Planned Missing Data in Longitudinal Dyadic Modeling

Alexander M. Schoemann [1] John K. Sakaluk [2]

[1]East Carolina University

[2]University of Victoria

Developmental Methods 2016

Overview

- Dyadic data analyses
- Missing data handling
- Planned missing data designs with dyadic data

Dyadic Data

- Data collected from two individuals
  - Usually individuals have a social relationship
  - Usually measure the same variables on each member of the dyad
  - Responses within dyad are not independent

Dyadic Data

- Distinguishability – whether dyad members can be "told apart"
  - AKA: Exchangeability
  - Distinguishable dyads (non-exchangeable): dyad where each member has a unique role
    - Heterosexual couples, parent and child, older and younger siblings
  - Indistinguishable dyads (exchangeable): dyad where both members have the same role
    - Homosexual couples, twins, friends, coworkers

Data Structures for dyadic data

- Three structures for dyadic data from a standard design
  - Individual (long)
  - Dyad (wide)
  - Pairwise
- Choice of the data structure depends on the analysis technique and the type of dyad

Individual data structure

- Each row represents an individual's score
  - There is a variable representing dyad membership
  - Between dyad variables are entered twice (once on each row)

Individual data structure

```
##   d p x y z
## 1 1 1 5 9 3
## 2 1 2 2 8 3
## 3 2 1 6 3 7
## 4 2 2 4 6 7
## 5 3 1 3 6 5
## 6 3 2 9 7 5
```

Dyad data structure

- Each row represents a dyad
  - Responses from different members of the dyad are in different variables

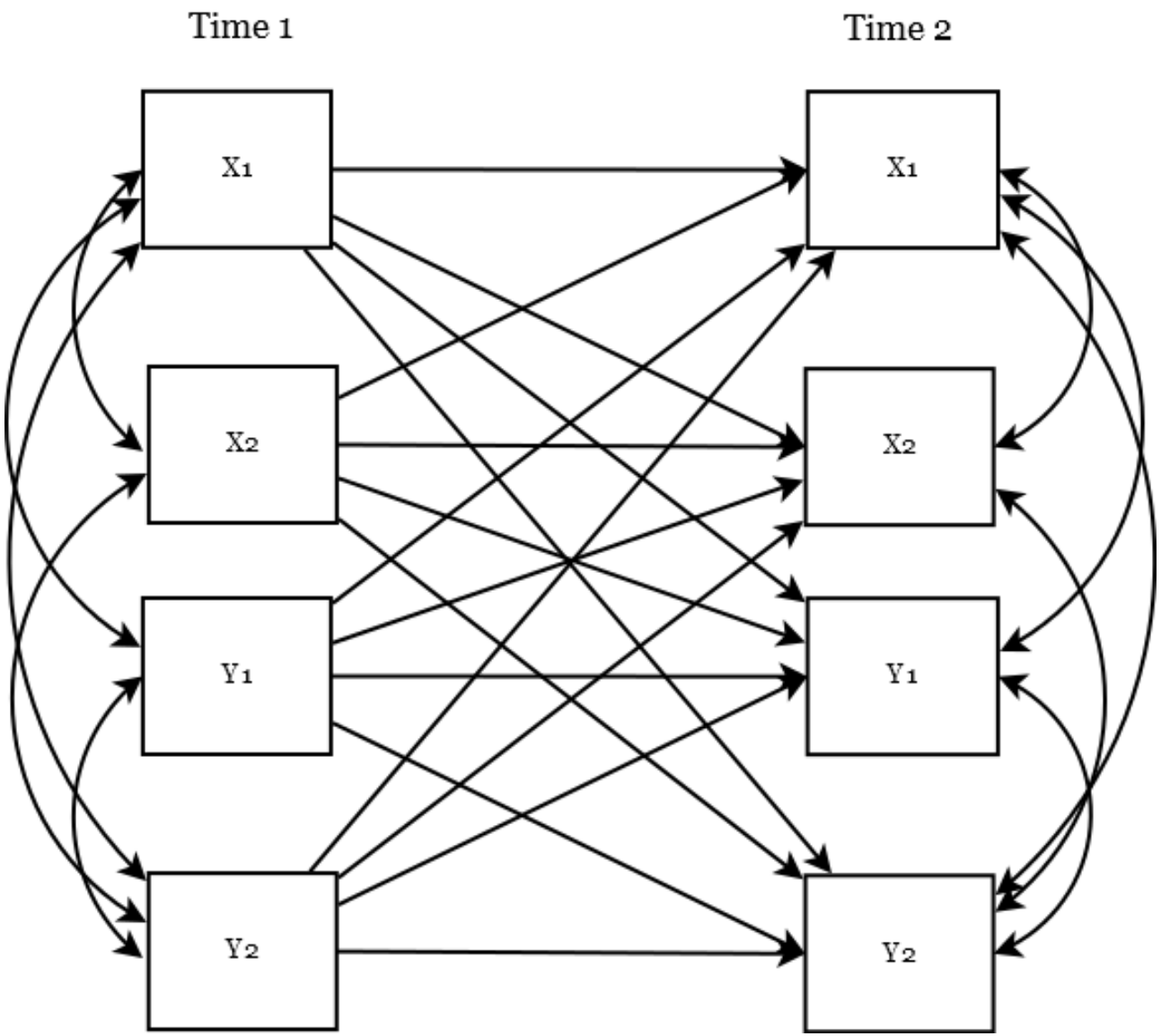
Dyad data structure

```
##   d x.1 y.1 z.1 x.2 y.2 z.2
## 1 1   5   9   3   2   8   3
## 3 2   6   3   7   4   6   7
## 5 3   3   6   5   9   7   5
```

