Day 3

- Change
- Systematic Mean-Level Change
- Reversible Change

Change

Change

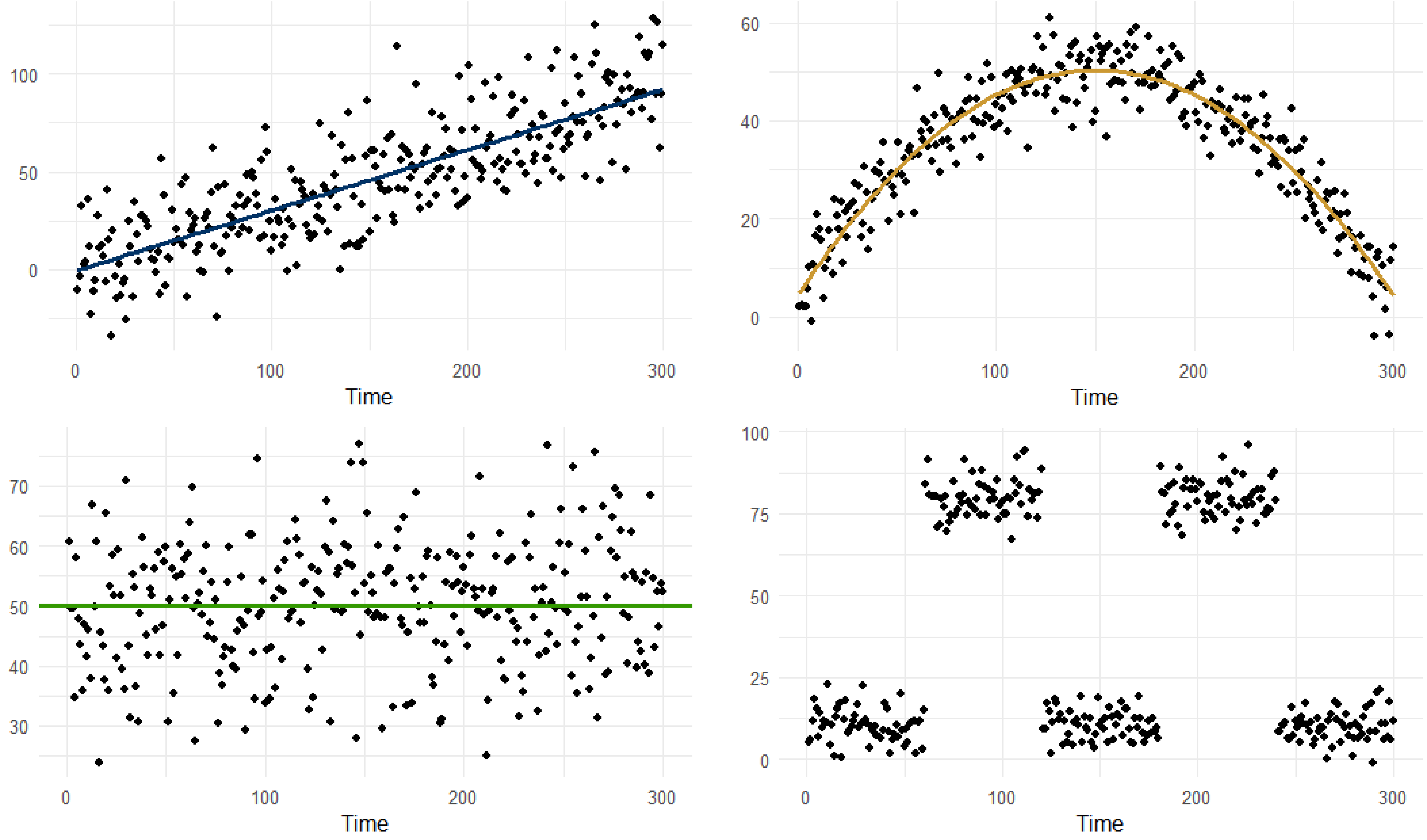
How do you define change?

