A Multilevel Simultaneous Equations Model for Within-Cluster Dynamic Effects, with an Application to Reciprocal Parent-Child and Sibling Effects

Fiona Steele*
Jenny Jenkins†
Jon Rasbash*

*University of Bristol, †University of Toronto

Symposium on Recent Advances in Methods for the Analysis of Panel Data, Lisbon, 15-16 June 2011

Outline

- Background: within-family dynamics
- Methods for studying dyadic relationships
  - Bivariate growth curve model
  - Bivariate autoregressive cross-lagged model
- Our approach: generalising the cross-lagged model
- Application to maternal depression and child delinquency

Within-Cluster relationships

- Substantial interest in influences of one individual on another within a social group:
  - Classroom: peer effects on child performance
  - Workplace: relationships between employees
  - Family: sibling relationships
- Often individuals in groups have different roles:
  - Teacher/Student
  - Boss/Employee
  - Parent/Child
- Aim is to develop a model for studying influence of each individual on another over time, recognising diferent roles

Application to Within-Family Dynamics

- Family system made up of set of individuals interacting together over time
- So behaviour of all family members interdependent with behaviour of one member *causing* behaviour of another
  - Parent $\rightarrow$ child
  - Child $\rightarrow$ parent
  - Child $\rightarrow$ child (i.e. sibling effect)
  - Mother $\rightarrow$ father
  - Father $\rightarrow$ mother

Challenges for Causal Inference with Observational Data

- How to distinguish between causal effects attributable to different family members?
- How to disentangle effects of unmeasured individual and family characteristics?
  - E.g. apparent sibling effect could be due to shared family characteristics (genetic or environmental) influencing both children
- How to disentangle genetic and environmental influences without genetically-informative design?

Previous Research

- Focused on parent $\leftrightarrow$ child *or* child $\leftrightarrow$ child (not both)
- Dyadic relationships only
  - Parent $\leftrightarrow$ child based on 1 parent and 1 child
  - Child $\leftrightarrow$ child based on 2 children, and usually only older $\rightarrow$ younger 'training' effects
- No allowance for effects of unmeasured family characteristics

Methods for Studying Dyadic Relationships

Two main approaches:

- Bivariate growth curve model
- Bivariate autoregressive cross-lagged model

Both can be framed as an SEM or as a multilevel model:

- SEM useful when outcomes measured by multiple indicators
- MLM can handle clustering, e.g. family effects

Bivariate Linear Growth Curve Model: Notation

Responses

y_{tj}^P	response at time t for parent in family j
y_{tj}^C	response at time t of child in family j

Residuals

u_{0j}^P and u_{1j}^P	random intercept and slope effects for parent
u_{0j}^C and u_{1j}^C	random intercept and slope effects for child
e_{tj}^P and e_{tj}^C	time-varying parent and child residuals

Suppose $t = 0$ at first measurement occasion.

Bivariate Growth Curve Model

Consider linear model:

$$y_{tj}^P = (\alpha_0^P + u_{0j}^P) + (\alpha_1^P + u_{1j}^P)t + e_{tj}^P$$

$$y_{tj}^C = (\alpha_0^C + u_{0j}^C) + (\alpha_1^C + u_{1j}^C)t + e_{tj}^C$$

where $(u_{0j}^P, u_{1j}^P, u_{0j}^C, u_{1j}^C) \sim$ multivariate normal.

Interested in P-C covariances, e.g. $\mathrm{cov}(u_{0j}^P, u_{1j}^C) > 0 \implies$ children of parents with above-average y^P at $t = 0$ tend to show faster change in y^C.

BUT tells us about association between trajectories, not the *dynamic* relationship between parent and child.

Bivariate Autoregressive Cross-Lagged Model

Model for reciprocal dynamic effects between parent and child outcomes:

$$\begin{aligned} y_{tj}^P &= \beta_0^P + \beta_1^P y_{t-1,j}^P + \beta_2^P y_{t-1,j}^C + e_{tj}^P \\ y_{tj}^C &= \beta_0^C + \beta_1^C y_{t-1,j}^C + \beta_2^C y_{t-1,j}^P + e_{tj}^C \end{aligned}$$

where interest centres on the cross-lagged effects:

- Child to parent: β_2^P
- Parent to child: β_2^C

Our Approach

We extend the bivariate AR cross-lagged model:

- Allow simultaneously for parent $\leftrightarrow$ child and child $\leftrightarrow$ child effects
- Include families with different size sibships (including one-child families)
- Generalisable to mixtures of one and two parent families
- Separate occasion, individual and family effects

Illustrate method in application to maternal depression and child delinquency.

