Reciprocal relationships, reverse causality, and temporal ordering: Testing theories with cross-lagged panel models

Charles C. Lanfear Thiago R. Oliveira

Abstract from panel; we may want to decide on specific illustrative case after we've laid out framework of paper, as we want our case to exhibit every potential issue.

Reciprocal causal relationships are a common feature of criminological theories. For example, police forces tend to use force more often in areas where crime concentrates, while at the same time legal cynicism theory suggests that cumulative exposures to police use-of-force can foster criminal activity. When multiple observations over time are available, cross-lagged panel models are commonly used to estimate these reciprocal effects. This is often done without careful attention to the assumptions that must be satisfied to produce valid estimates, such as correctly specified temporal lags, sufficient inter-temporal variation, and proper accounting for unobserved heterogeneity. Failure to satisfy these assumptions can produce severe issues including spurious associations and parameter estimates that are biased or even reversed in direction. In addition, reciprocal relationships violate causal assumptions based on graphical tools; criminological theories that suggest reciprocal causal relationships usually have an underlying macro-micro mechanism often not accounted for in empirical models. We provide guidance on how to align theory, model specification, and choice of estimator and illustrate this using an empirical example. We use data from Chicago at the census tract level and model the potential reciprocal relationship between police use-of-force and violent crime between 2004 and 2016. We finalise highlighting the importance of criminological theory and careful attention to empirical implications of theoretical premises when investigating reciprocal relationships

Introduction

- When do we expect reciprocal effects?
  - What are reciprocal effects *really*?
  - Separate theoretical and methodological concerns

- What must a model do to capture these?
- Use 3 wave model as example
- Might be worth taking a clear but extremely concise perspective on causality to ward off pedants
- Terminology
  - Strict exogeneity
  - Predetermined regressors (weakly exogenous, sequentially exogenous)
- Explanation of how what we're doing is different from recent reviews of CLPM literature, e.g., Zyphur et al. (2020).
  - Zyphur et al. elaborate a complete generalized cross-lagged panel model that accounts for a wider range of specifications. Here we focus on key problems that appear in applied literature.

Notation and diagrams

We should be specific and consistent in choosing graphical representations and notation.

DAGs are easy to understand but rule out contemporaneous reciprocal relationships. This is good in one respect, as causal effects are always time ordered. Perhaps we separate causal models from statistical models explicitly by using DAGs for causal models and SEM path diagrams for statistical models. Then using that, we talk about situations under which modeling something as contemporaneous is necessary.

Relatedly, something I've never seen done is using DAGs to really elaborate how we should think about waves of panel data as snapshots during temporal processes. For example, you might imagine a causal process like that below where we only observe at times 1, 3, and 5 despite the process operating over 1 unit lags. This is a discretization, of course, and you can imagine many processes are essentially continuous with hundreds of steps in between observations.

> In below diagrams, I'd prob make the observed ones x1, x2, x3, y1, y2, y3 and intervening ones something else, e.g., x1.5

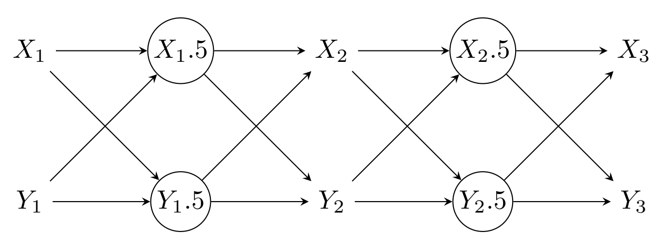

Figure 1: Directed acyclic graph representation of CLPM

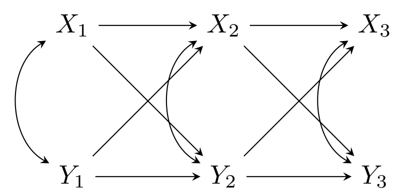

Figure 2: SEM path diagram representation of CLPM

- DAGs for theory, SEM path diagrams for estimation?

