Dynamic Structural Equation Modeling of Intensive Longitudinal Data Using Multilevel Time Series Analysis in Mplus Version 8 Parts 3 and 4

Bengt Muthén
bmuthen@statmodel.com

Tihomir Asparouhov & Ellen Hamaker

Workshop at Johns Hopkins University, August 17 - 18, 2017

We thank Noah Hastings for excellent assistance

Mplus Version 8: Methods for Analyzing Intensive Longitudinal Data

- Time series analysis (N = 1)
- Two-level time series analysis (N > 1)
  - Random effects varying across subjects (subject is level 2, so many more random effects than usual)
- Cross-classified time series analysis
  - Random effects varying across subjects and time
- Dynamic Structural Equation Modeling (DSEM)
  - General latent variable modeling
  - Bayesian estimation
  - Statistical background:
    - Asparouhov, Hamaker & Muthén (2017). Dynamic structural equation models. Technical Report, www.statmodel.com
    - Asparouhov, Hamaker & Muthén (2017). Dynamic latent class analysis. Structural Equation Modeling, 24, 257-269

The Version 8 Mplus User's Guide adds N=1 examples 6.23 - 6.28 and N > 1 examples 9.30 - 9.40, many with two parts (basic and advanced). Also webpage

Outline

- **Introduction to Bayesian analysis**
- Introduction to longitudinal analysis
- N=1 time series analysis
- Two-level time series analysis
- Cross-classified time series analysis
- Latent variable time series analysis

Prelude 1: Bayesian Analysis - A Thumbnail Sketch

All that's needed:

ESTIMATOR = BAYES;

- Bayesian advantages over ML
- An example: Estimating a mean
- Convergence of Bayes iterations
- Trace and autocorrelation plots
- Speed of Bayes in Mplus
- Bayes references

Bayesian Analysis: Advantages over ML

- Six key advantages of Bayesian analysis over frequentist analysis using maximum likelihood estimation:
  1. More can be learned about parameter estimates and model fit
  2. Large-sample theory is not needed and small-sample performance is better
  3. Parameter priors can better reflect results of previous studies
  4. Analyses are in some cases less computationally demanding, for example, when maximum-likelihood requires high-dimensional numerical integration
  5. In cases where maximum-likelihood computations are prohibitive, Bayes with non-informative priors can be viewed as a computing algorithm that would give essentially the same results as maximum-likelihood if maximum-likelihood estimation were computationally feasible
  6. New types of models can be analyzed where the maximum-likelihood approach is not practical (e.g. DSEM)

Why Are Bayesian Computations Possible Where ML Computations Are Not?

The general modeling features of DSEM make ML almost impossible, creating the need for Bayesian estimation.

An intuitive description of the computational difference between ML and Bayes (with non-informative priors):

- ML works with the joint distribution of all variables to find the parameter values that give the logL maximum
- Bayes works with a series of conditional distributions for the parameters to get (posterior) parameter distributions
- The joint distribution can be difficult to describe whereas the conditional distributions can be easier
- Bayes is sometimes the only feasible alternative when the joint distribution is hard to formulate

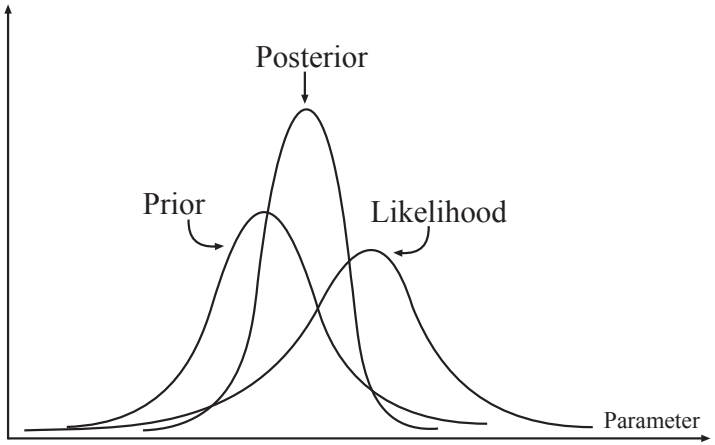


Figure: Informative prior

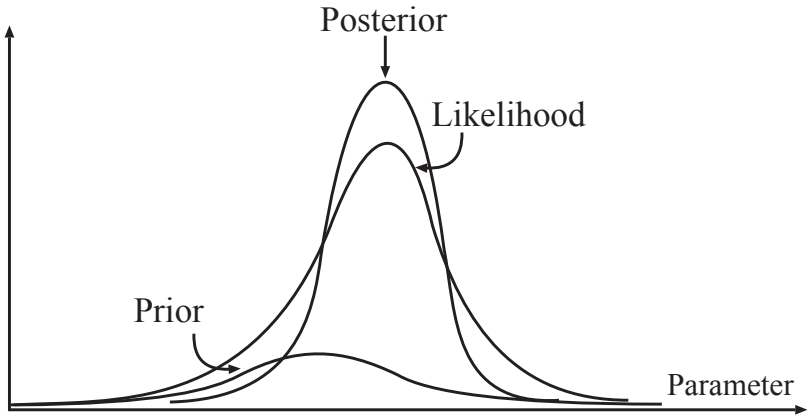


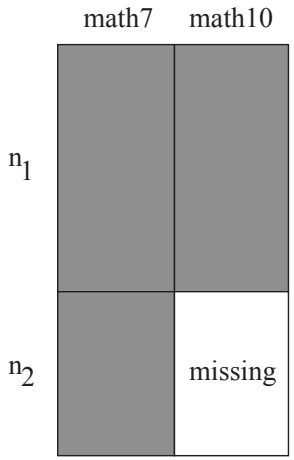
Figure: Non-informative prior

- Priors:
  - Non-informative priors (diffuse priors): Large variance (default in Mplus)
    - A large variance reflects large uncertainty in the parameter value. As the prior variance increases, the Bayesian estimate gets closer to the maximum-likelihood estimate
  - Weakly informative priors: Used for technical assistance
  - Informative priors:
    - Informative priors reflect prior beliefs in likely parameter values
    - These beliefs may come from substantive theory combined with previous studies of similar populations

Bayes' Advantage Over ML: Informative Priors

- Frequentists sometimes object to Bayes using informative priors
  - But they already do use such priors in many cases in unrealistic ways (e.g. factor loadings fixed exactly at zero)
- Bayes can let informative priors reflect prior studies
- Bayes can let informative priors identify models that are unidentified by ML which is useful for model modification (BSEM)
- The credibility interval for the posterior distribution is narrower with informative priors

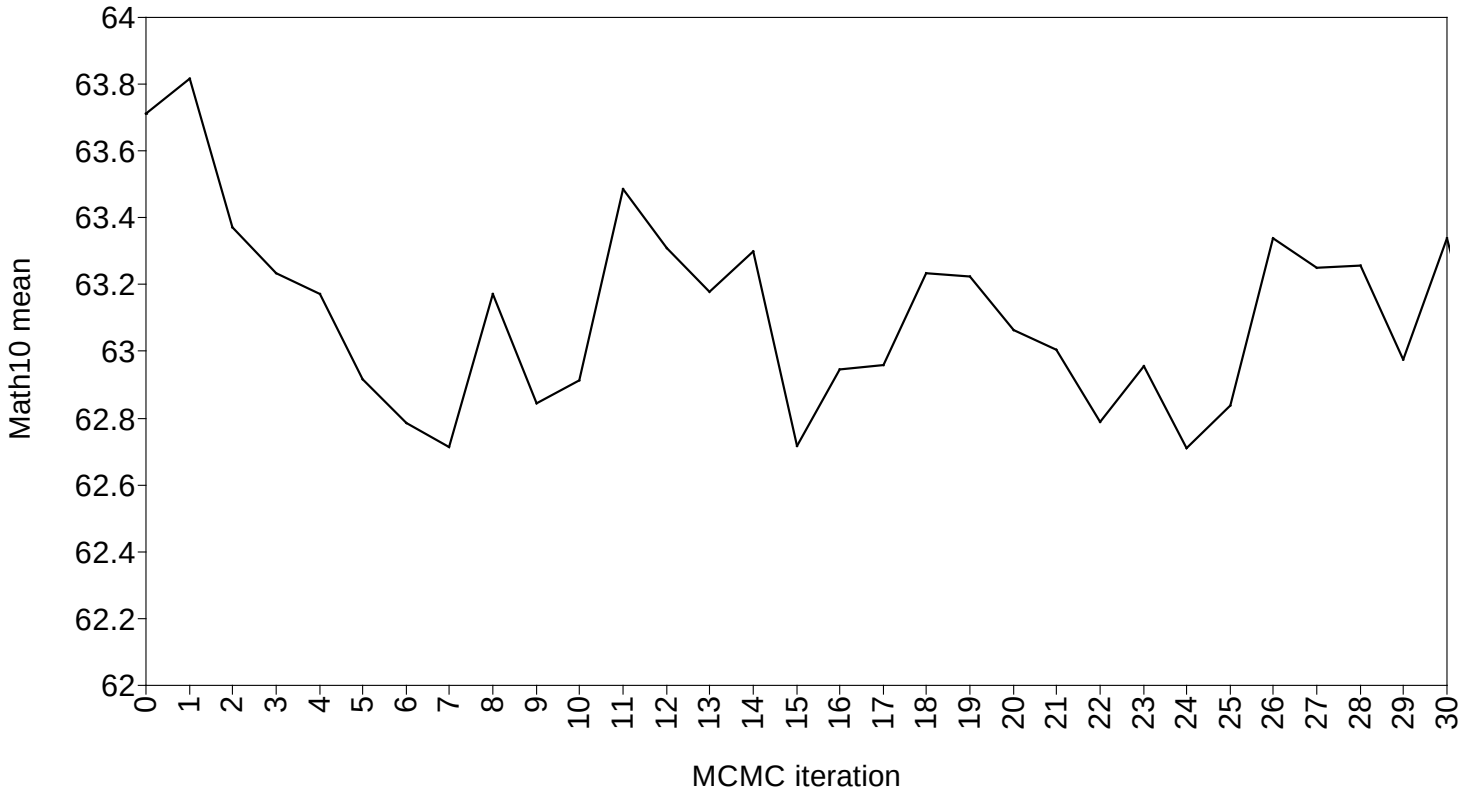
An MCMC Example: LSAY Math with Missing Data



- Three sets of unknowns assuming bivariate normality:
  - 2 means
  - 2 variances, and 1 covariance
  - n_2 missing values on math10

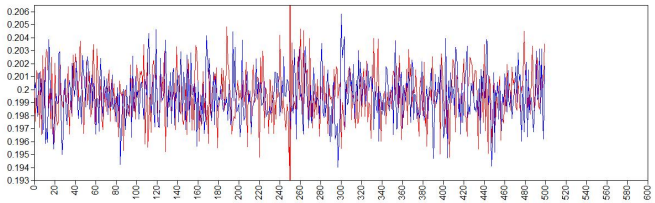
1. **Draw values for the two means** from the conditional distribution of the means conditioned on the variance-covariance parameters, the observed and missing data, and the priors.
2. **Draw values for the n_2 missing values on math10** from the conditional distribution of missing values conditioned on the mean parameters, the observed data, and the priors.
3. **Draw values for the two variance and covariance parameters** from the conditional distribution of the variance-covariance parameters conditioned on the mean parameters, the observed and missing data, and the priors.

Trace Plot: 20 MCMC Iterations for the LSAY Math10 Mean

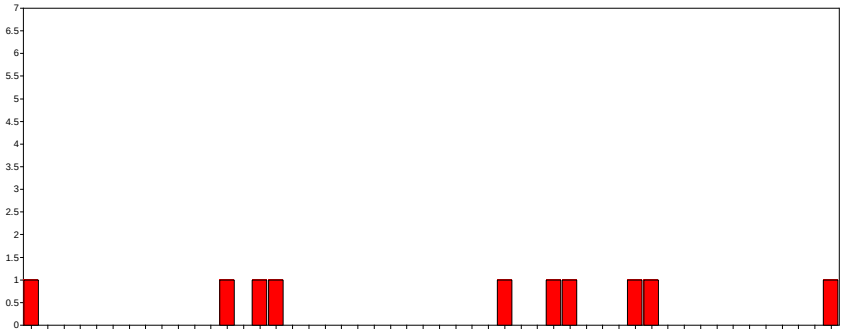


- Starting value for the mean is the listwise estimate of 63.7

Forming the Posterior Distribution of a Parameter Estimate

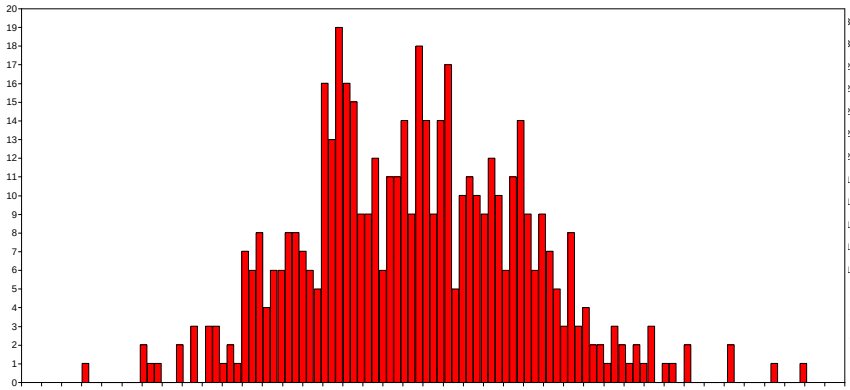


Posterior Distributions of the LSAY Math10 Mean Using Different Number of MCMC Iterations

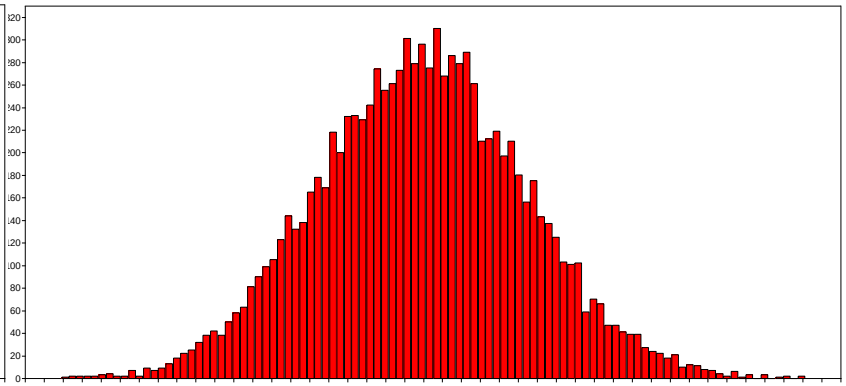


(a) 10 iterations

(b) 100 iterations

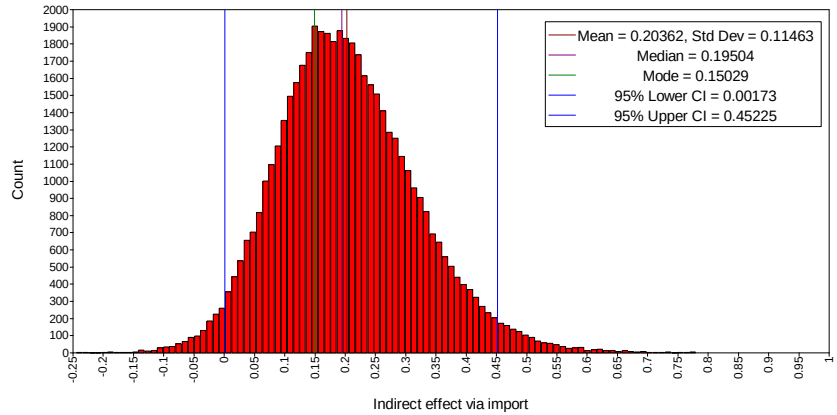


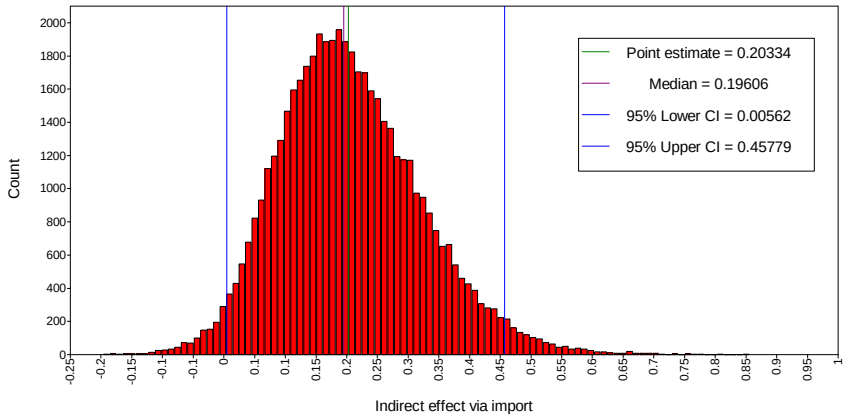
(c) 500 iterations



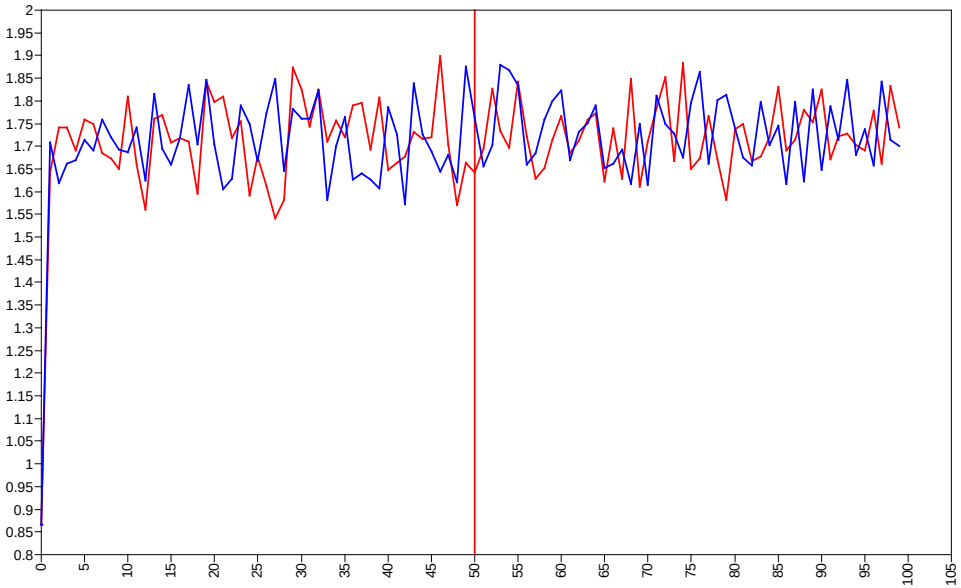
(d) 10,000 iterations

Bayes Posterior Distribution Similar to ML Bootstrap Distribution: Credibility versus Confidence Intervals