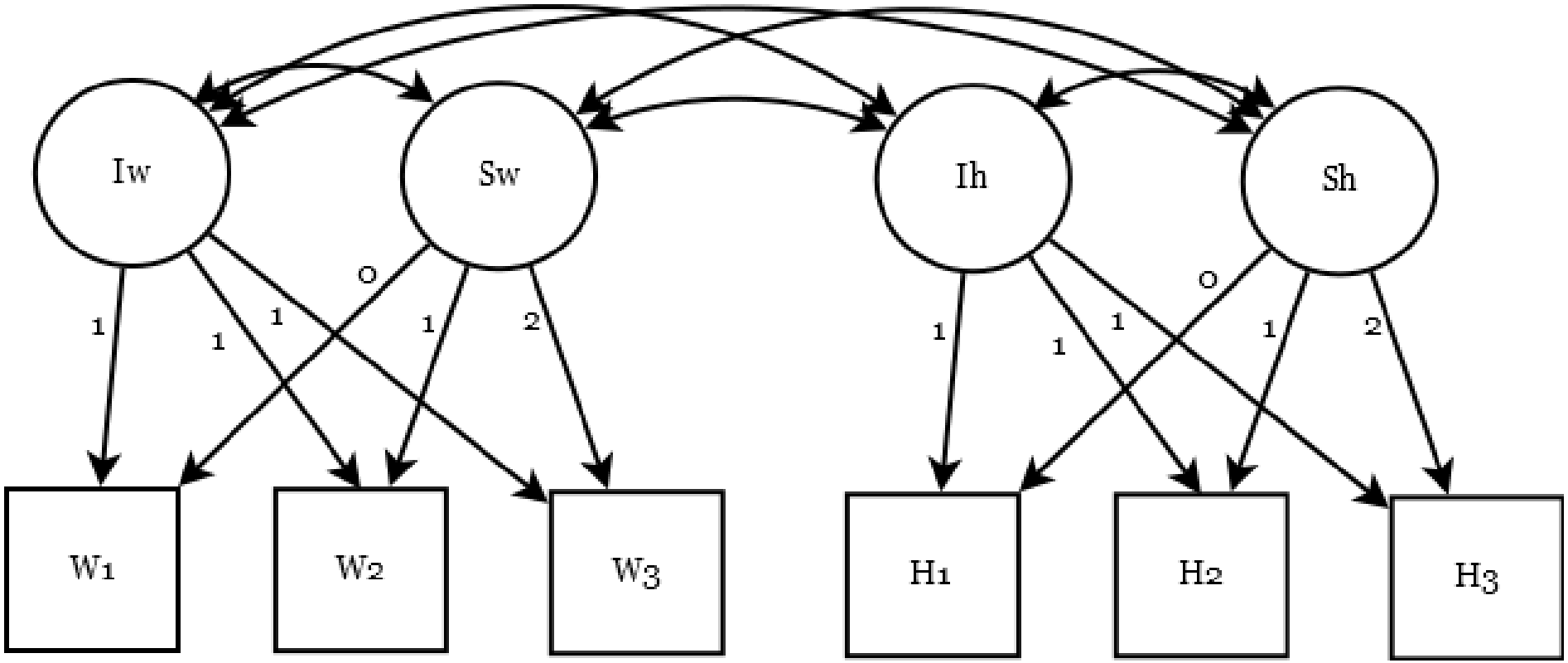
Longitudinal Dyadic Models

- Common models with dyads include:
  - Cross-lagged panel model Actor-Partner Interdependence Model (CP-APIM)
  - Growth curve models