Change

Introduction to Change

- Types of Constructs in Social Sciences
  - Fixed Constructs (Traits)
    - Aspects that do not change over time
      - Personality, IQ (though debatable)
  - Changing Constructs (States)
    - Aspects that do change
      - Mood, Concentration
  - Nothing is Truly Fixed
    - Over long time scales, even traits can change
      - Mountains change over time

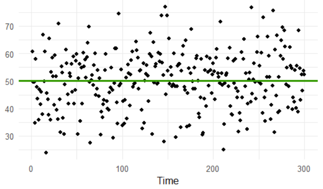
Change

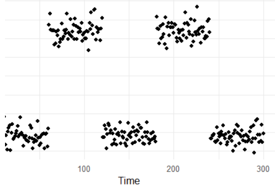

100
50
0
0
100
200
300
Time
60
40
20
0
0
100
200
300
Time
70
60
50
40
30
0
100
200
300
Time
100
75
50
25
0
0
100
200
300
Time

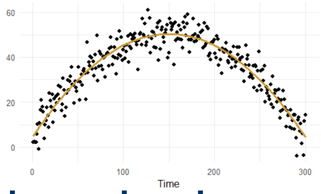
Change

Different Types of Change

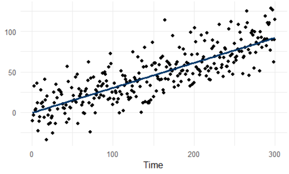
- Change isn't one thing
  - Change varies in how it manifests
- Types of Change
  - Long-term vs Short-term
- Reversible vs Non-reversible/Trend
  - Reversible changes can return to previous states







  - Non-reversible changes or trends show consistent movement in one direction



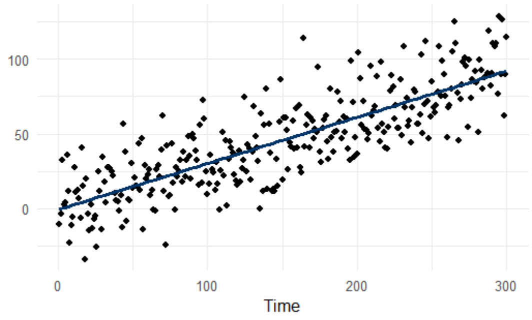
TILBURG UNIVERSITY

tesc

Change

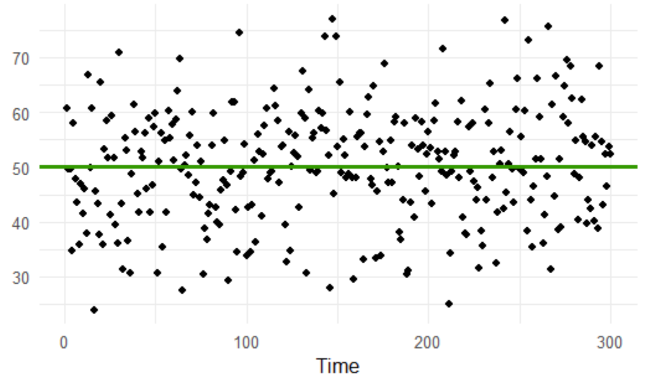
Multilevel Modeling for Investigating Systematic Mean-Level Change

- Traditional Multilevel Analysis
  - Useful for examining change across different levels
    - Within-person variability (how an individual changes over time)
    - Between-person differences (how individuals differ in their patterns of change)
  - We can model time as a predictor to study how (average) change occurs over time and whether the effects of time differ across individuals



Change

- Alternative Multilevel Analysis
  - We can also look at lagged (autoregressive) effects (using the value of a predictor from a previous time point to predict the current outcome) to model short-term fluctuations and temporal stability of the variable (e.g., how much yesterday's mood influences today's mood
  - With regular multilevel modeling and time as predictor, we would find slopes of zero, and the intercept would just be the average outcome value



Change

- Type of data needed
  - Intensive Longitudinal Data (ILD) or shorter longitudinal (e.g., panel) studies to track changes
    - ILD is characterized by many observations per person and observations taken close together in time
    - ILD is particularly helpful for revealing how individuals experience short-term fluctuations.
      - – Example: To study mood fluctuations throughout the day, collecting frequent data points is essential

Systematic Mean-Level Change

Systematic Mean-Level Change: Recall the Steps of a multilevel analysis

1. Check whether multilevel is necessary
2. Add first-level variables
3. Add second-level variables
4. Add random slopes
5. Add cross-level interactions (if random slopes are present)