Convergence: Trace Plot for Two MCMC Chains. PSR



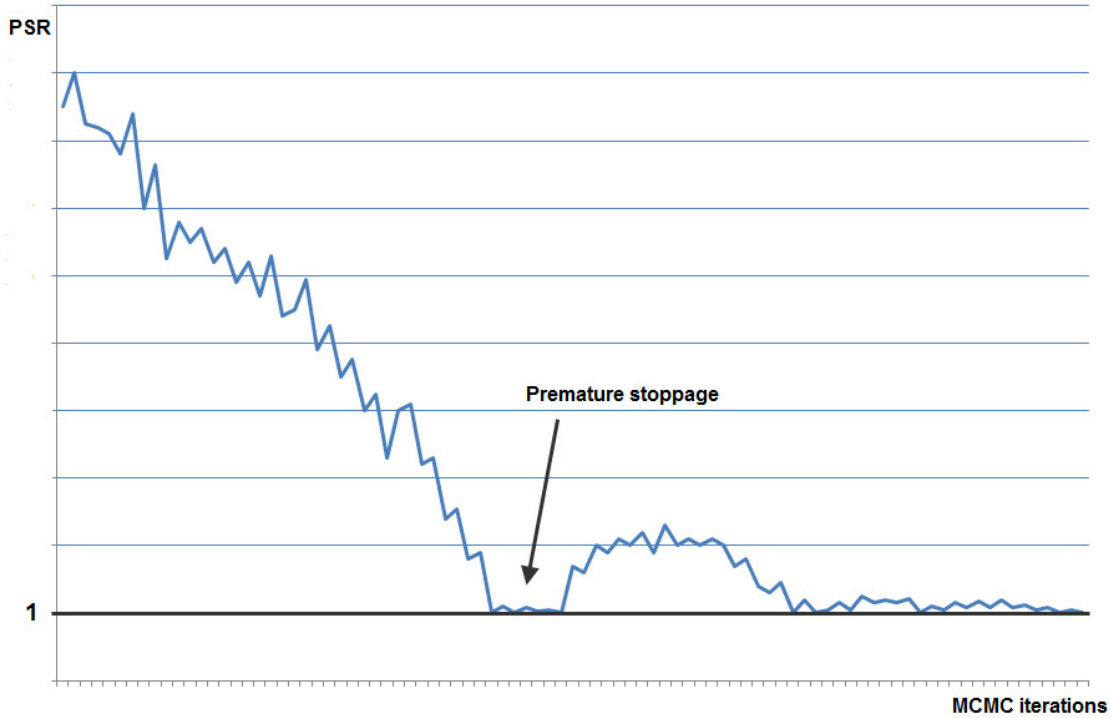
Potential scale reduction criterion (Gelman & Rubin, 1992):

$$PSR = \sqrt{\frac{W+B}{W}}, \qquad (1)$$

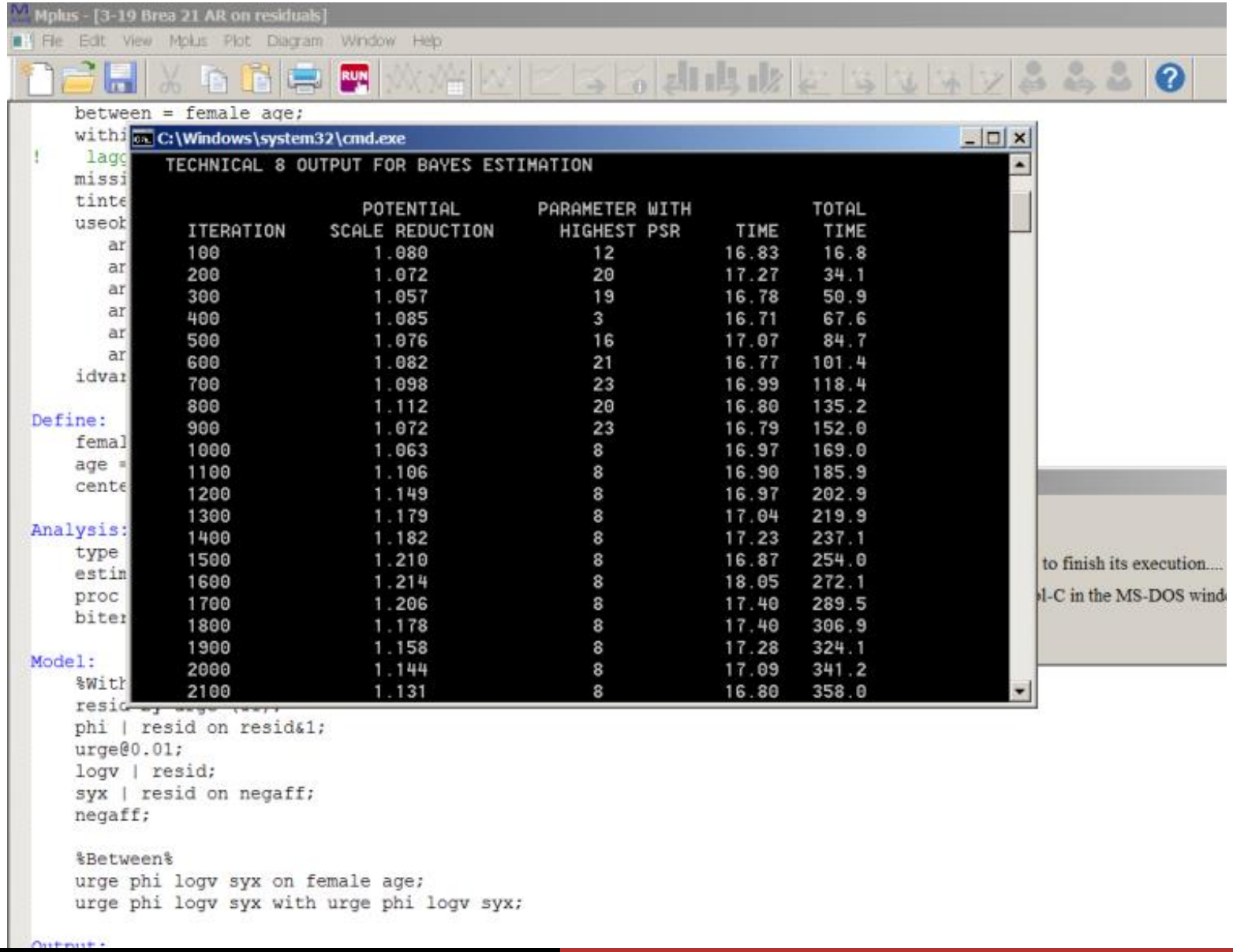
where W represents the within-chain variation of a parameter and B represents the between-chain variation of a parameter. **A PSR value close to 1 means that the between-chain variation is small relative to the within-chain variation** and is considered evidence of convergence.

Convergence of the Bayes Markov Chain Monte Carlo (MCMC) Algorithm

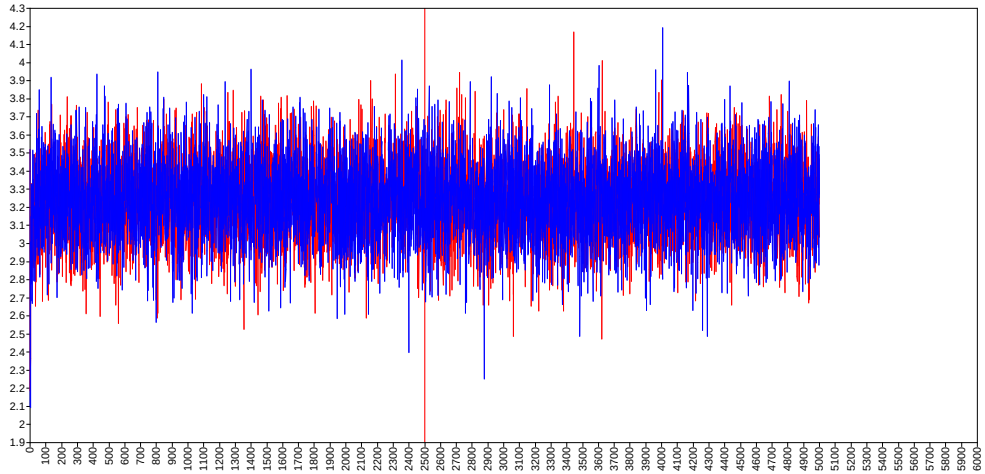
Figure: Premature stoppage of Bayes MCMC iterations using the Potential Scale Reduction (PSR) criterion

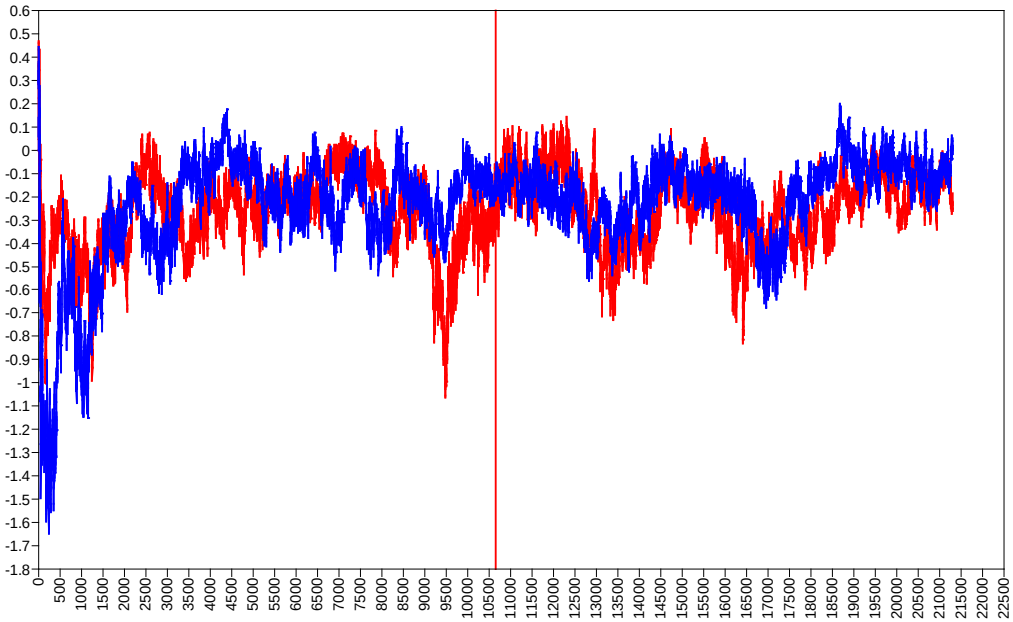


TECH8 Screen Printing of Bayes MCMC Iterations

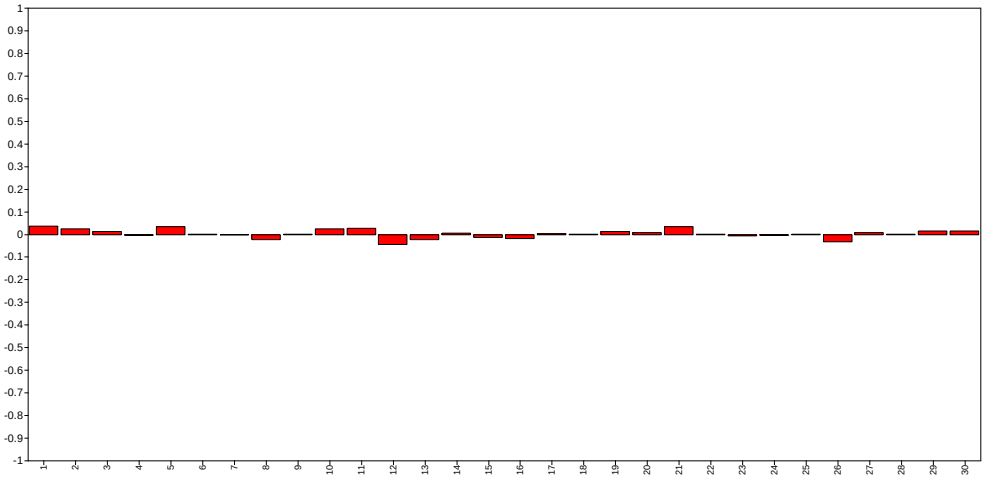


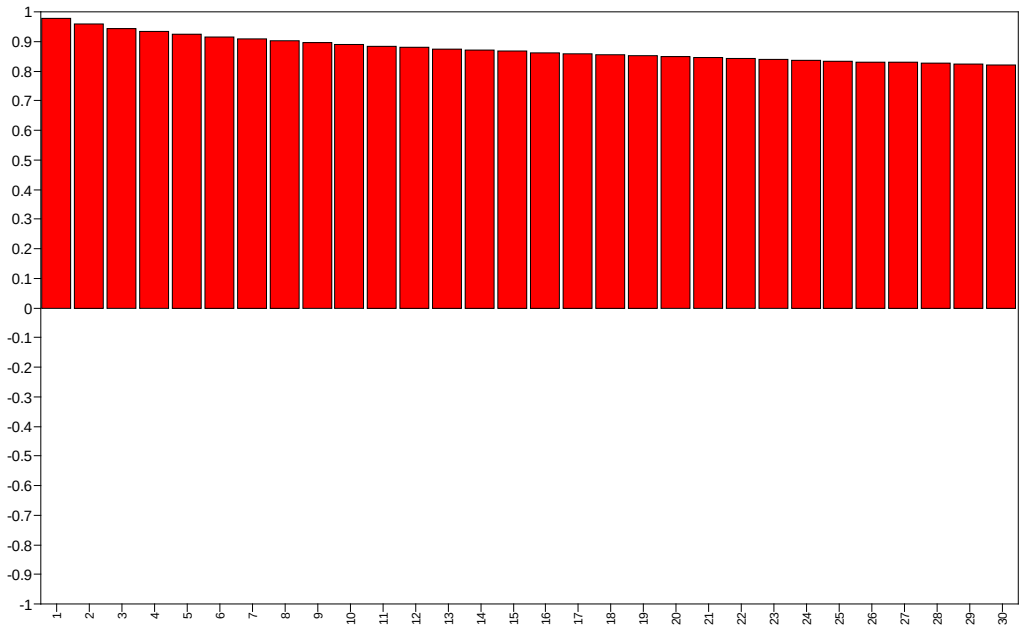
Trace Plots Indicating Good vs Poor Mixing





Autocorrelation Plots Indicating Good vs Poor Mixing





Speed Of Bayes In Mplus

Wang & Preacher (2014). Moderated mediation analysis using Bayesian methods. *Structural Equation Modeling*.

- Comparison of ML (with bootstrap) and Bayes: Similar statistical performance
- Comparison of Bayes using BUGS versus Mplus: Mplus is 15 times faster

- Reason for Bayes being faster in Mplus:
  - Mplus uses Fortran (fastest computational environment)
  - Mplus uses parallel computing so each chain is computed separately
  - Mplus uses the largest updating blocks possible - complicated to program but gives the best mixing quality
  - Mplus uses sufficient statistics when possible
- Mplus Bayes considerably easier to use

Nevertheless - It's Going To Be Slower Than Usual: Timings For The Runs In This Talk

Using smoking data with N = 230, T $\approx$ 150

- N=1 analysis of subject 227: 0 seconds
- First two-level analysis: 3:54
- Cross-classified analysis spotting a trend: 1:10
- Two-level trend analysis: 4:42
- Cross-classified trend analysis: 34 minutes
- Cross-classified ordinal factor analysis: 54 minutes (dichotomous 34 mins, continuous 16 mins)

Bengts PC as of June 2012: Dell XPS 8500, i7-3770 with 8 processors, CPU of 3.40 GHz, 12 GB RAM, 64-bit.