Notation

Consider family with 1 parent and 2 children.

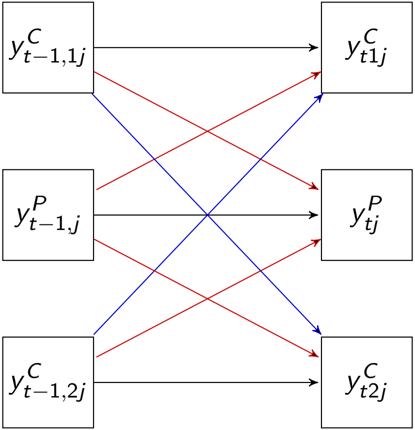
Responses

y_{tj}^{P}	response at time t for parent in family j
y_{tij}^{C}	response at time t of child i (=1,2) in family j

Residuals

e_{tj}^{P} and e_{tij}^{C}	time-varying parent and child
u_{ij}^{C}	time-invariant child
v_{j}^{P} and v_{j}^{C}	time-invariant family

Cross-lag SEM, 1 Parent and 2 Children, Times $t - 1$ and t

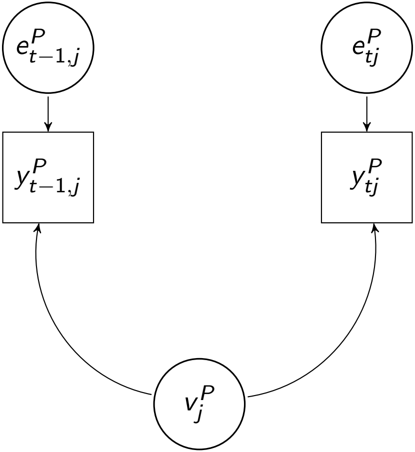


$\rightarrow$ individual lag

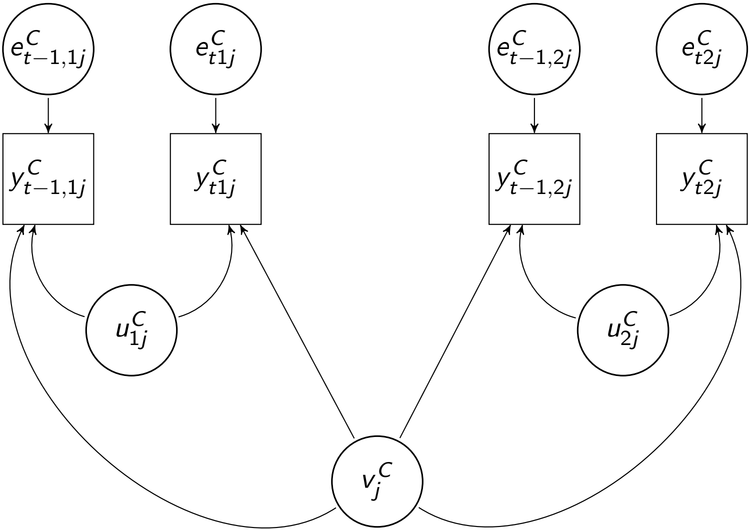
$\rightarrow$ sibling cross-lags

$\rightarrow$ parent $\leftrightarrow$ child cross-lags

Residual Structure of Multilevel SEM: Parent Model



Residual Structure of Multilevel SEM: Child Model



Basic Parent $\leftrightarrow$ Child Model

Multilevel model for 1 parent and 2 children:

$$
\begin{aligned}
y^P_{tj} &= \beta^P_0 + \beta^P_1 y^P_{t-1,j} + \beta^P_2 y^C_{t-1,+j} + v^P_j + e^P_{tj} \\
y^C_{tij} &= \beta^C_0 + \beta^C_1 y^C_{t-1,j} + \beta^C_2 y^P_{t-1,j} + v^C_j + u^C_{ij} + e^C_{tj}
\end{aligned}
$$

where $y^C_{t-1,+j} = y^C_{t-1,1j} + y^C_{t-1,2j}$

Assume $(v^P_j, v^C_j) \sim$ bivariate normal to allow for unmeasured family characteristics affecting both parent and child outcomes.