Step 1: Check whether multilevel is necessary

Step 1: Check whether multilevel is necessary

- Hypotheses
  - H_0: observations of the same person aren't more alike than observations of other persons; the level-2 variance (in the intercepts) is zero
  - H_1: observations of the same person are more alike than observations of other persons; the level-2 variance (in the intercepts) is larger than zero

Step 1 in R

- The dataset *gpa* contains longitudinal data on 200 college students
- The students' grade point average (GPA; range 1.7-4) has been recorded for six successive semesters
- At the same time, it was recorded whether the student held a job in that semester, and for how many hours. This is recorded in the time-varying variable job (= hours worked)
- We also have the student-level variables
  - high school GPA
  - biological sex (0 = male, 1 = female)

Step 1 in R

```
> # Check whether multilevel is necessary (using intercept-only models)
> IO <- lm(gpa ~ 1, gpa)
> IO_ML <- lmer(gpa ~ 1 + (1 | student), gpa)
> anova(IO_ML, IO)
refitting model(s) with ML (instead of REML)
Data: gpa
Models:
IO: gpa ~ 1
IO_ML: gpa ~ 1 + (1 | student)
      npar     AIC     BIC  logLik deviance  Chisq Df Pr(>Chisq)
IO       2 1167.28 1177.46 -581.64  1163.28
IO_ML    3  919.46  934.73 -456.73   913.46 249.82  1  < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Allowing for between-person differences in the intercepts significantly improves model fit

Step 1 in R

```
> summary(IO_ML)
Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: gpa ~ 1 + (1 | student)
   Data: gpa

REML criterion at convergence: 919.5

Scaled residuals:
    Min      1Q  Median      3Q     Max
-3.6504 -0.5496  0.0603  0.6356  2.5736

Random effects:
 Groups   Name        Variance Std.Dev.
 student  (Intercept) 0.05714  0.2390
 Residual             0.09759  0.3124
Number of obs: 1200, groups:  student, 200

Fixed effects:
            Estimate Std. Error        df t value Pr(>|t|)
(Intercept)  2.86500    0.01916 198.99999   149.6   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
> # "Variance student" / ("Variance residual" + "Variance student")
> 0.05714 /(0.05714 + 0.09759) # ICC = 0.3692884
[1] 0.3692884
> # --> variance explained by between-person differences
```

Together they make up the total variance: 0.05714 of the variance is at the person level and 0.09759 is the variance within persons

Average gpa across all students and time-points

Step 2: Add first-level variables

Step 2: Add first-level variables

- We want to check if time-point-specific variables (level 1) help explain within-person variation in the outcome
  - Add all first-level variables (+ their interactions) and check whether they are significant or not

- Hypotheses:
  - H_0: There is no relation between the dependent and the explanatory variable(s)
  - H_1: There is a (positive/negative) relation between the dependent and the explanatory variable(s)

Step 2 in R

```
> # Step 2: Add first-level variables
> ## TIME
> ML_level1_a <- lmer(gpa ~ 1 + time + (1 | student), gpa)
> summary(ML_level1_a)
Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: gpa ~ 1 + time + (1 | student)
   Data: gpa

REML criterion at convergence: 408.9

Scaled residuals:
    Min      1Q  Median      3Q     Max
-3.6169 -0.6373 -0.0004  0.6361  2.8310

Random effects:
 Groups   Name        Variance Std.Dev.
 student  (Intercept) 0.06372  0.2524
 Residual             0.05809  0.2410
Number of obs: 1200, groups:  student, 200

Fixed effects:
             Estimate Std. Error        df t value Pr(>|t|)
(Intercept) 2.599e+00  2.170e-02 3.223e+02   119.8   <2e-16 ***
time        1.063e-01  4.074e-03 9.990e+02    26.1   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Correlation of Fixed Effects:
     (Intr)
time -0.469
```

There is a significant effect of *time* on gpa
(the effect may look small, but it is also a small scale)

The intercept is interesting: expected GPA at t = 0

Step 2 in R