Bayes References

- Gelman et al. (2014). *Bayesian Data Analysis*, 3rd edition
- Lynch (2010). *Introduction to Applied Bayesian Statistics and Estimation for Social Scientists*
- Bayes technical reports on the Mplus website: See www.statmodel.com under Papers, Bayesian Analysis
- Muthén (2010). Bayesian analysis in Mplus: A brief introduction. Technical Report. www.statmodel.com
- Chapter 9 of Muthén, Muthén & Asparouhov (2016). *Regression and Mediation Analysis using Mplus*

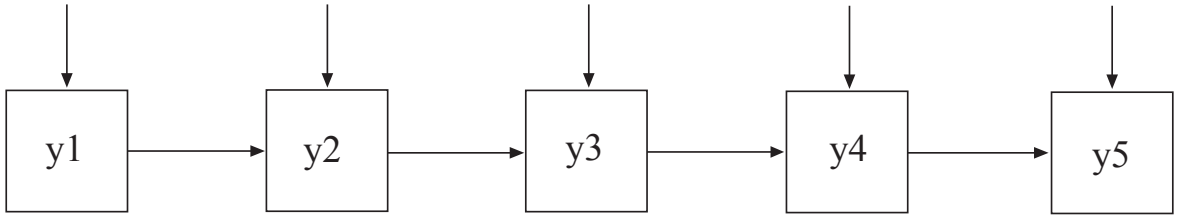
Outline

- Introduction to Bayesian analysis
- **Introduction to longitudinal analysis**
- N=1 time series analysis
- Two-level time series analysis
- Cross-classified time series analysis
- Latent variable time series analysis

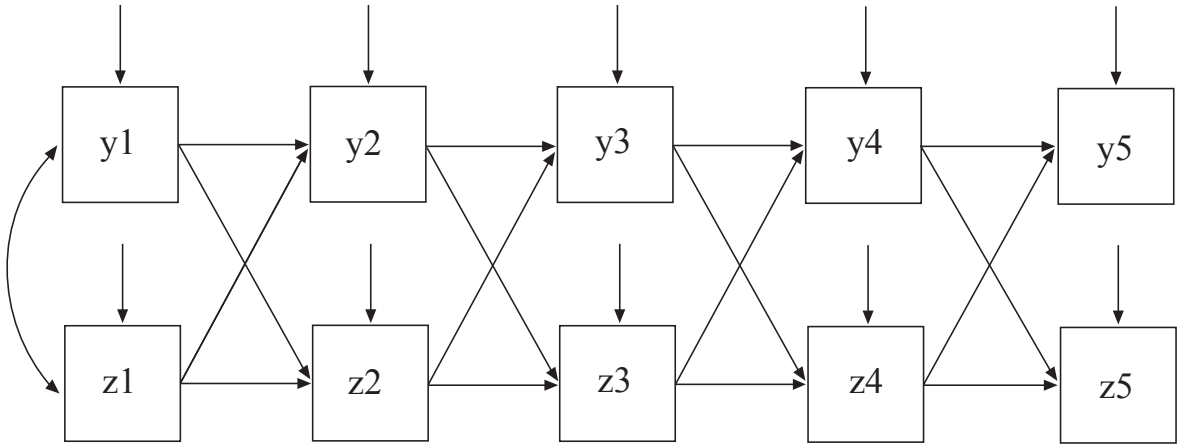
Prelude 2: Methods for Longitudinal Data

- Non-intensive longitudinal data:
  - T small (2 - 10) and N large
  - Modeling: Auto-regressive (cross-lagged) and growth modeling
- Intensive longitudinal data:
  - T large (30-200) and N smallish (even $N = 1$) but can be $1,000$. Often $T > N$
  - Modeling: We shall see

Common Methods for Non-Intensive Longitudinal Data
N large and *T* small (2 - 10):
(1) Auto-Regressive Modeling



Cross-lagged modeling (e.g. y = urge, z = negative affect):

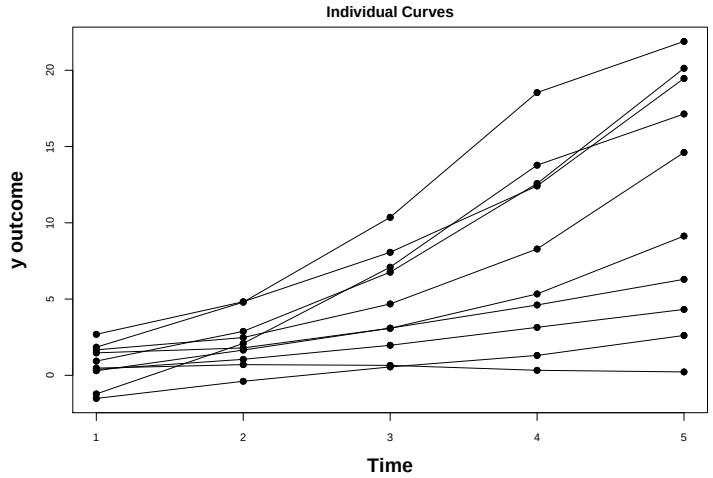


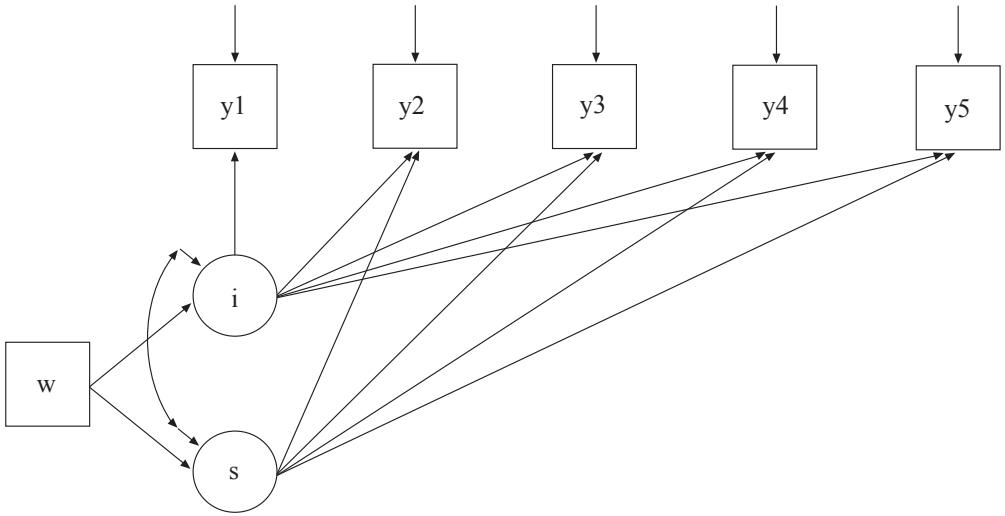
Recent References for Cross-Lagged Modeling

Extensions of the classic cross-lagged panel model:

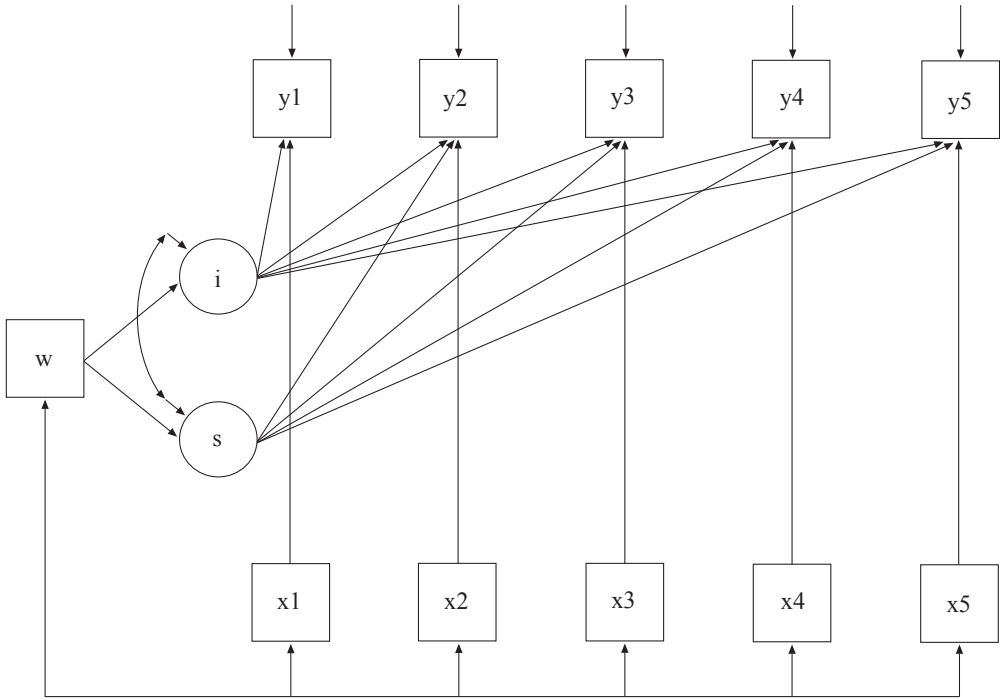
- Hamaker et al., *Psych Methods* 2015: The random intercepts cross-lagged panel model
- Curran et al., *J of Consulting & Clinical Psych* 2014: The separation of between-person and within-person components
- Berry and Willoughby, *Child Development* 2016: Rethinking the cross-lagged panel model (growth model added)
  - These models are fitted in Mplus

Common Methods for Non-Intensive Longitudinal Data: (2) Growth Modeling





Growth Modeling with Time-Varying Covariates

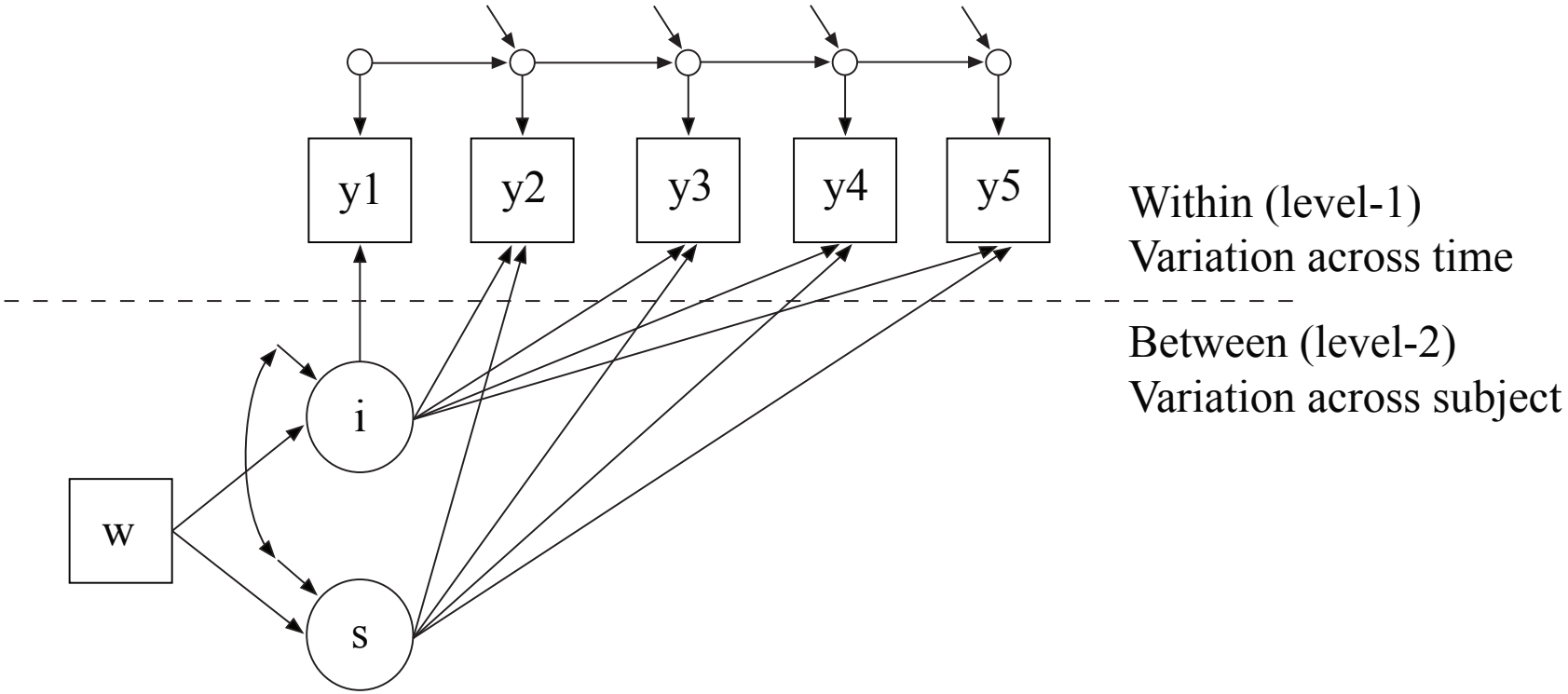


Why Is Regular Growth Modeling Not Sufficient For ILD?

There are 2 problems:

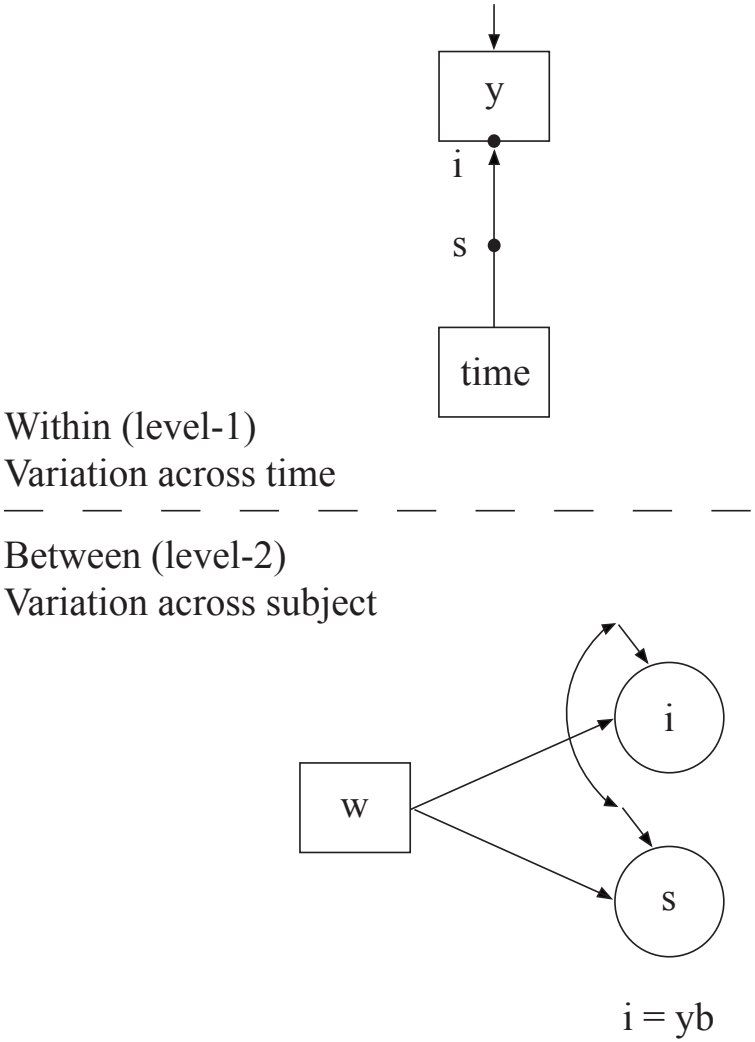
1. Correlation between time points not fully explained by growth factors alone due to closely spaced measurements - autocorrelation needs to be added
2. Time series are too long causing slow computations

Solving Problem 1. Add Residual (Auto) Correlation: Growth Modeling In Single-Level, Wide Format Version y as 5 columns in the data



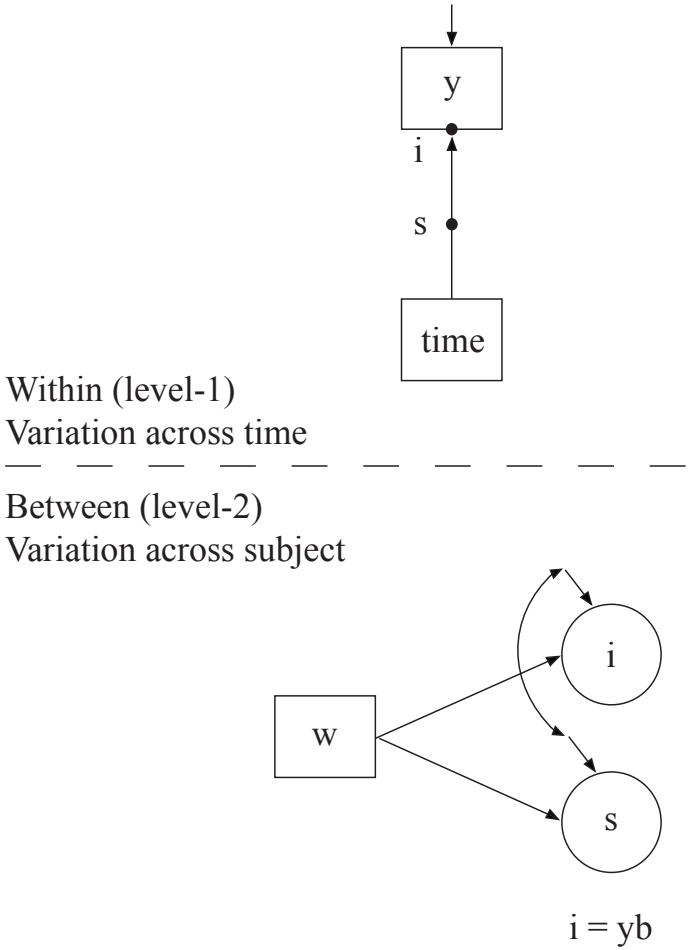
Mplus User's Guide ex6.17 - but cumbersome with large T.

Solving Problem 2. Switch From Single-Level to Two-Level, Long Format Version: y as 1 column in the data



Mplus User's Guide ex9.16

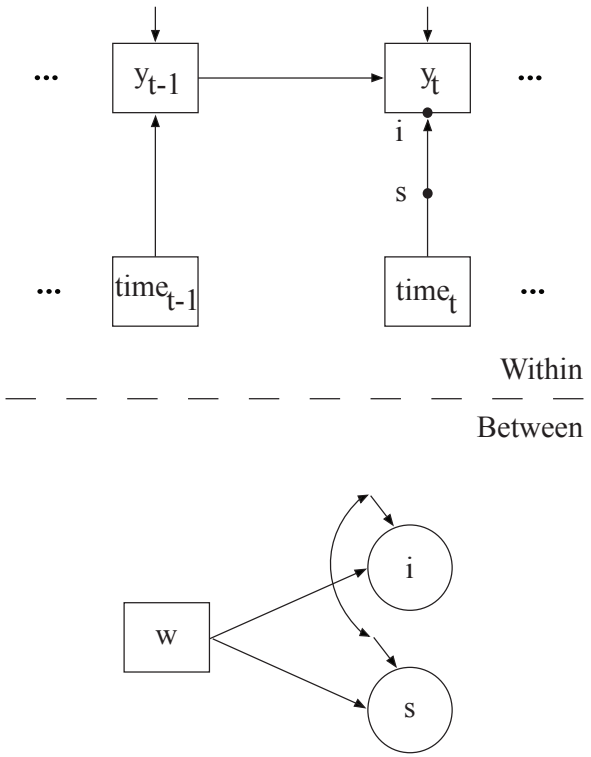
Growth Modeling In Two-Level, Long Format




```
VARIABLE:  CLUSTER = subject;
           WITHIN = time;
           BETWEEN = w;
ANALYSIS:  TYPE = TWOLEVEL RANDOM;
MODEL:     %WITHIN%
           s | y ON time;
           %BETWEEN%
           y s ON w; ! y is the same as i
           y WITH s;
```

But where is the autocorrelation? And how can it be made random?

Solution: Two-Level Time Series Analysis With A Trend Allowing Autocorrelation and Many Time points



Autoregression for the residuals instead?

Hamaker (2005). Conditions for the equivalence of the autoregressive latent trajectory model and a latent growth curve model with autoregressive disturbances. *Sociological Methods & Research*.

