Despite their growing influence, MSMs are not without limitations. They rely on the assumption of no unmeasured confounding—an assumption that, while common in observational research, is fundamentally untestable. They can be sensitive to misspecification of the treatment model, and extreme weights can introduce instability, particularly in small samples or when positivity is only weakly satisfied. Researchers in the social sciences have actively engaged with these challenges, developing diagnostics and alternative estimation strategies that improve the practical reliability of MSMs.

As longitudinal data become increasingly available and social scientists continue to grapple with questions about cumulative and dynamic causal processes, MSMs are likely to play an even larger role in empirical research. Their capacity to handle the complex temporal dependencies that characterize social life makes them an indispensable addition to the methodological toolkit of the social sciences.

References

- Bacak, V. and Kennedy, E. H. (2015). Marginal structural models: An application to incarceration and marriage during young adulthood. *Journal of Marriage and Family*, 77(1):112–125.
- Bijlsma, M. J., Mortensen, L. H., et al. (2022). Marginal structural models for life-course theories and social epidemiology: Definitions, sources of bias, and simulated illustrations. *American Journal of Epidemiology*, 191(2):349–359.

- Blackwell, M. (2013). A framework for dynamic causal inference in political science. *American Journal of Political Science*, 57(2):504–520.
- Blackwell, M. (2020). Estimating controlled direct effects through marginal structural models. *Political Science Research and Methods*.
- Blackwell, M. and Glynn, A. N. (2018). How to make causal inferences with time-series cross-sectional data under selection on observables. *American Political Science Review*, 112(4):1067–1082.
- Cole, S. R. and Hernán, M. A. (2008). Constructing inverse probability weights for marginal structural models. *American Journal of Epidemiology*, 168(6):656–664.
- Elwert, F. (2013). Graphical causal models. In Morgan, S. L., editor, *Handbook of Causal Analysis for Social Research*, pages 245–273. Springer, Dordrecht.
- Hernán, M. A., Brumback, B., and Robins, J. M. (2000). Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. *Epidemiology*, 11(5):561–570.
- Hernán, M. A., Brumback, B., and Robins, J. M. (2001). Marginal structural models to estimate the joint causal effect of nonrandomized treatments. *Journal of the American Statistical Association*, 96(454):440–448.
- Hernán, M. A. and Robins, J. M. (2020). *Causal Inference: What If*. Chapman & Hall/CRC, Boca Raton.
- Hong, G. (2015). *Causality in a Social World: Moderation, Mediation and Spill-over*. Wiley, Chichester.
- Imai, K. and Ratkovic, M. (2015). Robust estimation of inverse probability weights for marginal structural models. *Journal of the American Statistical Association*, 110(511):1013–1023.

- Lechner, M. (2009). Sequential causal models for the evaluation of labor market programs. *Journal of Business & Economic Statistics*, 27(1):71–83.
- Morgan, S. L., editor (2013). *Handbook of Causal Analysis for Social Research*. Springer, Dordrecht.
- Naimi, A. I., Cole, S. R., and Kennedy, E. H. (2017). An introduction to G methods. *International Journal of Epidemiology*, 46(2):756–762.
- Robins, J. M. (1986). A new approach to causal inference in mortality studies with a sustained exposure period—application to control of the healthy worker survivor effect. *Mathematical Modelling*, 7(9–12):1393–1512.
- Robins, J. M. (1997). Causal inference from complex longitudinal data. In Berkane, M., editor, *Latent Variable Modeling and Applications to Causality*, volume 120 of *Lecture Notes in Statistics*, pages 69–117. Springer, New York.
- Robins, J. M. (1999). Marginal structural models versus structural nested models as tools for causal inference. In Halloran, M. E. and Berry, D., editors, *Statistical Models in Epidemiology, the Environment, and Clinical Trials*, volume 116 of *The IMA Volumes in Mathematics and its Applications*, pages 95–133. Springer, New York.
- Robins, J. M., Hernán, M. A., and Brumback, B. (2000). Marginal structural models and causal inference in epidemiology. *Epidemiology*, 11(5):550–560.
- Sampson, R. J., Laub, J. H., and Wimer, C. (2006). Does marriage reduce crime? A counterfactual approach to within-individual causal effects. *Criminology*, 44(3):465–508.
- Sharkey, P. and Elwert, F. (2011). The legacy of disadvantage: Multigenerational neighborhood effects on cognitive ability. *American Journal of Sociology*, 116(6):1934–1981.
- Shinozaki, T. and Suzuki, E. (2020). Understanding marginal structural models for time-varying exposures: Pitfalls and tips. *Journal of Epidemiology*, 30(9):377–389.

- Thoemmes, F. and Ong, A. D. (2016). A primer on inverse probability of treatment weighting and marginal structural models. *Emerging Adulthood*, 4(1):40–59.
- Van Der Wal, W. M. (2022). Marginal structural models to estimate causal effects of right-to-carry laws on crime. *Statistics and Public Policy*, 9(1):163–174.
- VanderWeele, T. J. (2015). *Explanation in Causal Inference: Methods for Mediation and Interaction*. Oxford University Press, New York.
- Wodtke, G. T. (2020). Regression-based adjustment for time-varying confounders. *Sociological Methods & Research*, 49(4):906–946.
- Wodtke, G. T., Harding, D. J., and Elwert, F. (2011). Neighborhood effects in temporal perspective: The impact of long-term exposure to concentrated disadvantage on high school graduation. *American Sociological Review*, 76(5):713–736.
- Zhou, X. and Wodtke, G. T. (2020). Residual balancing: A method of constructing weights for marginal structural models. *Political Analysis*, 28(4):487–506.