CP-APIM



Growth Curve Model



Missing data in dyadic data

- Rarely discussed by dyadic researchers
  - Missing data strategies only mentioned in about 30% of dyadic papers
  - Deletion strategies tend to dominate (followed by FIML)
- Dependence in dyadic data provides special challenges with missing data.
  - Techniques need to incorporate distinguishability when recovering missing data

Patterns of missingness in longitudinal dyadic data

- Three patterns of missingness
  1. Missing data by item
  2. Missing data by time
  3. Missing data by person
     - More on this later!

Planned missing data

- Missing data do not have to be a problem!
- Two types of planned missing data designs:
  - Time based planned missing data designs
    - Control participant entry into the study (e.g., cohort sequential design)
  - Participant based planned missing data designs
    - Randomly assign participants to receive only a subset of items

Planned missing data

- For dyadic data both planned missing data designs can be used
  - Participant based designs to reduce questionnaire length (e.g. 3-Forms planned missing data designs)
  - Time based planned missing data designs (e.g. control when dyads are measured in a longitudinal study)

Planned missing data

- A third type of planned missing data design is possible with dyadic data: dyad based planned missing
  - Some dyads have data from both members
  - Some dyads have data only collected from one dyad member
- This design can lead to cost savings/power increases compared to assessing all dyad members

Dyadic planned missing data

Dyad	Person 1	Person 2
1	X	X
2	X	X
3	X	O
4	O	X

Dyadic planned missing data

- Missing data in this designs can be assigned or naturally occurring
  - When missingness is assigned (dyads are randomly assigned to have 1 or 2 members measures) missingness is MCAR
  - When missingness is natural (only 1 member of a dyad responds) missingness is (probably) MAR or MNAR
    - Researchers should measure dyad/partner variables related to non-response

Dyadic planned missing data

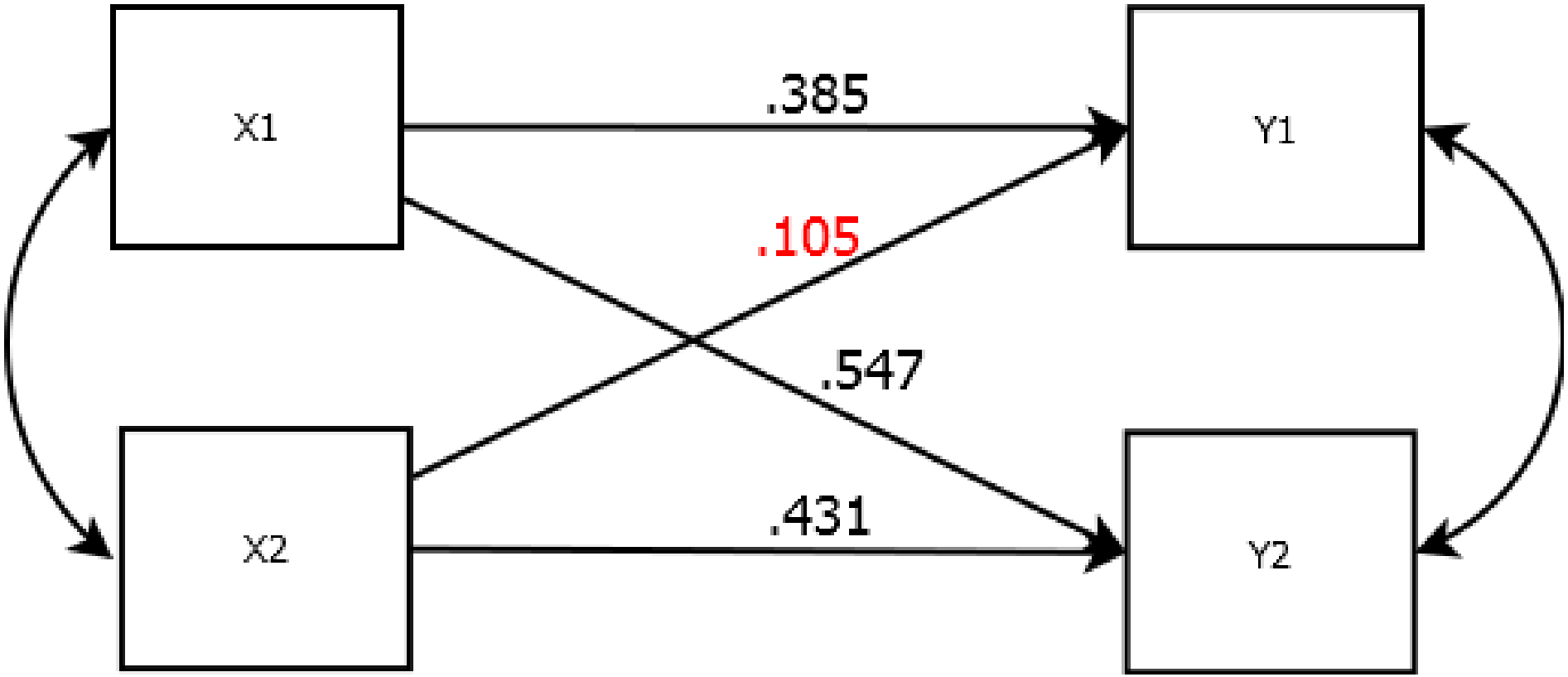
- Missing data should (if possible) be balanced across individuals
  - Equal missing for both members in distinguishable dyads
  - Missingness equally distributed across members for indistinguishable dyads
    - Data management