Example: Smoking Cessation (EMA)
Overview of Analyses

- N = 1 time series analysis
- Two-level time series analysis
- Cross-classified time series analysis - looking for trends over time and finding trend functions
- Adding trend to two-level time series analysis
- Cross-classified time series analysis with a trend
- Time-varying effect modeling (TVEM) using cross-classified time series analysis

Outline

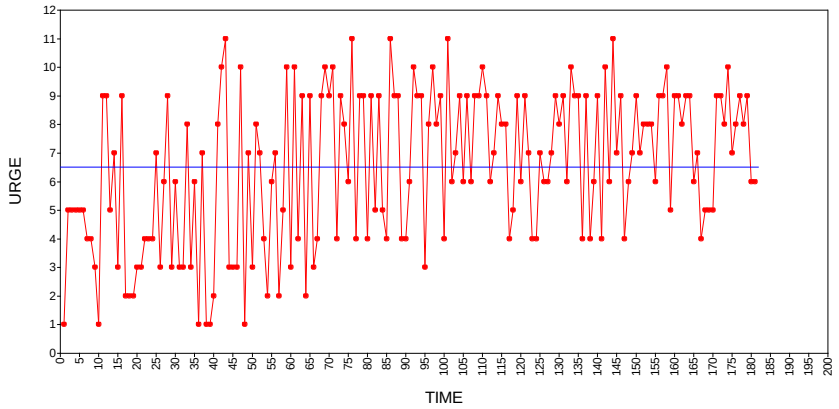
- Introduction to Bayesian analysis
- Introduction to longitudinal analysis
- **N=1 time series analysis**
- Two-level time series analysis
- Cross-classified time series analysis
- Latent variable time series analysis

EMA Example: Smoking Urge Data

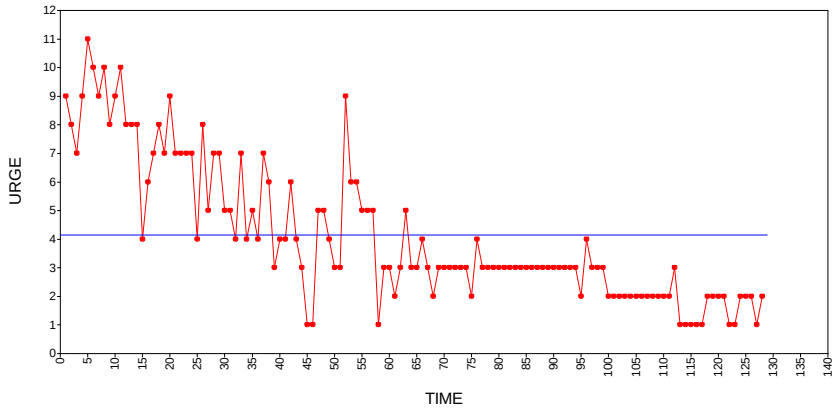
- Shiffman smoking cessation data
- N = 230, T $\approx$ 150: Random prompts from Personal Digital Assistant (hand held PC) approx. 5 times per day for a month
- Variables: Smoking urge (0-10 scale), negative affect (unhappy, irritable, miserable, tense, discontent, frustrated-angry, sad), gender, age, quit/relapse
- Shiyko et al. (2012). Using the time-varying effect model (TVEM) to examine dynamic associations between negative affect and self confidence on smoking urges. *Prevention Science*, 13, 288-299

N = 1 Time Series Analysis Of Subjects 227 And 5

Smoking urge plotted against time for subject 227 (didn't quit)



Smoking urge plotted against time for subject 5 (did quit)



N = 1 Time Series Analysis Of Subjects 227 And 5

	Estimate	Posterior S.D.	One-Tailed P-Value	95% C.I. Lower 2.5%	95% C.I. Upper 2.5%	Significance
urge ON						
urge&1	0.112	0.068	0.060	-0.027	0.240	
negaff	1.196	0.178	0.000	0.810	1.542	*
Intercepts						
urge	4.882	0.494	0.000	3.899	5.865	*
Residual Variances						
urge	5.719	0.635	0.000	4.646	7.070	*

	Estimate	Posterior S.D.	One-Tailed P-Value	95% C.I. Lower 2.5%	95% C.I. Upper 2.5%	Significance
urge ON						
urge&1	0.822	0.050	0.000	0.723	0.918	*
negaff	-0.257	0.408	0.272	-1.087	0.516	
Intercepts						
urge	0.517	0.377	0.074	-0.247	1.230	
Residual Variances						
urge	2.007	0.272	0.000	1.566	2.617	*

Time Interval (TINTERVAL option)

- Used to create a new time variable and insert missing data records when data are misaligned with respect to time:
  - due to missed measurement occasions that are not assigned a missing value flag
  - due to random measurement occasions
- The creation of the new time variable involves both substantive and statistical considerations

For more details, technical discussion and simulations, see Asparouhov, Hamaker, Muthén (2017) at www.statmodel.com.

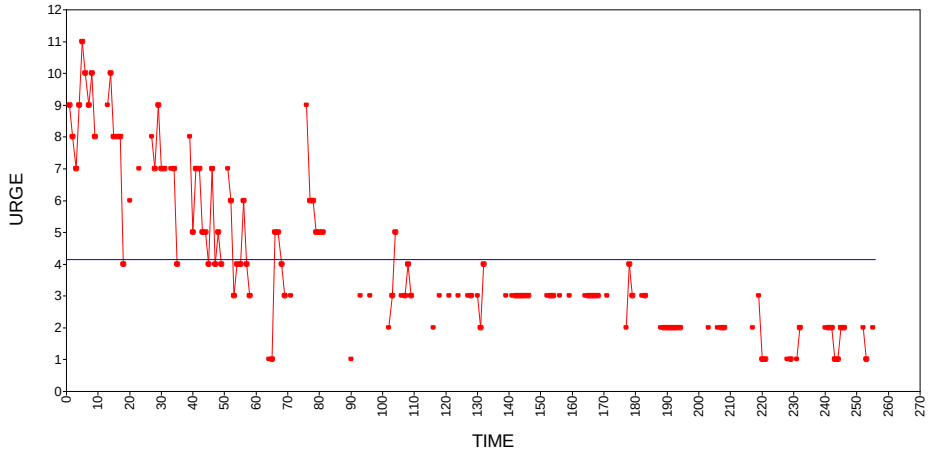
A Tinterval Example for One Subject

- Observed time given in fractions of a day - the 7 time points of the first day for one subject are shown in the table below
- An interval of 0.08 is used corresponding approximately to 2 hours (2/24 = 0.0833), that is, bin size = 0.08
- The lowest observed time is 0.32 (0.32*24 is 7:41 am); this is the mid point of the first bin, with the new time value 1 used in the analysis

Observed time	Bins	New time	Outcome
0.32	0.28 - 0.36	1	observed
0.39	0.36 - 0.44	2	observed
0.51	0.44 - 0.52	3	observed
0.59	0.52 - 0.60	4	observed
0.62	0.60 - 0.68	5	observed
	0.68 - 0.76	6	**missing**
0.77	0.76 - 0.84	7	observed
	0.84 - 0.92	8	**missing**
0.93	0.92 - 1.00	9	observed

N = 1 Time Series Analysis Using Tinterval= timeqd(0.08)

Subject 5 (did quit): Tinterval results in missing data records inserted to resolve different time distances between measurements



	Estimate	Posterior S.D.	One-Tailed P-Value	95% C.I. Lower 2.5%	95% C.I. Upper 2.5%	Significance
Subject 5 without Tinterval						
urge ON						
urge&1	0.822	0.050	0.000	0.723	0.918	*
negaff	-0.257	0.408	0.272	-1.087	0.516	
Subject 5 with Tinterval (0.08)						
urge ON						
urge&1	0.844	0.037	0.000	0.772	0.917	*
negaff	-0.158	0.382	0.328	-0.930	0.577	

Mplus Input For Subject 5 Time Series Regression

```
TITLE:      Shiffman smoking urge data N = 1 model for subject 5 (quit=1)
DATA:       FILE = combined_relapsers_quitters_03-17-17.csv;
VARIABLE:   NAMES = subject t day urge craving negaff arousal timeqd
            gender age quit;
            ! quit = 1 for quitters, 0 for relapsers
            USEVARIABLES = urge negaff;
            LAGGED = urge(1);
            MISSING = ALL(999);
            TINTERVAL = timeqd(0.08);
            USEOBSERVATIONS = subject EQ 5;
            IDVARIABLE = _recnum;
ANALYSIS:   ESTIMATOR = BAYES;
            PROCESSORS = 2;
            BITERATIONS = (1000);
MODEL:      urge ON urge&1 negaff;
            negaff;
OUTPUT:     TECH1 TECH8 STANDARDIZED TECH4 RESIDUAL;
PLOT:       TYPE = PLOT3;
```

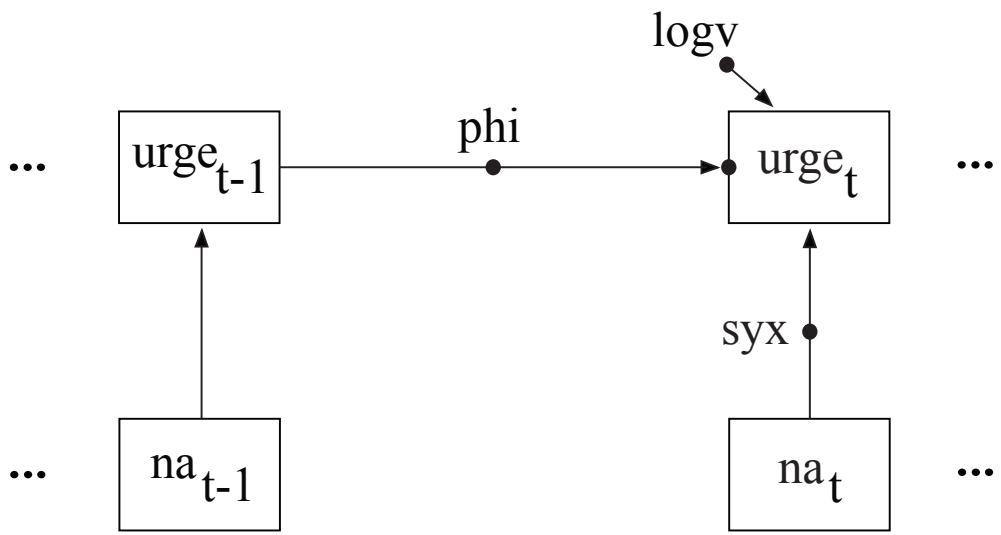
Outline

- Introduction to Bayesian analysis
- Introduction to longitudinal analysis
- N=1 time series analysis
- **Two-level time series analysis**
- Cross-classified time series analysis
- Latent variable time series analysis

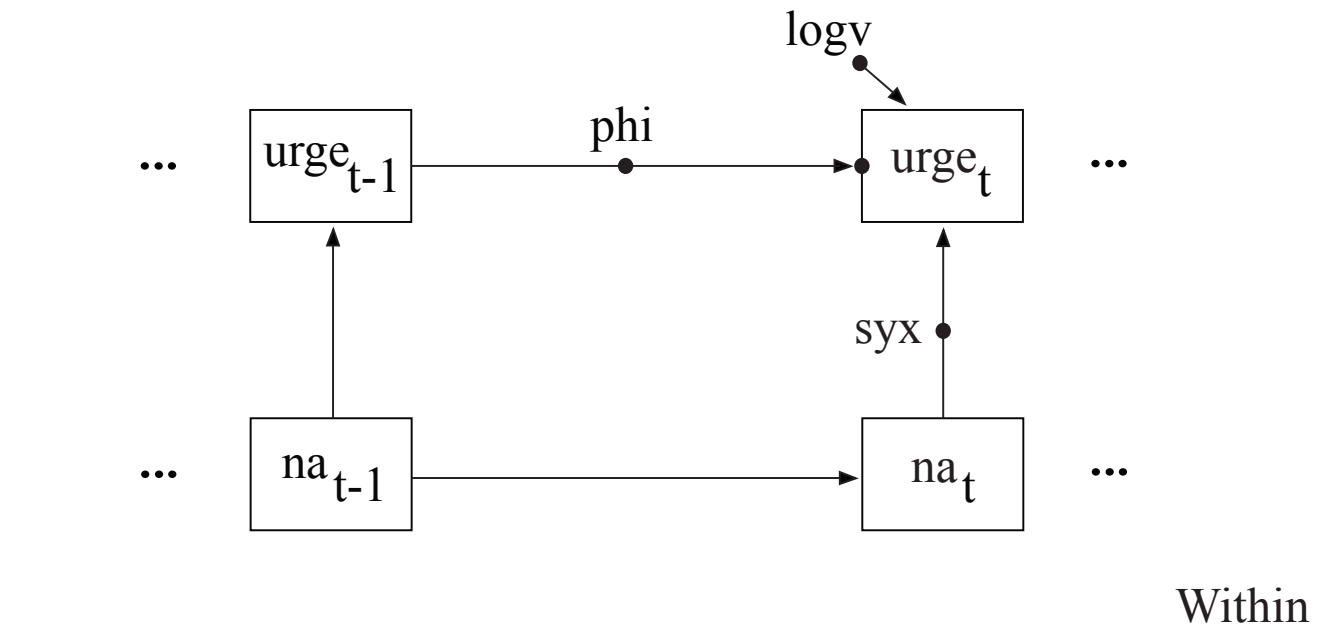
Two-Level Time Series Analysis ($N > 1$)

- Analysis of all N = 230 smoking data subjects
  - Allowing for parameter variation across subjects using random effects

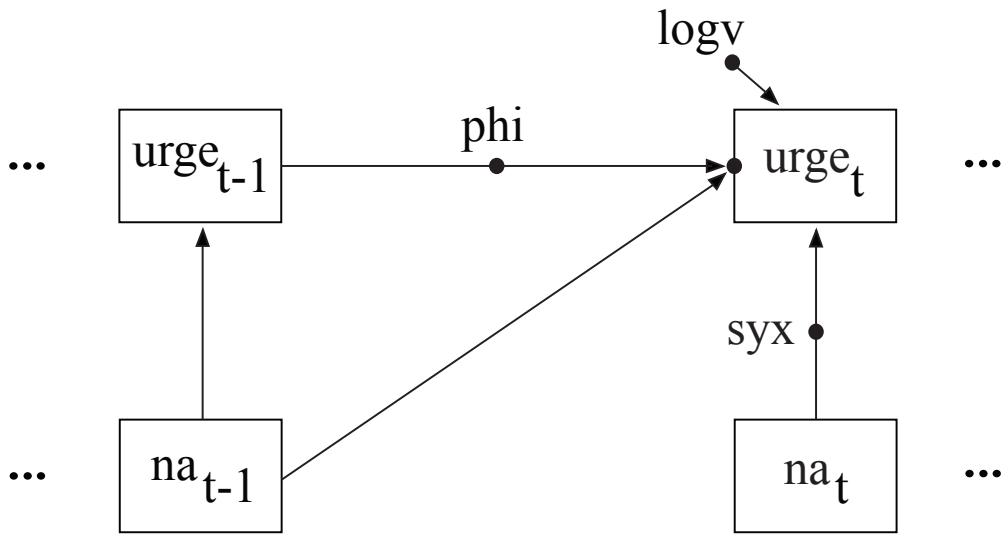
Two-Level Time Series Analysis: Regression of Smoking Urge on Negative Affect (na) Using 4 Random Effects



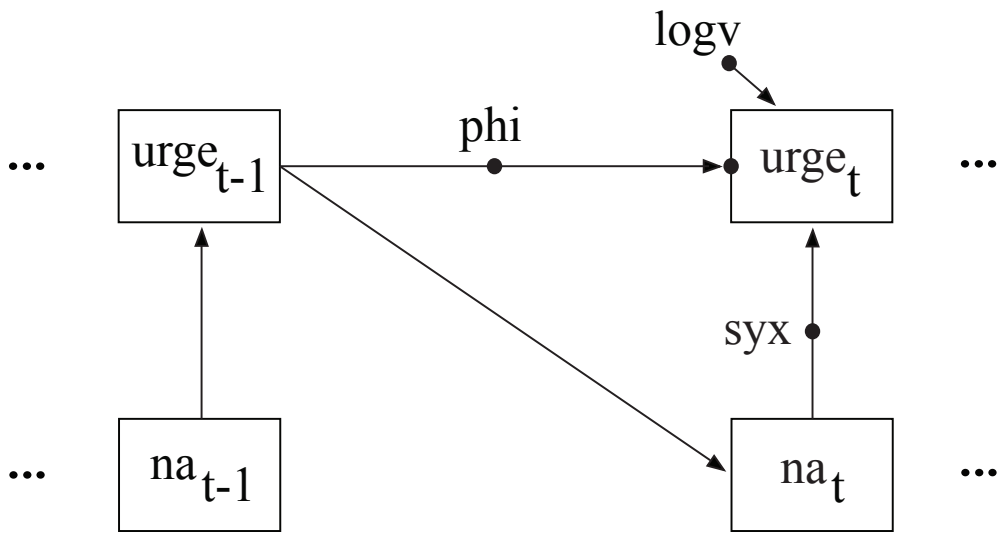
Two-Level Time Series Analysis: Regression of Smoking Urge on Negative Affect (na) Using 4 Random Effects



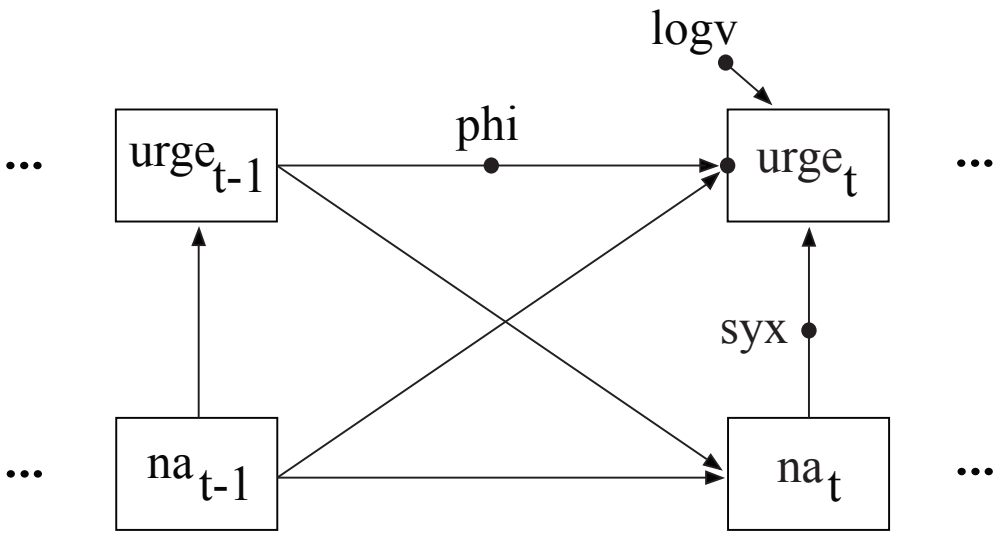
Two-Level Time Series Analysis: Regression of Smoking Urge on Negative Affect (na) Using 4 Random Effects



