Statistica Sinica **36** (2026), 1-6
doi:https://doi.org/10.5705/ss.202025.0152

DISCUSSION ON "CAUSAL AND COUNTERFACTUAL VIEWS OF MISSING DATA MODELS"

Shanshan Luo[1] and Zhi Geng*[1,2]

[1]*Beijing Technology and Business University and* [2]*Peking University*

1. Introduction

We sincerely congratulate Professors Nabi, Bhattacharya, Shpitser, and Robins on their interesting contribution, which addresses challenges in missing data analysis through the framework of potential outcomes. In this paper, the authors explain how missing data problems can be framed as causal inference problems: the complete variables are viewed as counterfactual outcomes, the missingness indicators are treated as treatment variables, and the partially observed variables are interpreted as combinations of potential outcomes and treatments. The authors introduce missing data directed acyclic graphs (m-DAGs), review several missing data models from previous literature that can be represented using m-DAGs, and present identifiability results for various graph structures.

In this discussion, we explore more identifiable missing not at random (MNAR) models within the m-DAG framework, with a particular focus on causal graphs involving binary outcomes. We observe that in many of the figures presented in the main text, the complete counterfactual variables are not allowed to have direct effects on their own missingness indicators. In our discussion, we consider alternative scenarios in which the missingness mechanism may be directly affected by the outcome variable itself, as in the self-censoring problem (Nabi, Bhattacharya and Shpitser, 2020; Li et al., 2023), and examine the associated identifiability issues. We first study identifiability without effects between missingness indicators, and then consider how auxiliary variables can facilitate identification with effects between missingness indicators. The Supplementary Material providing the complete proofs of the theoretical results is available online at `https://gitlab.com/pipishan95/mdag_SM`.

2. Identification in Self-Censoring m-DAGs

In this discussion, we focus on an MNAR mechanism in which the missingness of a variable depends on its unobserved outcome. This type of mechanism is

*Corresponding author. E-mail: zhigeng@pku.edu.cn