Dyadic planned missing data

- Example power analysis with MCAR planned missing data
  - Total budget $10000
    - Dyads: $50 per dyad
    - Singles: $10 per person
    - With no planned missing n = 200 dyads

Dyadic planned missing data

- Population model



Dyadic planned missing data

N dyad	N individuals	Power
200	0	.637
190	50	.643
176	120	.667
111	445	.552

Longitudinal dyadic planned missing data

- Longitudinal designs provide additional complexity with planned missing data designs
  - Different members of a dyad can be missing at different times of measurement
  - Time and dyad planned missing data can be integrated into the same study
    - Accelerated longitudinal designs with dyad planned missing

Dyadic planned missing data

Dyad	P1 T1	P2 T1	P1 T2	P2 T2
1	X	X	O	X
2	X	O	X	X
3	X	O	O	X
4	O	X	X	O

Longitudinal dyadic planned missing data

- Recommendations
  - Alternate which dyad member is missing across time points
    - e.g. if husband is missing at time 1, wife would be missing at time 2
  - Try to assess all dyads together for at least one time point

Future directions

- Provide guidance on dyadic planned missing data designs
  - Ratio of dyads to singles
  - Distribution of singles across dyad members
  - Distribution of singles across time

Thank you!

- Slides from today at:

`http://MARlab.org/Supplemental_Materials/`

- email: `schoemanna@ecu.edu`

Simulation population values I

```
## This is lavaan 0.5-20

## lavaan is BETA software! Please report any bugs.

##                lhs op           rhs label    est
## 1   SATISFACTION  ~    ACT_HOUSE     a -0.591
## 2   SATISFACTION  ~   PART_HOUSE     p  0.888
## 3         PSATIS  ~    ACT_HOUSE     p  0.888
## 4         PSATIS  ~   PART_HOUSE     a -0.591
## 5   SATISFACTION ~~ SATISFACTION    v2  2.382
## 6         PSATIS ~~       PSATIS    v2  2.382
## 7      ACT_HOUSE ~~    ACT_HOUSE    v1  1.060
## 8     PART_HOUSE ~~   PART_HOUSE    v1  1.060
## 9      ACT_HOUSE ~~   PART_HOUSE        0.417
## 10 SATISFACTION ~1                int1  4.791
```

Simulation population values II

```
## 11        PSATIS ~1                    int1  4.791
## 12     ACT_HOUSE ~1                    int2  1.630
## 13    PART_HOUSE ~1                    int2  1.630
## 14 SATISFACTION ~~       PSATIS              0.812
```