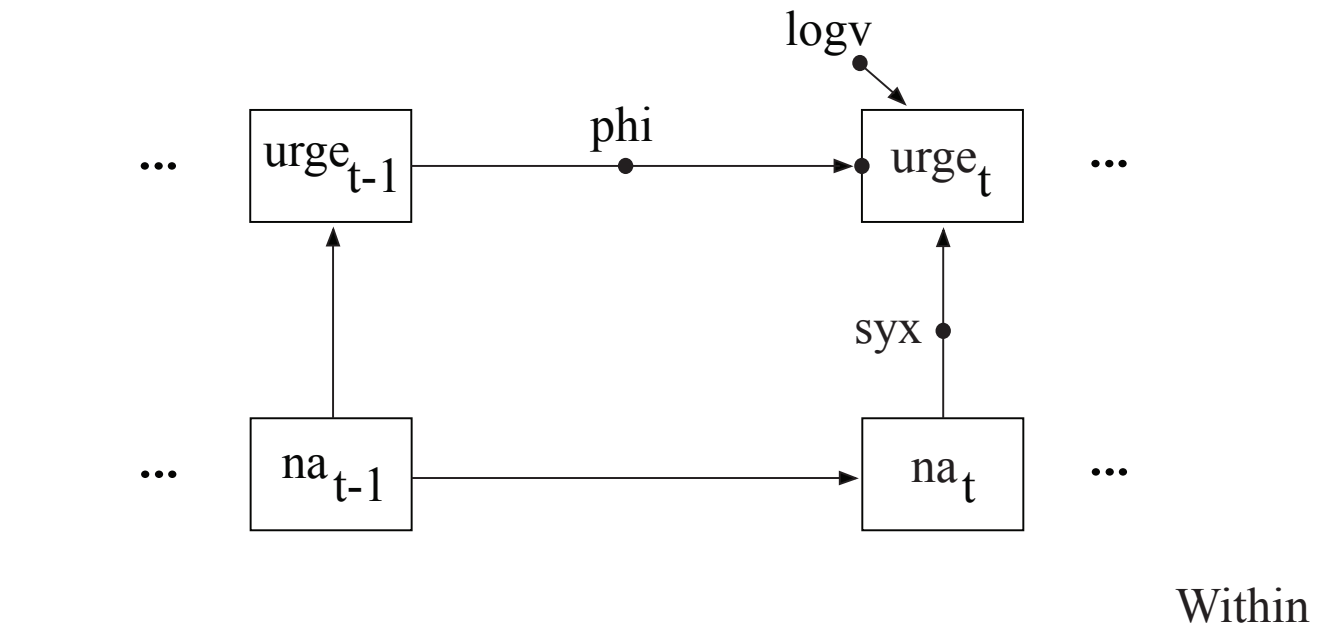
Two-Level Time Series Analysis: Regression of Smoking Urge on Negative Affect (na) Using 4 Random Effects



Two-Level Time Series Analysis: Regression of Smoking Urge on Negative Affect (na) Using 4 Random Effects



Two-Level Time Series Analysis: Regression of Smoking Urge on Negative Affect (na) Using 4 Random Effects



Mplus Input for Two-Level Regression Analysis

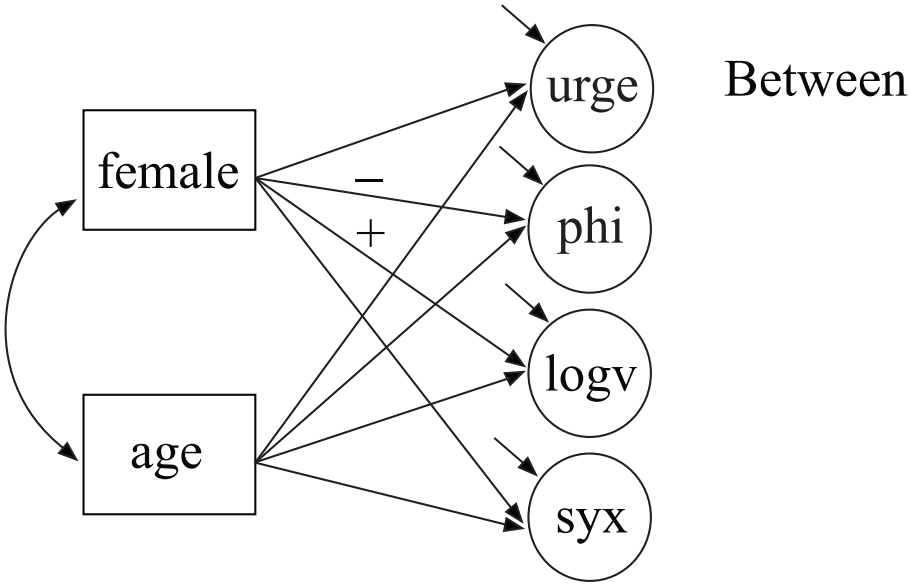
```
VARIABLE:   NAMES = subject t day urge craving negaff arousal timeqd
            gender age quit;
            !quit = 1 for quitters, 0 for relapsers
            USEVARIABLES = urge negaff age female;
            CLUSTER = subject;
            BETWEEN = female age;
            WITHIN = negaff;
            LAGGED = urge(1) negaff(1);
            MISSING = ALL(999);
            TINTERVAL = timeqd(0.08);
DEFINE:     female = gender - 1;
            age = (age-44.3)/10.1;
            CENTER negaff(GROUPMEAN);
ANALYSIS:   TYPE = TWOLEVEL RANDOM;
            ESTIMATOR = BAYES;
            PROCESSORS = 2;
            BITERATIONS = (1000);
```

Mplus Input for Two-Level Regression Analysis, Cont'd

```
MODEL:      %WITHIN%
            phi | urge ON urge&1;
            logv | urge;
            syx | urge ON negaff;
            negaff ON negaff&1;
            %BETWEEN%
            urge phi logv syx ON female age;
            urge phi logv syx WITH urge phi logv syx;
OUTPUT:     TECH1 TECH8 FSCOMPARISON STANDARDIZED
            TECH4 RESIDUAL;
PLOT:       TYPE = PLOT3;
            FACTORS = ALL;
```

Run time: 3:54 (3:13 without FACTORS=ALL)

Between-Level Results



- phi ON female not significant unless both logv and syx are allowed to be random

New Output Warnings

*** WARNING
One or more individual-level variables have no variation within a cluster for the following clusters.
Variable Cluster IDs with no within-cluster variation

URGE 160 12 60 192 186 49

WARNING: PROBLEMS OCCURRED IN SEVERAL ITERATIONS IN THE COMPUTATION
OF THE STANDARDIZED ESTIMATES FOR SEVERAL CLUSTERS. THIS IS MOST LIKELY DUE
TO AR COEFFICIENTS GREATER THAN 1 OR PARAMETERS GIVING NON-STATIONARY MODELS.
SUCH POSTERIOR DRAWS ARE REMOVED. THE FOLLOWING CLUSTERS HAD SUCH PROBLEMS:

160 115 205

BETWEEN-LEVEL FACTOR SCORE COMPARISONS

Results for Factor PHI

Ranking	Cluster	Factor Score	Ranking	Cluster	Factor Score	Ranking	Cluster	Factor Score
1	115	0.941	2	11	0.936	3	205	0.912
4	113	0.907	5	138	0.906	6	4	0.901

Checking the Time Interval Value

- The choice of 0.08 gives a warning:

  THE VALUE SPECIFIED IN THE TINTERVAL OPTION MAY BE TOO BIG. THE MAXIMUM DISCREPANCY BETWEEN THE ACTUAL TIME AND THE TIME RECODED BY THE TINTERVAL OPTION IS 9.517 IN CLUSTER 33.

- 0.08 corresponds to 2 hours: 2/24 = 0.08
- Because in this run 0.08 is represented as one time unit in the analysis, 9.517 corresponds to 9.517*2 = 19 hours displacement
- Given the large displacement, the data for cluster (subject) 33 should be inspected: first 30 observations made in less than 2-day span!? (aim: 5/day)
- 0.08 should perhaps be changed to say 0.04

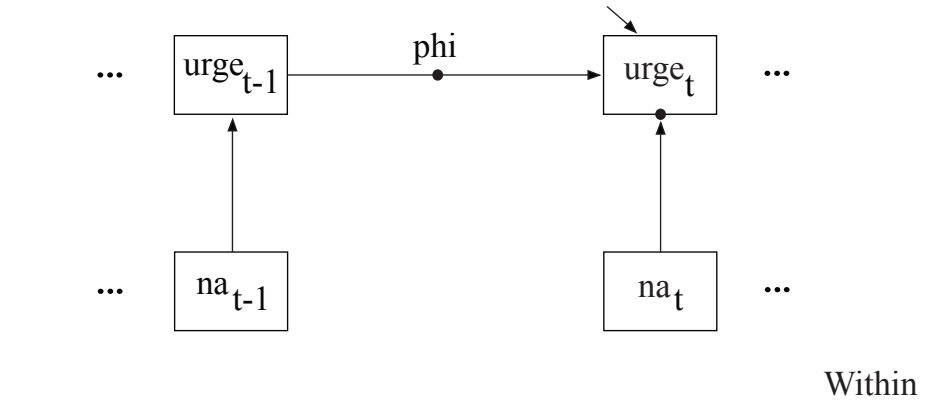
Checking Sensitivity to Tinterval Choices

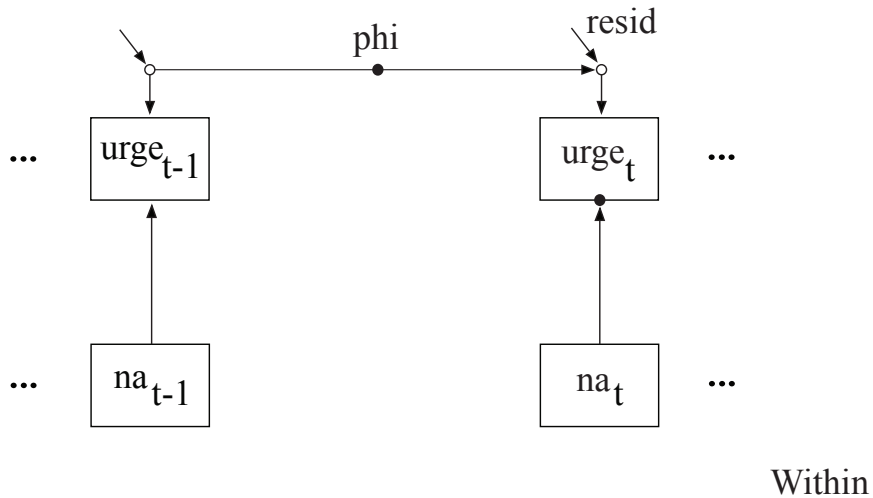
Univariate analysis of urge ON urge&1 with non-random ϕ and only the mean random

Tinterval	ϕ	Coverage	Time (secs)
0.08	0.325	0.41	22
0.06	0.344	0.31	25
0.04	0.373	0.20	40

- If an AR(1) model holds, ϕ with interval t (say 0.04), results in ϕ^2 for interval $2t$ (0.08)
- Because 0.373^2 is not equal to 0.325, a pure AR(1) model does not hold (for instance, urge ON urge&2 is significant)
- Smaller time interval gives more missing data, i.e. lower coverage
  - 10 - 15 % coverage is good
  - Coverage as low as 5% is ok
- Smaller time interval gives longer run time

Technical Interlude: Ampersand Versus Hat Where Should The Autocorrelation Be Applied?





Comparing Regular vs Residual Auto-Correlation

Regular AR(1):

```
urge ON negaff;
phi | urge ON urge&1;
```

Residual AR(1) in V8.1:

```
urge ON negaff;
phi | urge^ ON urge^1;
```

Model	DIC	pD
Regular AR (for the whole outcome)	45,3727	69623
Residual AR	45,6347	70440

- DIC: Deviance information criterion
  - Bayesian counterpart to BIC (lower is better)
- pD: Effective number of parameters
  - Also includes latent variables and missing data
- DIC requires a large number of iterations (20K used here); only for continuous outcomes

Outline

- Introduction to Bayesian analysis
- Introduction to longitudinal analysis
- N=1 time series analysis
- Two-level time series analysis
- **Cross-classified time series analysis**
- Latent variable time series analysis

Cross-Classified Time Series Analysis ($N > 1$)

- Two between-level cluster variables: subject crossed with time (one observation for a given subject at a given time point).
- Generalization of the two-level model providing more flexibility: random effects can vary across not only subject but also time

Consider the two-level model with a random intercept/mean:

$$y_{it} = \underbrace{\alpha + \alpha_i}_{\text{Between subject}} + \underbrace{\beta\, y_{w,it-1} + \varepsilon_{it}}_{\text{Within subject}}. \quad (2)$$

The corresponding cross-classified model is:

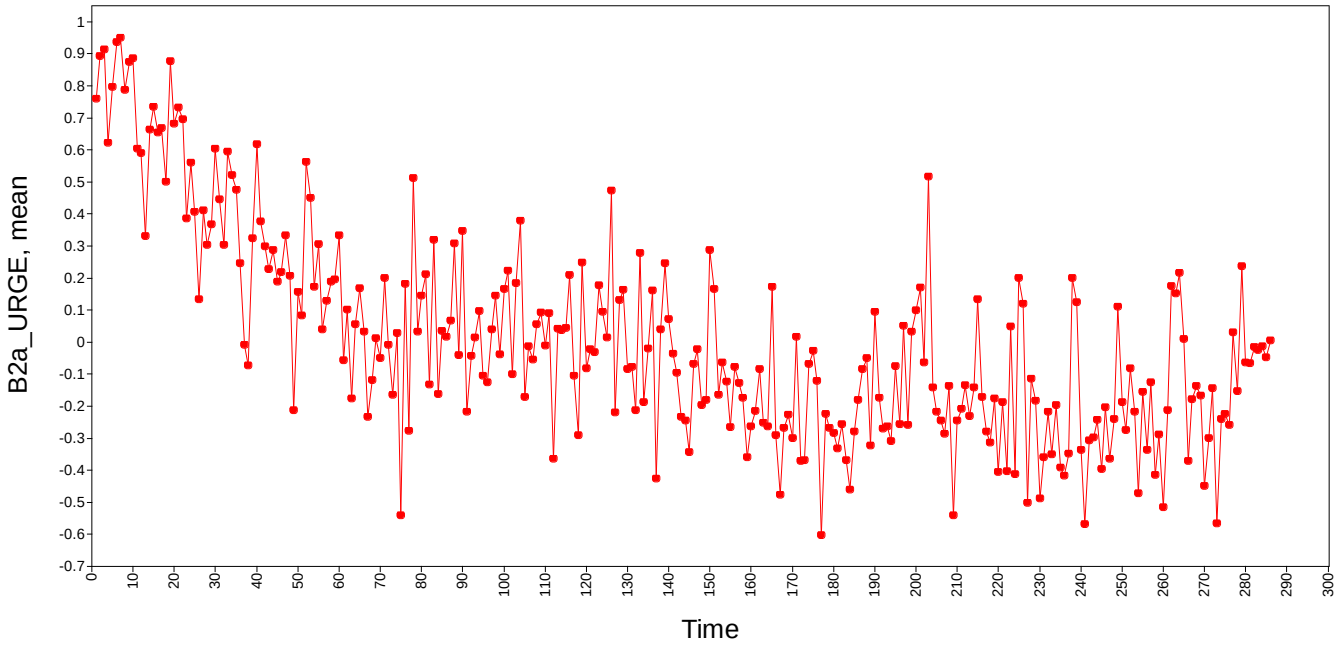
$$y_{it} = \underbrace{\alpha + \alpha_i}_{\text{Between subject}} + \underbrace{\alpha_t}_{\text{Between time}} + \underbrace{\beta\, y_{w,it-1} + \varepsilon_{it}}_{\text{Within subject}}. \quad (3)$$

- The Bayes MCMC algorithm is more complex and considerably slower

Cross-Classified Analysis: A Quick Way to Spot a Trend in a Variable

```
TITLE:      Shiffman smoking urge data, checking for trend in urge
DATA:       FILE = combined_relapsers_quitters_03-17-17.csv;
VARIABLE:   NAMES = subject t day urge craving negaff arousal timeqd gender age
            quit; !quit = 1 for quitters, 0 for relapsers
            USEVARIABLES = urge;
            CLUSTER = subject timeqd;
            LAGGED = urge(1);
            MISSING = ALL(999);
            TINTERVAL = timeqd(0.08);
ANALYSIS:   TYPE = CROSSCLASSIFIED RANDOM;
            ESTIMATOR = BAYES;
            PROCESSORS = 2;
            BITERATIONS = (1000);
MODEL:      %WITHIN%
            urge ON urge&1;
            %BETWEEN subject%
            urge;
            %BETWEEN timeqd%
            urge;
OUTPUT:     TECH1 TECH8;
PLOT:       TYPE = PLOT3;
            FACTORS = urge(50);
```

Cross-Classified Analysis of Trend: Time Series Plot of Urge Factor Scores



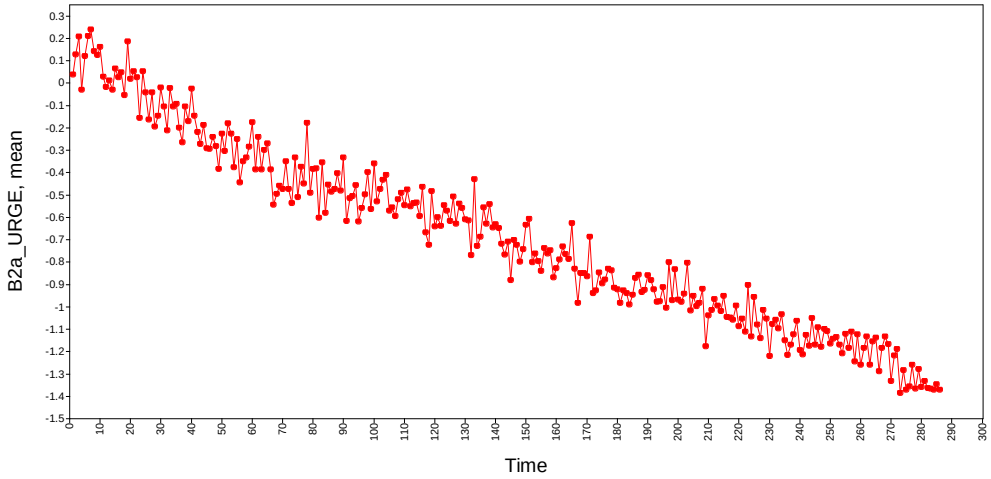
- Run time is only 1:10 with fixed AR(1)
- The trend can be modeled according to some functional form
  - In a cross-classified analysis
  - In a two-level analysis