Motivating example

We've written up ESC using legal cynicism as motivating example, but if we find it to be too complicated, we could also use the Airbnb paper for ESC (doesn't make sense for JDLCC though). It is nice because it has a DAG to motivate an estimation strategy with contemporaneous effects:

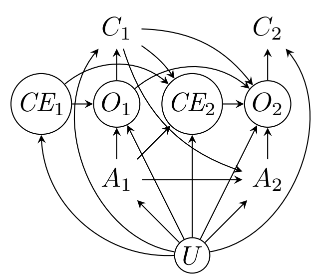

Figure 3: Theoretical model. CE is collective efficacy, O is criminal opportunity, C is crime, A is short-term lettings, U is omitted confounders

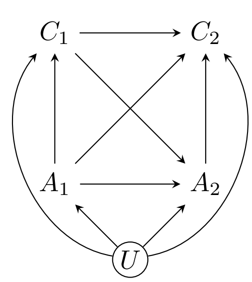

Figure 4: Estimation model. C is crime, A is short-term lettings, U is omitted time- and unit-specific confounders

Alternatively, we might select something with ambiguous contemporaneous directionality to motivate different estimation approaches, e.g., the Vaisey & Miles both effects approach or the

Allison unanalyzed correlation approach.

Issue 1: Temporal Order

- Unusual form of bias as documented by Vaisey & Miles that can flip signs:
  - If true model is $y = \beta x_t + \alpha_i + e_{it}$ and you fit $y = \beta^* x_{t-t} + \alpha_i + e_{it}$, $E(\beta^*) = -0.5\beta$, i.e., as Allison (2022) says "a positive contemporaneous effect of x gets transformed into an artifactually negative lagged effect".
- Probably the most underappreciated issue as it may produce misleading results that reject or overturn theories
  - Papers that find negative effects when theory suggests positive ones may be the result of this bias.

Solutions

- Theoretical ideal: Get it right
- Vaisey & Miles: Do both lagged and contemporaneous
  - If no contemporaneous effect, only lagged, you're probably fine. If contemporaneous shows up, you can't definitively determine direction of contemporaneous effect.
- Allison: Partial correlation model with lagged effect and contemporaneous residual covariance
  - This is least biased approach if parameter of interest is βx_{t-1}, but cannot estimate βx_t
  - This still does not adjudicate between direction of contemporaneous effects; if effect is entirely contemporaneous, all relationship will be in residual covariance.

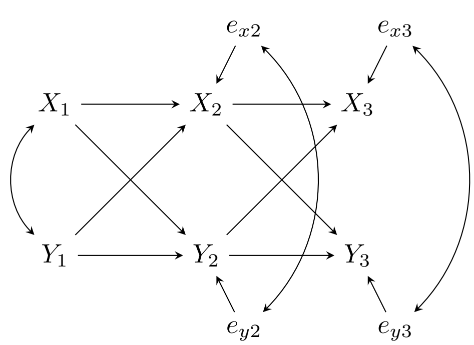

Figure 5: Allison (2022) recommendation

Issue 2: Unobserved heterogeneity

- Probably the most commonly encountered issue
- Might also raise treatment heterogeneity here as different but related issue not as easily addressed in CLPMs as in TWFE
  - Exception is time-specific treatment heterogeneity which is easily permitted by allowing wave-specific parameters as in Allison's model
  - A potentially useful application for wave-specific parameters is when you have unevenly spaced data; I think latent growth models would get at his too but I'm not very familiar with them.

Solutions

- RI-CLPM and related approaches
  - These are correlated stable trait models; I'm not very familiar
  - Seems to address time stable unit heterogeneity
- Allison's FE dynamic panel model
  - Can handle both under same assumptions and problems
  - Allison's ML-SEM approach is related to this Mundlak model; augments it with predetermined regressors.

Issue 3: Inter-temporal variation

- Maybe show some illustrative math here to show that $Var(Y_2|Y_1) \rightarrow 0$ as $\rho(Y_1, Y_2) \rightarrow 1$ approaches 1; i.e., you rapidly run out of anything to explain.
  - This is essentially a multicollinearity problem: Unbiased and consistent but high variance in estimates makes any particular estimate suspect.
- Rarely appreciated outside time series literature but common when waves are narrowly spaced
- With random measurement error, the error becomes proportionally larger as time periods get narrower
  - May mean you can address in part with measurement models
  - Note the endogenous variables are regressors, so measurement error biases parameter estimates as well

Solutions

- This is first and foremost a data problem, so you want to attack it during design of data collection if possible
  - Sample to maximize change
  - Look for shocks
- If you can't collect new data, look at different aggregations or wave skips
  - Small change at a large unit may mask larger changes in small units
  - Should prob only consider skipping waves if it makes sense theoretically and/or you're facing serious empirical underidentification problems.

General Advice

- Allison's FE model should prob be the default as it is most robust

Conclusion