Allowing for Sibling (Child $\leftrightarrow$ Child) Effects

Add sum of lagged responses for siblings of child i to model for y^C_{tij}.

Initial assumptions about parent $\leftrightarrow$ child and child $\leftrightarrow$ child effects:

- Each child has same effect on the parent
- Parent has same effect on each child
- Each child has same effect on each sibling

Can relax assumptions to allow each effect to depend on characteristics of child (e.g. age, sex), sibling pair (e.g. age difference) or parent/family (e.g. SES).

Initial Conditions

A problem when start of measurement does not coincide with start of process under study.

- Unmeasured time-invariant factors influencing $y_2, \ldots, y_T$ are likely to also influence y_1, leading to correlation between y_1 and random effects
- Can show that in an AR(1) model, the dependence of y_t on previous y operates entirely through y_1
- A solution is to specify a model for y_1 and estimate jointly with model for $y_2, \ldots, y_T$

Modelling the Initial Condition

Consider the model for the parent's outcome.

For $t > 1$:

$$y_{tj}^{P} = \beta_0^{P} + \beta_1^{P} y_{t-1,j}^{P} + \beta_2^{P} y_{t-1,+j}^{C} + v_j^{P} + e_{tj}^{P}$$

For $t = 1$:

$$y_{1j}^{P} = \alpha_0^{P} + \lambda_v^{P} v_j^{P} + e_{1j}^{P}$$

So a common set of unmeasured individual characteristics v_j^P influences y_1 and y_t given y_{t-1} $(t > 1)$, but their effects can differ.

Estimation

- Multilevel SEM is a type of multivariate response model, but need flexibility to allow for different hierarchical structures for parent and children
- Structure data in ‘long’ form with 1 record per occasion per family member
- For mum and children observed at 3 occasions, define 4 dummy variables: **mumt1**, **mumt23**, **kidt1** and **kidt23**
- Interact dummies with covariates to estimate equations for mum and child at $t = 1$ and $t = 2, 3$
- Options include MLwiN and aML

Application to Maternal Depression and Child Delinquency

- Avon Brothers and Sisters Study (ABSS): 175 families, 416 children, 1381 measurements
- 3 waves spaced 2 years apart
- Parent outcome (y^P): maternal depression (malaise inventory)
- Child outcome (y^C): delinquency (child behaviour checklist)
- All measures based on maternal report

Child Effects on Maternal Depression

	Model 1	Model 2
Lag child delinquency	-0.004	-0.058*
Family-level correlation		
$\text{corr}(v_j^P, v_j^C)$	0	0.710***

Model 1: equations for y^P and y^C separately estimated
Model 2: equations estimated simultaneously

Notes: (i) * $p < 0.10$, *** $p < 0.001$; (ii) adjusting for maternal lags, time and family size; (iii) outcome has mean of 0.2 and SD of 0.15.

Mother and Sibling Effects on Child Delinquency

	Estimation of y^P and y^C equations	
	Separate	Simultaneous
Girl	0.023	0.025
Mother effects		
Lag maternal depression	0.194***	−0.005
Lag maternal depression × girl	−0.232***	−0.244***

Tested for sibling effects, allowing effect to depend on birth order and age difference, but not significant.

Notes: (i) *** p $<$ 0.001; (ii) adjusting for child lags, time and family size; (iii) outcome has mean of 0.1 and SD of 0.15.

Further Investigation of Sibling Effects

Previous research has found ‘training’ effects from older to younger child.

Standard SEM includes a single residual term, while multilevel approach decomposes residual variation into occasion, individual and family components.

Compare standard SEM with multilevel SEM in analysis of sibling pairs (2-child families only).

Sibling Effects on Child Delinquency

	Standard SEM	Multilevel SEM
Lag younger sib y †	0.155	−0.182
Lag younger sib × | age diff |	0.044	0.047
Lag older sib y †	0.248**	−0.030
Lag older sib × | age diff |	−0.015	−0.004

† Age difference centred at 3 years.

So apparent training effect explained by shared dependency of both siblings' behaviour on unmeasured family characteristics.

Note: Both models allow for residual correlation between siblings at any t.

Discussion

- Important to jointly model parent and child outcomes, especially when using single-informant data
- Important to allow for unmeasured family characteristics
- Valuable to apply methods to more comprehensive data: larger sample size, more measurements, closer together in time, and from multiple informants