Imposing Linear/Quadratic Trend in Cross-Classified

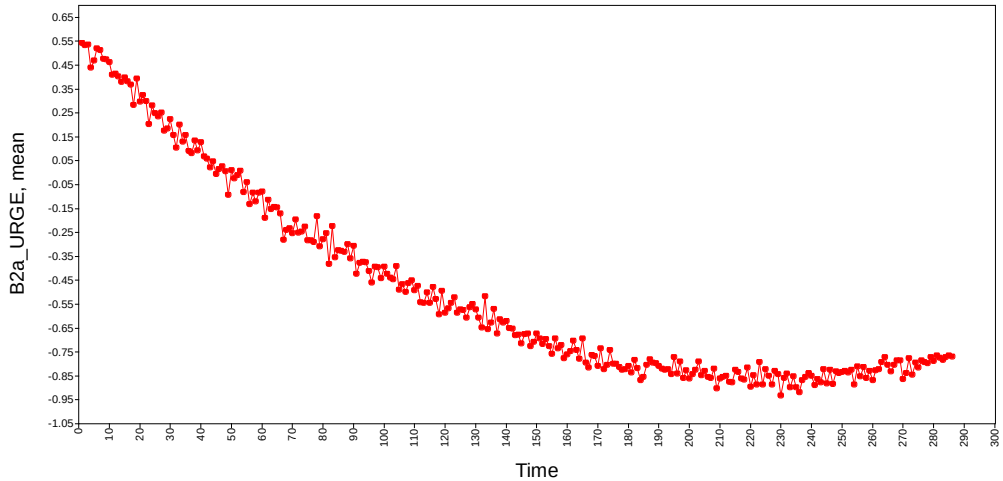
```
VARIABLE:   NAMES = subject t day urge craving negaff arousal timeqd gender age
            quit; !quit = 1 for quitters, 0 for relapsers
            USEVARIABLES = urge time time2;
            CLUSTER = subject timeqd;
            BETWEEN = (timeqd) time time2;
            LAGGED = urge(1);
            MISSING = ALL(999);
            TINTERVAL = timeqd(0.08);
DEFINE:     time = timeqd/100;
            time2 = time*time;
ANALYSIS:   TYPE = CROSSCLASSIFIED RANDOM;
            ESTIMATOR = BAYES;
            PROCESSORS = 2;
            BITERATIONS = (2000);
MODEL:      %WITHIN%
            urge ON urge&1;
            %BETWEEN subject%
            urge;
            %BETWEEN timeqd%
            urge;
            urge ON time time2;
OUTPUT:     TECH1 TECH8;
PLOT:       TYPE = PLOT3;
            FACTORS = urge(50);
```

Imposing Linear/Quadratic Trend in Cross-Classified

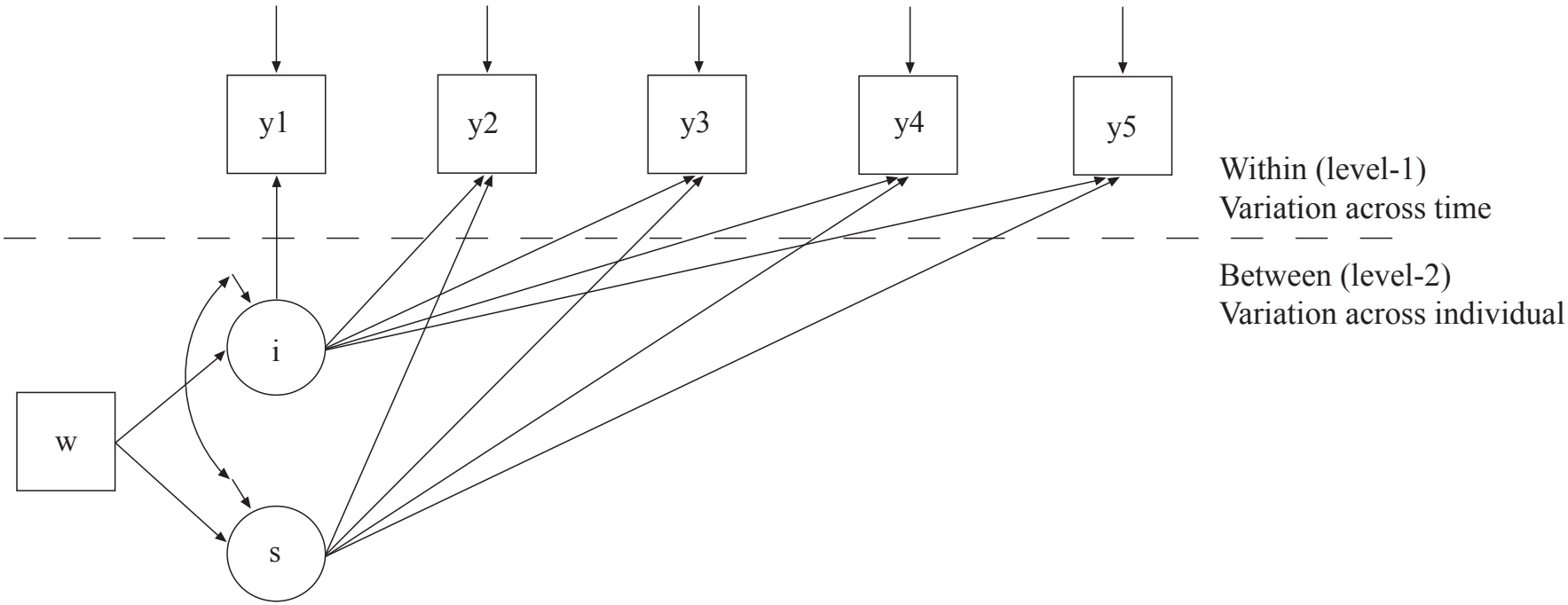
Linear:



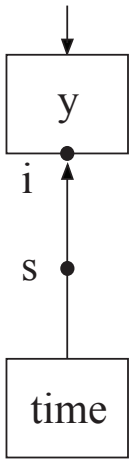
Quadratic:



Modeling the Trend: Recall How Growth Modeling Can Be Transformed From Wide To Long

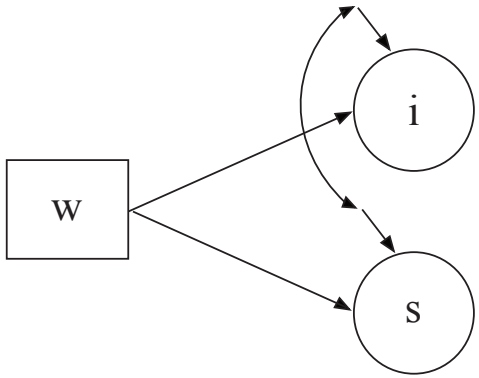


Growth Modeling: Two-Level, Long Format Version

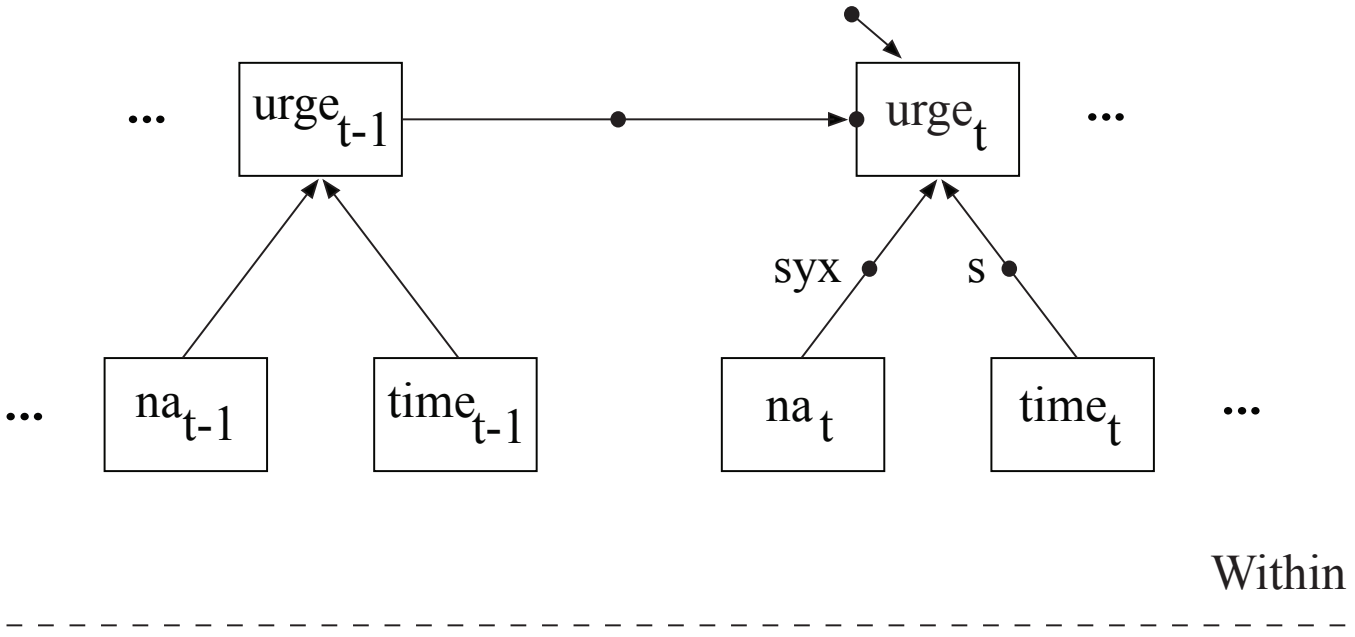


Within (level-1)
Variation across time

Between (level-2)
Variation across individual



Two-Level Time Series Analysis of Smoking Urge data
Adding a Trend for Urge.
- Growth Analysis with a Time-Varying Covariate



- Interpretation of s not the usual one; direct effect at each time
- An alternative formulation places the autoregression on the residuals (Hamaker, 2005; *SM&R*), resulting in the usual s interpretation

Mplus Input For Two-Level Trend Analysis

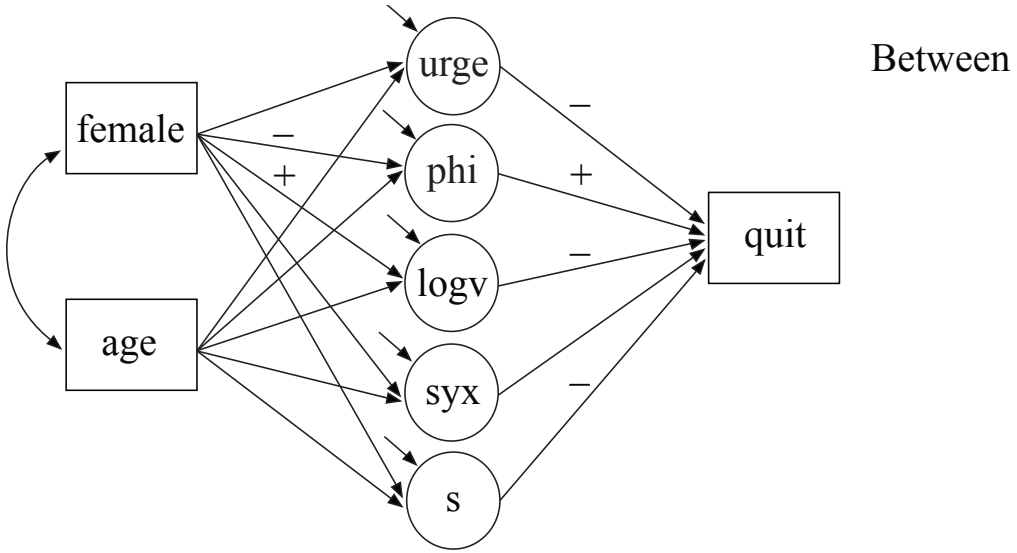
```
VARIABLE:   NAMES = subject t day urge craving negaff arousal timeqd
            gender age quit;
            USEVARIABLES = urge quit negaff age female time;
            CLUSTER = subject;
            BETWEEN = female age quit;
            CATEGORICAL = quit;
            WITHIN = time negaff;
            LAGGED = urge(1) negaff(1);
            MISSING = ALL(999);
            TINTERVAL = timeqd(0.08);
DEFINE:     female = gender - 1;
            age = (age-44.3)/10.1;
            time = timeqd/100-1.305;
            CENTER negaff(GROUPMEAN);
ANALYSIS:   TYPE = TWOLEVEL RANDOM;
            ESTIMATOR = BAYES;
            PROCESSORS = 2;
            BITERATIONS = (1000);
```

Mplus Input For Two-Level Trend Analysis, Cont'd

```
MODEL:      %WITHIN%
            phi | urge ON urge&1;
            syx | urge ON negaff;
            logv | urge;
            s | urge ON time;
            negaff ON negaff&1;
            time;
            %BETWEEN%
            urge syx s phi logv ON female age;
            urge syx s phi logv WITH urge syx s phi logv;
            quit ON urge syx s phi logv female age;
OUTPUT:     TECH1 TECH8 FSCOMPARISON STANDARDIZED TECH4
            RESIDUAL;
PLOT:       TYPE = PLOT3;
            FACTORS = ALL;
```

Run time: 4:42 (3:56 without FACTORS = ALL)

Results for Two-Level Regression Analysis of Smoking Urge Data: Adding a Trend for Urge. - Growth Analysis with a Time-Varying Covariate



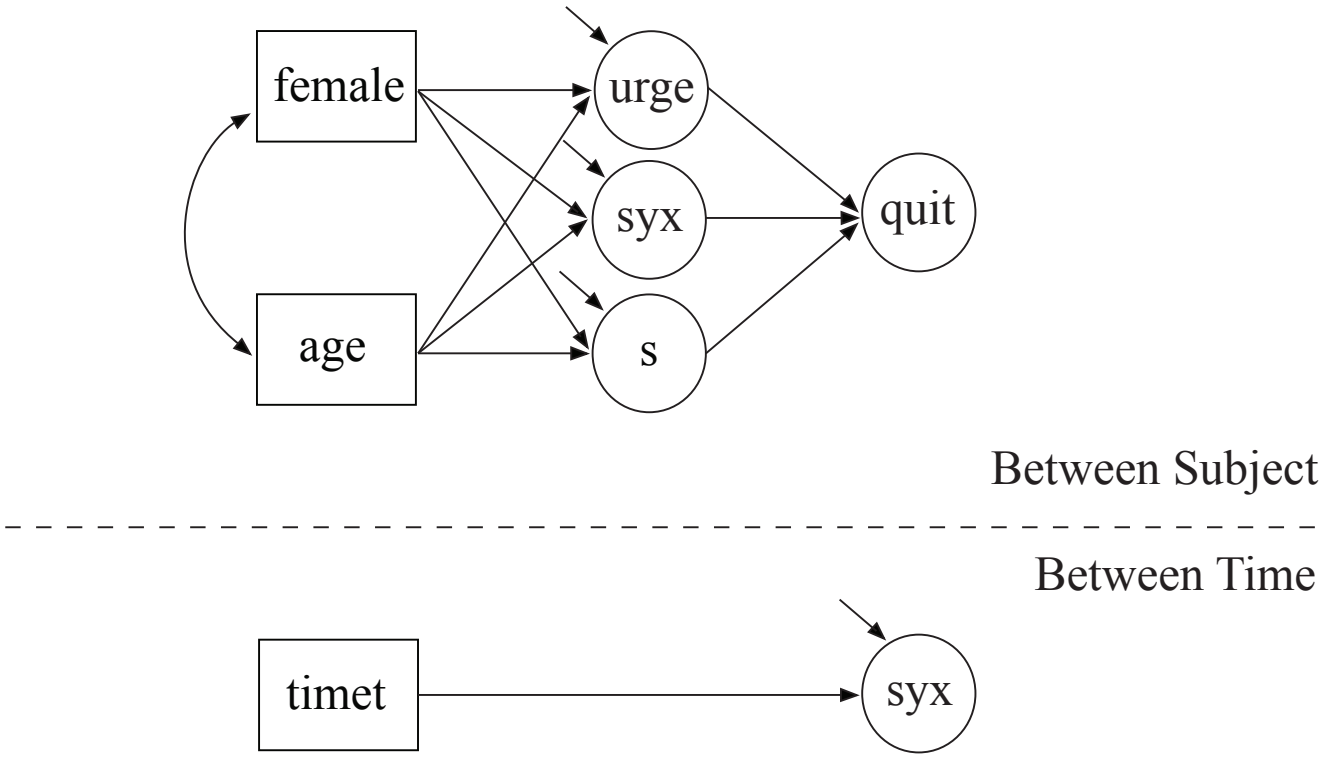
Quit (binary) regressed on random effects:

- higher urge gives lower quit probability
- higher autocorrelation gives higher quit probability
- higher residual variance gives lower quit probability
- higher trend slope gives lower quit probability

Time-Varying Effect Modeling (TVEM) Using Cross-Classified Analysis

- Cross-classified modeling allows parameters to change over time
- An example is a regression slope
  - Does the influence of negative affect on smoking urge decline over time?

