Examining the causal direction of synchronous effects in structural panel models: Using residual-predictor independence tests to identify model mis-specifications

Xintong Li

University of Missouri

Xintong Li

Assessment Resource Center

College of Education

University of Missouri

303D Townsend Hall, Columbia, MO 65201

phone: (573) 882-9320

email: xl579@mail.missouri.edu

Background

Establishing cause-effect relations between variables is one of the focal interests of applied researchers in educational context. Though conventionally causal direction is examined with experimental studies, observational data are increasingly drawing the attention of educational researchers, as more algorithms for causal discovery using observational data was proposed in the recent decades. In addition, observational data can typically be collected with less cost, less ethical concerns, and more authenticity when collecting from an operating educational system, such as a teacher evaluation system or an administrative system, etc. Structural panel models, especially Cross lagged Panel Models (CLPMs), are often used when observational data are collected over time to examine causal dominance. CLPM is a longitudinal structural equation model (SEM) widely used in educational context that allows researchers to test causal predominance by comparing cross-lagged relationships across multiple time points while controlling for correlations within time points and construct stability across time points (Kearney, 2017; Newsom, 2015). However, correlational methods, such as SEM, shed no light on the direction of effects of two particular variables, as they are based on variances and covariance, whose inherent symmetric properties prevent researchers from distinguishing directionally competing causal models (Dodge & Rousson, 2000; Wiedermann & Hagmann, 2016). That is to say, when a synchronous effect exist in a panel model, such as a survey bias, SEM is unable to distinguish its causal direction by simply reversing the “arrow” in the model. In addition, the mis-specification of the direction of a synchronous effect may also lead to erroneous conclusions on the cross-lagged effects. The recent development of machine learning literature has witnessed various asymmetric properties proposed to discover the causal direction in synchronous observational data. Predictor-residual independence is one of them for causal discovery in non-linear regression models or linear models with non-normal variables (Mooij, Peters, Janzing, Zscheischler & Schölkopf, 2016; Wiedermann, & von Eye, 2015). Various tests were proposed to examine this property, among which the Breusch–Pagan test, non-linear correlation tests and the Hilbert Schmidt Independence Criterion (HSIC) consistently show adequate performance with cross-sectional data. Moreover, statistical tools has been developed for empirical researchers (Wiedermann & Li, 2018; Wiedermann & Li, 2019), which can be used to examine the causal direction of synchronous effects and avoid model mis-specification of structural panel models.

Research Questions

1. Will a mis-specified direction of synchronous effect or the neglecting of the synchronous effect lead to erroneous causal conclusions in cross-lagged panel models?
2. Can Breusch–Pagan test, non-linear correlation tests and Hilbert Schmidt Independence Criterion (HSIC) inform the selection of structural panel models with competing direction of a synchronous effect?
3. Can the proposed tests be used with empirical data to identify the direction of a synchronous effect?

A Monte Carlo simulation study was conducted to answer the first two research question. An empirical inquiry was included to answer the third research question, which consists of student ratings of teachers collected from an authentic teacher evaluation system in Missouri, i.e. the Network for Educator effectiveness (NEE; neeadvantage.com).

Simulation Design

A two-time-point panel model is simulated with the sample size of 800, and 8 simulation conditions are included based on two synchronous effects, the presence of a cross-lagged effect, and two shape parameters of time 1 predictor, which is generated from a gamma distribution. 500 iterations is conducted for each condition. The performance of three structural panel models are estimated, including the "true" model (Model T), the mis-specified synchronous model (Model M) and cross-lagged model without synchronous effect (Model C).

Participants

For the empirical inquiry, 632 teachers' student ratings of performance were collected across two school years, i.e. 2017-2018 and 2018-2019. Student level information is strictly anonymous, and the teacher level data are based on mean aggregation of student ratings.

Data Collection and Analysis

The empirical data are obtained from an authentic teacher evaluation system, the Network for Educator Effectiveness, with IRB approval. Two focal measures are included in the empirical data, i.e. cognitive engagement (CE) and teacher-student relation (TSR). Both are latent factor scores based on measures from the Teacher Effectiveness Student Survey (TESS) used by NEE. Teacher years of experience is also included as a potential confounder variable. TESS includes 3 screening items that are evenly distributed among the survey items in each survey session to identify inattentive responses (Cornell, Klein, Konold, & Huang, 2012). Surveys that fail (score < 2) two of the three screening items are flagged for manual review. To minimize the threats to validity due to high turnover in some districts where students move during a school year, a screening question is also included asking if they have been in the class for at least a month. If they answer "no," their survey is eliminated. Further, responses that are finished within unrealistic time (three standard deviations from the mean) are also flagged for manual review.

Findings

As for the simulation study, model fit measures support all the three structural panel models. **Model T** has adequate performance in recovering the synchronous effect and the cross-lagged effect. In contrast, when the synchronous effect is mis-specified, **Model M** failed to recover the true cross-lagged effect. **Model C**, on the other hand, return erroneous results favoring one of the cross-lagged effects when none of them is present.

As for the empirical study, as expected, three types of predictor-residual independence tests, i.e. heteroscedasticity tests (Breusch & Pagan, 1979), non-linear correlation tests (Wiedermann & Li, 2018) and the Hilbert-Schmidt Independence Criterion (HSIC; Gretton et al., 2008), all support a sychronous TSR $\rightarrow$ CE effect, which suggests that teachers tend to have better CE pratices when they maintain a good relation with the students. Another possible contributor is the halo effect of teachers when they have better relation with students.

Conclusions

Examining the direction of synchronous effect in a structural panel model is necessary to avoid potential model mis-specification and consequent biases. The proposed predictor-residual independence tests have the potential to examine the direction of effect and avoid erroneous conclusions in structural panel models.