Cross-Classified Regression Analysis of Smoking Urge Data: Adding a Trend for Urge **and the Negaff Regression Slope Cross-Classified** Growth Analysis With a Time-Varying Covariate



Mplus Input for Cross-Classified Regression Analysis with an Urge Trend and a Negaff Slope Trend

```
VARIABLE:   NAMES = subject t day urge craving negaff arousal timeqd
            gender age quit;
            ! quit = 1 for quitters, 0 for relapsers
            USEVARIABLES = urge quit negaff age female timew timet;
            CLUSTER = subject timeqd;
            BETWEEN = (subject) female age quit (timeqd) timet;
            CATEGORICAL = quit;
            WITHIN = negaff timew;
            LAGGED = urge(1);
            MISSING = ALL(999);
            TINTERVAL = timeqd(0.08);
DEFINE:     female = gender - 1;
            age = (age-44.3)/10.1;
            timew = timeqd/100-1.305;
            timet = timew;
            CENTER negaff(GROUPMEAN subject);
ANALYSIS:   TYPE = CROSS RANDOM;
            ESTIMATOR = BAYES;
            PROCESSORS = 2;
            BITERATIONS = (1000);
```

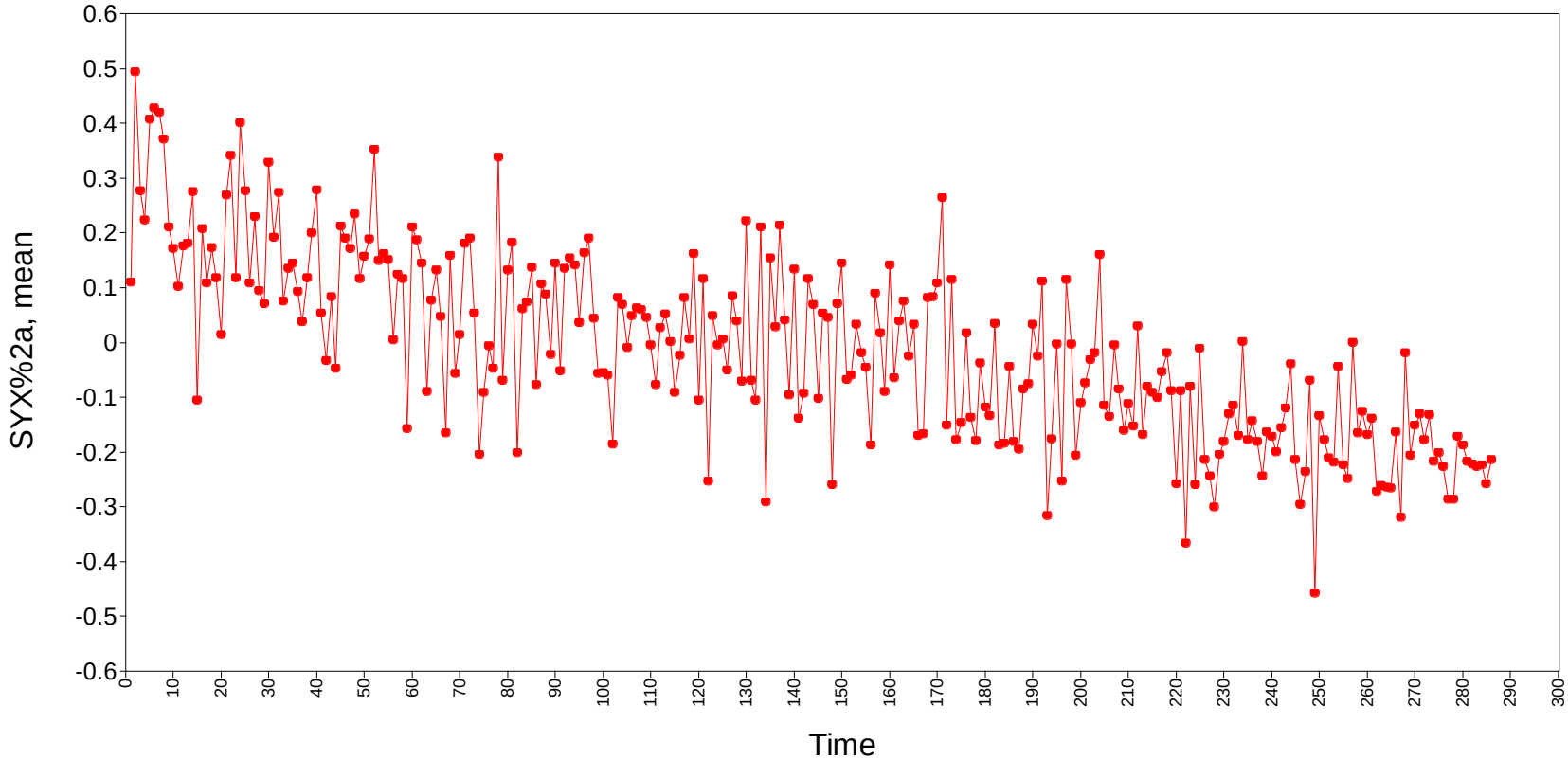
Input Continued

```
MODEL:      %WITHIN%
            urge ON urge&1;
            syx | urge ON negaff;
            s | urge ON timew;
            negaff;
            timew;
            %BETWEEN subject%
            urge syx s ON female age;
            urge syx s WITH urge syx s;
            quit ON urge syx s female age;
            s; [s];
            %BETWEEN timeqd%
            syx ON timet;
            urge WITH syx;
            s@0;
OUTPUT:     TECH1 TECH8;
PLOT:       TYPE = PLOT3;
            FACTORS = ALL;
```

- Run time: 34 minutes

Trend in Slope for Urge Regressed on Negative Affect

%BETWEEN timeqd%
syx ON timet; ! the estimate is significant negative



- The effect of negative affect on smoking urge is reduced over time
- syx%2a is the between-time factor score part of the syx slope (its mean is zero) - the full syx slope mean is expressed as on the next slide

Estimated syx Slope Mean At Different Time Points: A Time-Varying Effect (TVEM)

$$\hat{E}(syx|timet, female, age) = 0.474 \underbrace{-0.156 * timet}_{\text{from the between time level}} \underbrace{+0.072 * female + 0.000 * age}_{\text{from the between subject level}}$$

- DEFINE: timet = timeqd/100-1.305; with range -1.2 to +1.2
- timet = -1, males: $\hat{E}(syx|*) = \mathbf{0.63}$
- timet = 0, males: $\hat{E}(syx|*) = \mathbf{0.47}$
- timet = 1, males: $\hat{E}(syx|*) = \mathbf{0.32}$

Outline

- Introduction to Bayesian analysis
- Introduction to longitudinal analysis
- N=1 time series analysis
- Two-level time series analysis
- Cross-classified time series analysis
- **Latent variable time series analysis**

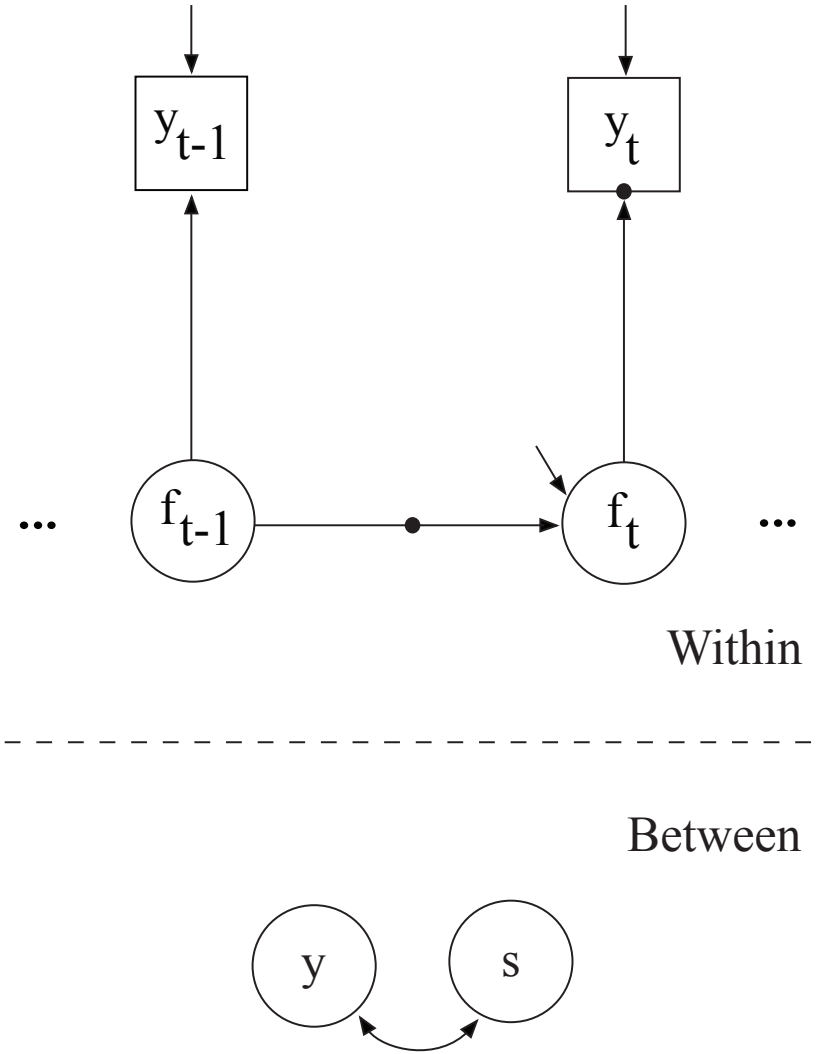
Time Series Analysis with Latent Variables: Latent Variables on the Within Level

- So far we have focused on latent variables on the between level in the form of random effects
  - Although on Within we have used the latent variable within-level decomposition of the outcome, centering by y_{bi}:

$$y_{wit} = y_{it} - y_{bi}$$

- Now we introduce within-level factors:
  - Factors defined by single indicators with measurement error
  - Residual factors in ARMA(1,1)
  - Factors defined by multiple indicators
  - Two-level and Cross-classified analysis
- Categorical latent variables (version 8.x, although an SEM article is already online; Asparouhov, Hamaker, Muthén, 2017):
  - Transition modeling (Hidden Markov, regime switching, time-series LTA) with latent class variables
  - Growth mixture modeling

Single-Indicator Measurement Error Model

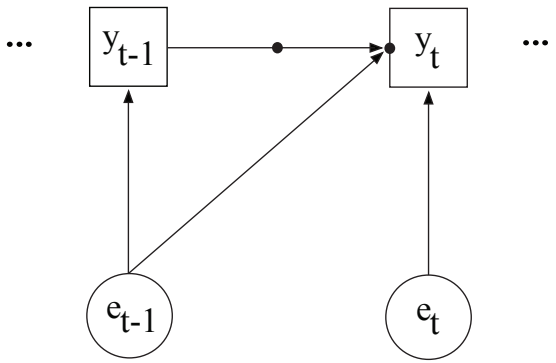


- Two types of errors:
  - Dynamic errors carry over from occasion to occasion (unobserved influences)
  - Measurement errors don't carry over (making errors answering; white noise)
- The model is identified unlike regular factor analysis due to auto-regressive feature (like panel data modeling a la Werts, Linn & Jöreskog, 1977)
- Schuurman et al. (2015) Frontiers of Psych; $N = 1$

Input for Measurement Error Model

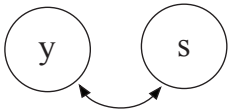
TITLE:	this is an example of a two-level time series analysis with a first-order autoregressive AR(1) factor analysis model for a single continuous indicator and measurement error
DATA:	FILE = ex9.33.dat;
VARIABLE:	NAMES = y subject;
	CLUSTER = subject;
ANALYSIS:	TYPE = TWOLEVEL RANDOM;
	ESTIMATOR = BAYES;
	PROCESSORS = 2;
	BITERATIONS = (5000);
MODEL:	%WITHIN%
	f BY y@1(&1);
	s | f ON f&1;
	%BETWEEN%
	y WITH s;
OUTPUT:	TECH1 TECH8;
PLOT:	TYPE = PLOT3;

A Related Model: ARMA(1,1)



Within

Between



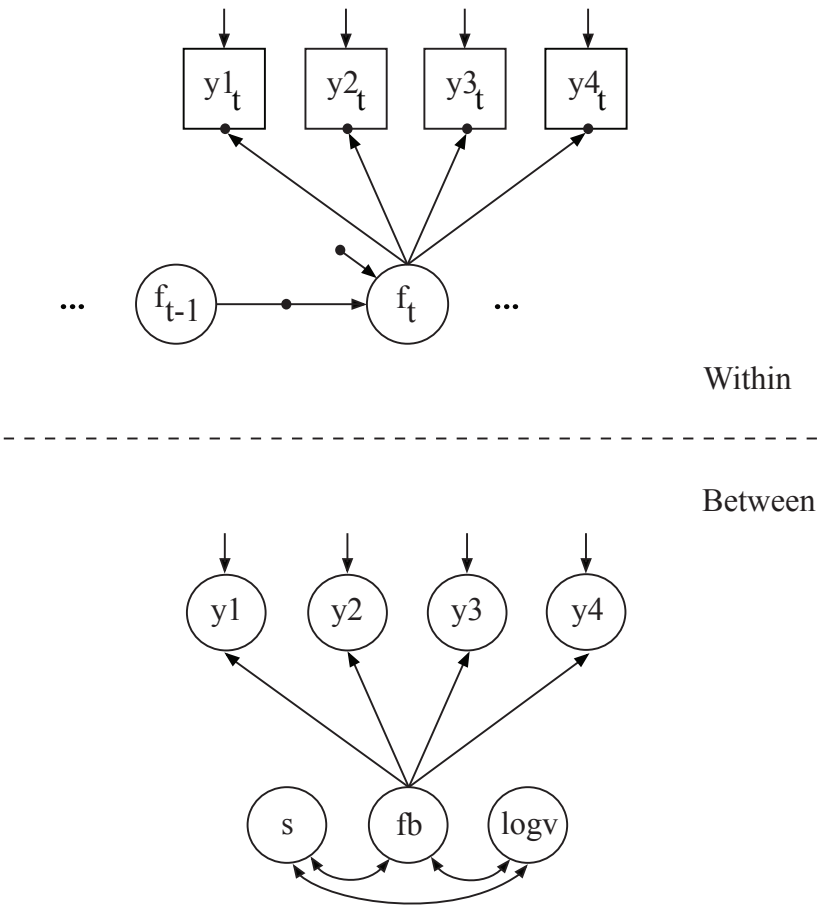

```
MODEL:      %WITHIN%
            s | y ON y&1;
            e BY y@1 (&1);
            y@.01;
            y ON e&1;
```

- AR stands for autoregressive and MA stands for moving average (Shumway & Stoffer, 2011)

Thoughts on Measurement Error versus ARMA(1,1)

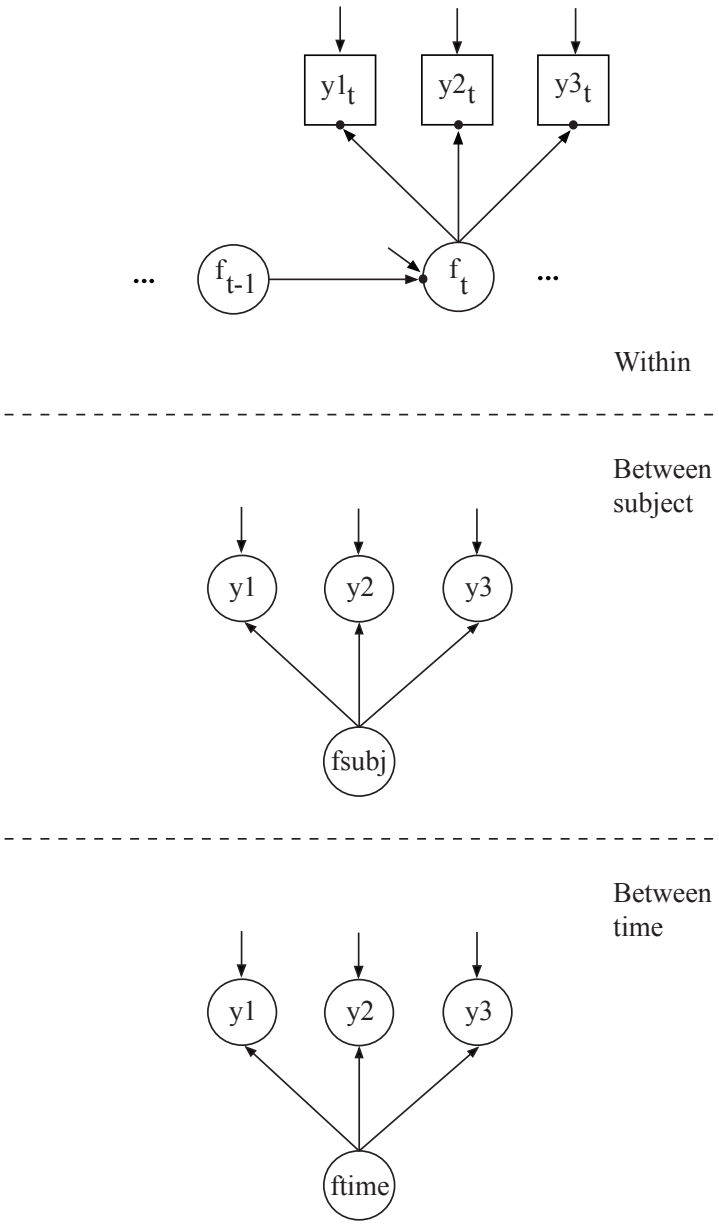
- Granger and Morris (1976) and Schuurman et al. (2015) show that for N = 1 time series analysis, ARMA (1, 1) is an alternative representation of the data used in the measurement error model; formulas show translation of parameters
  - In the Mplus implementation the measurement error formulation converges more smoothly than ARMA(1,1)
  - The $N = 1$ versions of these models require a large T, say $T > 100$
  - Preliminary simulations indicate that the $N > 1$ versions have good performance at $T = 50$, reasonable performance at $T = 25$, and maybe acceptable performance at $T = 14$: Suitable for daily diary designs
- AR models assume exponential decays in autocorrelation - the measurement error model allows a slower, more realistic decay (Asparouhov, 2017)
- A preliminary observation: it appears to be difficult to add random variance to the factor in the measurement error model
- Research questions: How does performance compare to having multiple indicators (e.g. 10 NA items)? Is random variance easier there?

Two-Level Factor Analysis: UG ex9.34



- Random intercepts become latent factor indicators on Between
- The figure shows a DAFS (direct autoregressive factor score) model on Within
- An alternative is the WNFS (white noise factor score) model which uses y1-y4 ON f&1 instead of f ON f&1
- A combination model is also identified (may need large T)
- N = 1 factor analysis: Engle & Watson (1981) in JASA, Molenaar (1985) in Psychometrika

Cross-Classified Factor Analysis: UG ex9.40



Measurement Non-Invariance Across Subjects Using Two-Level Factor Analysis: Random Intercepts

- For a certain item measured for individual i at time t, two-level factor analysis (see, e.g., Muthén, 1994) considers

$$y_{it} = \underbrace{\nu + \lambda_b f_{bi} + \varepsilon_{bi}}_{\text{Between}} + \underbrace{\lambda_w f_{wit} + \varepsilon_{wit}}_{\text{Within}}. \quad (4)$$

- This can be re-expressed as

$$Level\ 1: \ y_{it} = \nu_i + \lambda_w f_{wit} + \varepsilon_{wit}, \quad (5)$$

$$Level\ 2: \ \nu_i = \nu + \lambda_b f_{bi} + \varepsilon_{bi}, \quad (6)$$

which is a random intercept model, that is, there is measurement non-invariance across subjects wrt the intercepts (Jak et al., 2013, 2014; Muthén & Asparouhov, 2017). IRT typically uses $\lambda_w = \lambda_b$, $\varepsilon_{bi} = 0$,

$$Level\ 1: \ y_{it} = \nu + \lambda f_{it} + \varepsilon_{wit}, \quad (7)$$

$$Level\ 2: \ f_{it} = f_{bi} + f_{wit}, \quad (8)$$

- that is, no non-invariance and one single factor dimension

Measurement Non-Invariance of Intercepts and Loadings Across Subjects and Time

Mplus Version 8 offers:

- Two-level analysis:
  - Random intercepts varying across subjects
  - Random loadings varying across subjects: s1 - s10 | f BY y1 - y10
  - Asparouhov & Muthén (2015), Fox (2010)
- Cross-classified analysis:
  - Random intercepts varying across subjects and **time**
  - Random loadings varying across subjects
  - Version 7.4 had cross-classified analysis with random intercepts and loadings but not auto-correlation needed for ILD
  - For an example of random loadings varying across subjects, see the Mplus Version 8 User's Guide ex9.40 part2

Subject-Specific Reliability

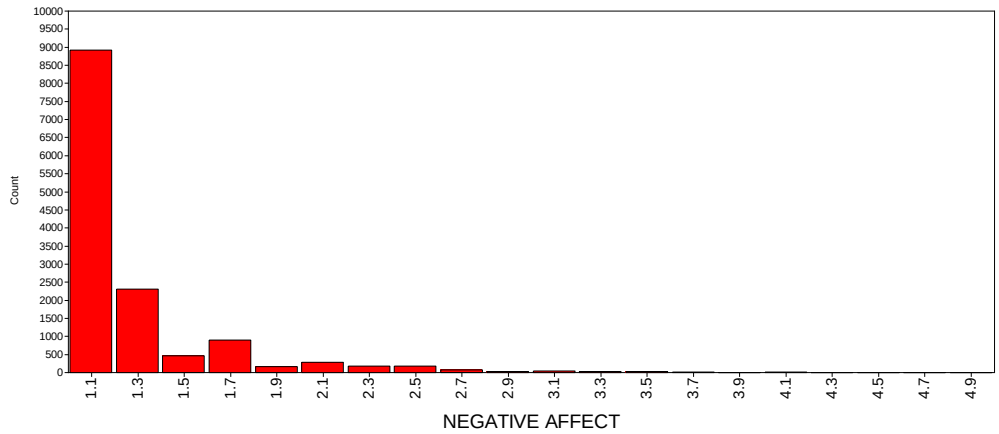
- The two-level and cross-classified factor analysis models imply
  - Measurement intercept and loadings possibly varying across subject and time
  - Factor variances and residual variances varying across subject and time
- This implies that reliabilities of test scores (based on a set of items) vary across subject and time
  - Hu, Nesselroade et al. (2016). Test reliability at the individual level. *Structural Equation Modeling*.
- But why not instead look at the precision with which the factor scores can be estimated?
  - Mplus Version 8 Monte Carlo simulations give correlations between true scores and estimated scores

Example: Item Factor Analysis (IRT) Using 10 Negative Affect Items

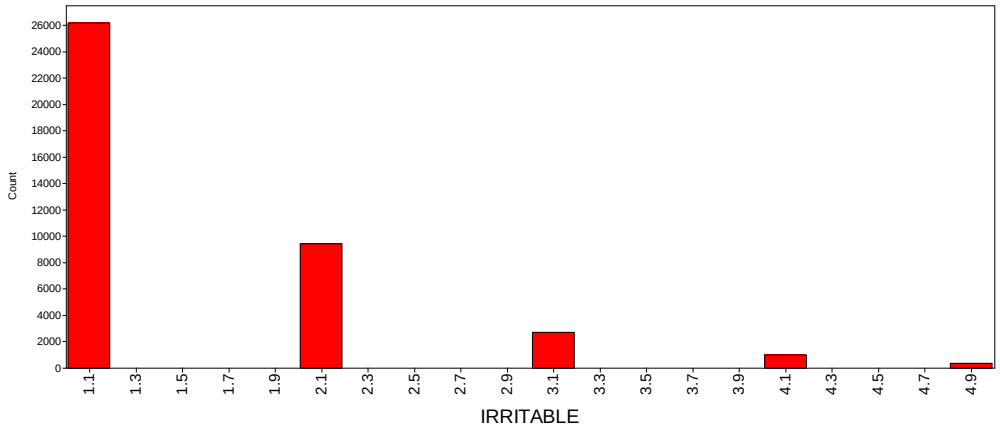
- Data from the older cohort of the Notre Dame Study of Health & Well-being (Bergeman): N = 270, T = 56 (daily measures on consecutive days)
- Wang, Hamaker, Bergeman (2012). Investigating inter-individual differences in short-term intra-individual variability. *Psychological Methods*
- Predictors and distal outcomes of negative affect development over the 56 days
- 10 NA items (5-cat scale): afraid, ashamed, guilty, hostile, scared, upset, irritable, jittery, nervous, distressed (average score used in article). Wide format would have 56*10 variables
- Question format: Today I felt... (1 = Not at all, ..., 5 = Extremely)
- 1-factor DAFS model

Negative Affect Distributions of NA in Bergeman Data

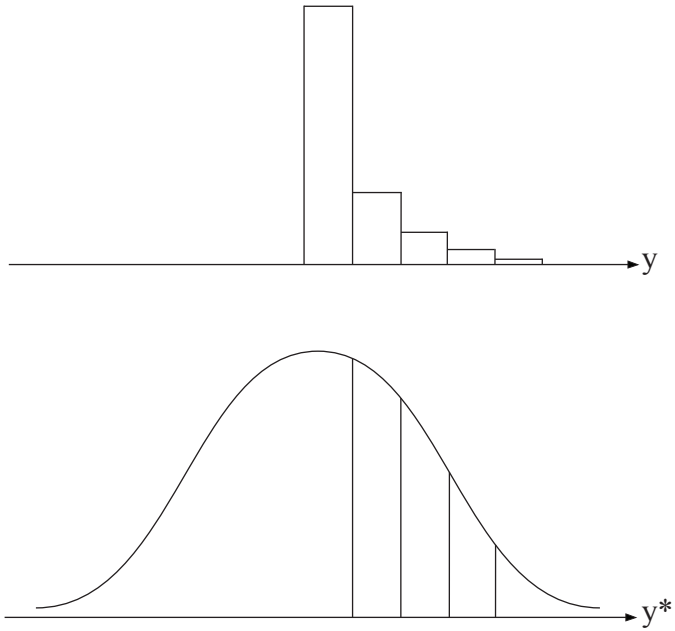
Average score (55% at floor value of 1 - Not at all for all 10 items):



Typical item distribution (66% at lowest value - Not at all):



Ordered Categorical Item Modeling: Proportional Odds Model (Graded Response Model)



- Despite non-normal y, we can have normality of:
  - The latent response variable y*
  - Any factors in the model
  - The between-level random effects

Mplus Input for **Cross-Classified** Factor Analysis with One Factor for 10 Ordinal NA Items

TITLE:	Bergeman twolevel
DATA:	FILE = bergeman.csv;
VARIABLE:	NAMES = subject gender age hosp1 chrhlth1 Somhlth1 slfhlth1 psqi neo day afraid1 unhappy1 annoyd1 ashmd1 guilty1 angry1 sad1 hostile1 scared1 upset1 irrtbl1 deprsd1 jttry1 drowsy1 slugish1 worrid1 nervs1 lonely1 fatiged1 distrsd1 nPANAS1;
	USEVARIABLES = afraid1 scared1 nervs1 jttry1 guilty1 ashmd1 irrtbl1 hostile1 upset1 distrsd1;
	CATEGORICAL = afraid1-distrsd1;
	CLUSTER = subject day;
	MISSING = all(999);
	TINTERVAL = day(1);
ANALYSIS:	**TYPE = CROSSCLASSIFIED RANDOM;**
	ESTIMATOR = BAYES;
	PROCESSORS = 2;
	BITERATIONS = (5000);
	THIN = 10;

Mplus Input for Cross-Classified Factor Analysis Continued

- IRT-style loading equality, setting the factor metric on the subject level

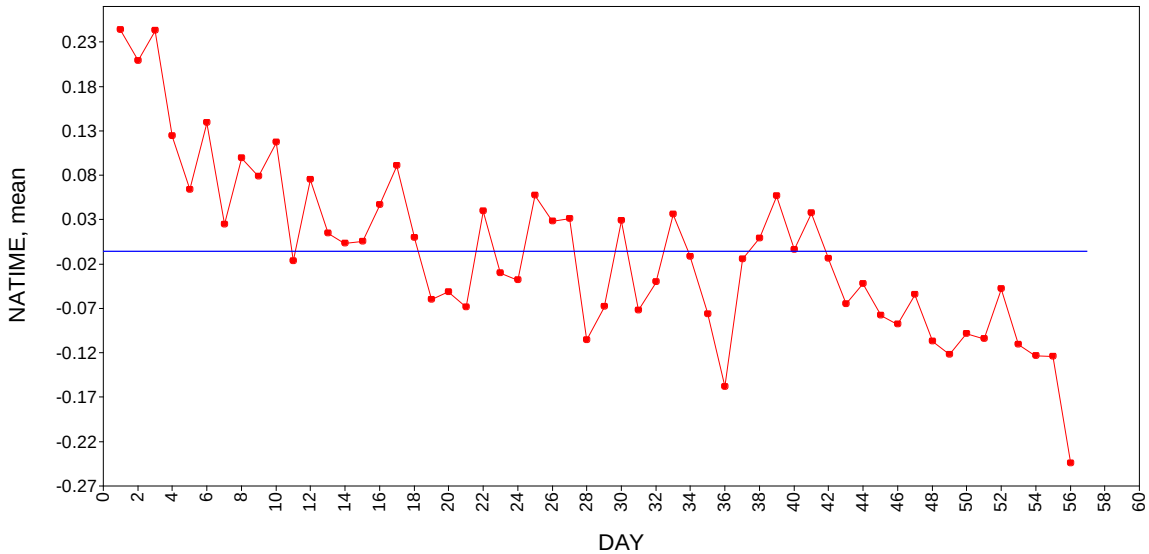
```
MODEL:      %WITHIN%
            na_w BY afraid1-distrsd1* (&1 1-10);
            na_w ON na_w&1;
            %BETWEEN SUBJECT%
            na_subj BY afraid1-distrsd1* (1-10);
            na_subj@1;
            %BETWEEN DAY%
            na_time BY afraid1-distrsd1* (1-10);
OUTPUT:     TECH1 TECH8 STDY STDYX TECH4 RESIDUAL
            FSCOMPARISON;
PLOT:       TYPE = PLOT3;
            FACTORS = ALL;
```

Run time: 54 minutes (dichotomized: 34 minutes; continuous: 16 minutes)

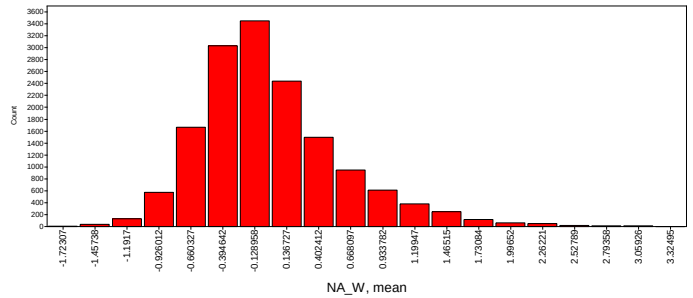
Results of Cross-Classified Factor Analysis with One NA Factor for 10 Ordinal Items

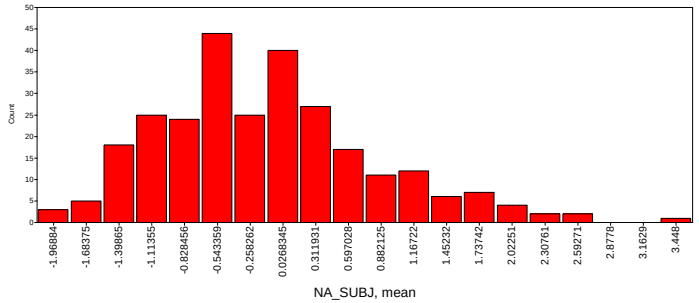
$$na\,factor_{it} = \underbrace{\alpha + \alpha_i}_{\text{Between subject}} + \underbrace{\alpha_t}_{\text{Between time}} + \underbrace{\beta\, y_{w,it-1} + \varepsilon_{it}}_{\text{Within subject}}. \quad (9)$$

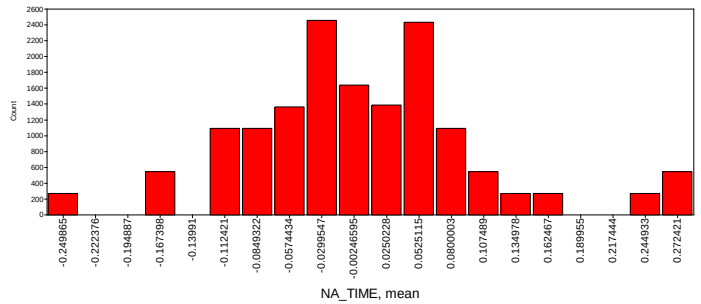
- V(na_subject) = 1.00, V(na_time) = 0.012, V(na_w) = 0.66
- The factor score plot for the na_time factor (on the between day level) shows a drop of 40% of the total factor SD over the 56 days:



Posterior Distributions for the Factor Scores on Within, Between Subject, and Between Time







DSEM Web Page

Feel free to submit your papers to be posted here:

http://www.statmodel.com/TimeSeries.shtml

References

Asparouhov, T. & Muthén, B. (2015). General random effect latent variable modeling: Random subjects, items, contexts, and parameters. In Harring, J. R., Stapleton, L. M., & Beretvas, S. N. (Eds.), Advances in multilevel modeling for educational research: Addressing practical issues found in real-world applications. Charlotte, NC: Information Age Publishing, Inc.

Asparouhov, T., Hamaker, E.L. & Muthén, B. (2017). Dynamic structural equation models. Technical Report. www.statmodel.com

Asparouhov, T., Hamaker, E.L. & Muthén, B. (2017). Dynamic latent class analysis. Structural Equation Modeling, 24, 257-269.

Bolger, N. & Laurenceau, J-P. (2013). Intensive longitudinal methods: An introduction to diary and experience sampling research. New York: Guilford.

Engle, R. & Watson, M. (1981). A one-factor multivariate time series model of metropolitan wage rates. JASA, 76, 774-781.

Fox, J.P. (2010). Bayesian item response theory. Springer.

Gelman, A., Carlin, J.B., Stern, H.S., & Rubin, D.B. (2014). Bayesian data analysis. Third edition. New York: Chapman & Hall.

Granger, C.W.J. & Morris, M.J. (1976). Time series modelling and interpretation. Journal of the Royal Statistical Society, Series A, 139, 246-257.

Hamaker, E.L. (2005). Conditions for the equivalence of the autoregressive latent trajectory model and a latent growth curve model with autoregressive disturbances. Sociological Methods & Research.

Hamaker, E.L.,Schuurman, N.K., & Ziljmans, E.A.O. (2017). Using a few snapshots to distinguish mountains from waves: Weak factorial invariance in the context of trait-state research. Multivariate Behavioral Research.

Hamaker, E.L. & Wichers, M. (2017). No time like the present: Discovering the hidden dynamics in intensive longitudinal data. Current Directions in Psychological Science.

Hu, Nesselroade et al. (2016). Test reliability at the individual level. Structural Equation Modeling.

References Continued

Jak, S., Oort, F.J., & Dolan, C.V. (2013). A test for cluster bias: Detecting violations of measurement invariance across clusters in multilevel data. Structural Equation Modeling, 20, 265-282.

Jak, S., Oort, F.J., & Dolan, C.V. (2014). Measurement Bias in Multilevel Data. Structural Equation Modeling, 21:1, 31-39, DOI: 10.1080/10705511.2014.856694

Lynch, S. M. (2010). Introduction to applied Bayesian statistics and estimation for social scientists. Springer.

Molenaar, P. (1985). A dynamic factor model for the analysis of multivariate time series. Psychometrika, 50, 181-202.

Muthén, B. (1994). Multilevel covariance structure analysis. In J. Hox & I. Kreft (eds.), Multilevel Modeling, a special issue of Sociological Methods & Research, 22, 376-398.

Muthén, B. & Asparouhov, T. (2017). Recent methods for the study of measurement invariance with many groups. Alignment and random effects. Forthcoming in a special measurement invariance issue of Sociological Methods & Research.

Muthén, B., Muthén, L, & Asparouhov, T. (2016). Regression and mediation analysis using Mplus. wwww.statmodel.com.

Schuurman, N.K., Houtveen, J.H., & Hamaker, E.L. (2015). Incorporating measurement error in n=1 psychological autoregressive modeling. Frontiers in Psychology, 6, 1-15.

Schuurman, N.K., Ferrer, E., de Boer-Sonnenschein, M., & Hamaker, E.L. (2016). How to compute cross-lagged associations in a multilevel autoregressive model. Psychological Methods, 21, 206-221.

Shiyko et al. (2012). Using the time-varying effect model (TVEM) to examine dynamic associations between negative affect and self confidence on smoking urges. Prevention Science, 13, 288-299

Shumway, R.H. & Stoffer, D.S. (2011). Time series analysis and its applications. New York: Springer.

Wall, M.M., Guo, J., & Amemiya, Y. (2012). Mixture factor analysis for approximating a nonnormally distributed continuous latent factor with continuous and dichotomous observed variables. Multivariate Behavioral Research, 47, 276-313.

Walls, T.D. & Schafer, J.L (2006). Models for intensive longitudinal data. Oxford Univ Press.

Wang, Hamaker, & Bergeman (2012). Investigating inter-individual differences in short-term intra-individual variability. Psychological Methods.