```
> VarianceIO <- as.data.frame(VarCorr(IO_ML))
> VarianceLv1 <- as.data.frame(VarCorr(ML_level1_a))
>
> VarianceIO
       grp        var1 var2       vcov     sdcor
1  student (Intercept) <NA> 0.05713534 0.239030
2 Residual        <NA> <NA> 0.09759000 0.312394
> VarianceLv1
       grp        var1 var2       vcov     sdcor
1  student (Intercept) <NA> 0.06371891 0.2524261
2 Residual        <NA> <NA> 0.05808854 0.2410156
>
> # Explained Variance on Level 1
> (VarianceIO[2, 4] - VarianceLv1[2, 4]) / VarianceIO[2, 4]
[1] 0.4047695
> # 0.4047695
>
> # Explained Variance on Level 2
> (VarianceIO[1, 4] - VarianceLv1[1, 4]) / VarianceIO[1, 4]
[1] -0.1152277
> # Update the ICC ("Variance student" / ("Variance residual" + "Variance student"))
> 0.06372/(0.06372 + 0.05809) # ICC = 0.5231
[1] 0.5231098
```

before adding *time* as a predictor

after adding *time* as a predictor

Explained variance can't be negative! (it is due to fixed measurement occasions)

What should you do now? When adding more level-1 predictors, use the model WITH time as your reference model, and NOT the intercept-only model! The same holds for calculating the ICC.

Step 2 in R

```
> ML_level1_b <- lmer(gpa ~ 1 + time + job + (1 | student), gpa)
> summary(ML_level1_b)
Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: gpa ~ 1 + time + job + (1 | student)
   Data: gpa

REML criterion at convergence: 330

Scaled residuals: 
     Min       1Q   Median       3Q      Max 
-3.03460 -0.60192 -0.00864  0.64432  2.88770 

Random effects:
 Groups   Name        Variance Std.Dev.
 student  (Intercept) 0.05274  0.2297  
 Residual             0.05524  0.2350  
Number of obs: 1200, groups:  student, 200

Fixed effects:
              Estimate Std. Error         df t value Pr(>|t|)    
(Intercept)  2.970e+00  4.413e-02  1.187e+03  67.304   <2e-16 ***
time         1.025e-01  3.993e-03  9.932e+02  25.660   <2e-16 ***
job         -1.714e-01  1.813e-02  1.089e+03  -9.452   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Correlation of Fixed Effects:
     (Intr) time  
time -0.314       
job  -0.889  0.102
```

There is a significant effect of *job* on gpa, after controlling for time

TILBURG UNIVERSITY

tesc

Step 2 in R

```
> # Now, we re-calculate the explained variances on Level 1 and 2
> VarianceTime <- as.data.frame(VarCorr(ML_level1_a))
> VarianceLv1 <- as.data.frame(VarCorr(ML_level1_b))
>
> VarianceTime
       grp        var1 var2       vcov     sdcor
1  student (Intercept) <NA> 0.06371891 0.2524261
2 Residual        <NA> <NA> 0.05808854 0.2410156
> VarianceLv1
       grp        var1 var2       vcov     sdcor
1  student (Intercept) <NA> 0.05274081 0.2296537
2 Residual        <NA> <NA> 0.05523791 0.2350275
>
> # Explained Variance on Level 1
> (VarianceTime[2, 4] - VarianceLv1[2, 4]) / VarianceTime[2, 4]
[1] 0.04907393
> # 0.04907393
>
> # Explained Variance on Level 2
> (VarianceTime[1, 4] - VarianceLv1[1, 4]) / VarianceTime[1, 4]
[1] 0.1722895
> # 0.1722895
```

We use the model with time as baseline model

TILBURG UNIVERSITY

tesc

Step 3: Add second-level variables

Step 3: Second-level variables

- We want to check if person-level variables (level 2) can explain between-person variation in the outcome
  - Add all second-level variables (+ their interactions) and check whether they are significant or not

- Hypotheses:
  - H_0: There is no relation between the explanatory variable(s) and the mean score of the dependent variable
  - H_1: There is a (positive/negative) relation between the explanatory variable(s) and the mean score of the dependent variable

Step 3 in R

```
> # Step 3: Add second-level variables
> ML_level1_2 <- lmer(gpa ~ 1 + time + job + (1 | student) + highgpa + sex, gpa)
> summary(ML_level1_2)# both person-level variables have significant partial effects
Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: gpa ~ 1 + time + job + (1 | student) + highgpa + sex
   Data: gpa

REML criterion at convergence: 314.8

Scaled residuals:
     Min       1Q   Median       3Q      Max
-2.92840 -0.59430 -0.02621  0.63278  2.95643

Random effects:
 Groups   Name        Variance Std.Dev.
 student  (Intercept) 0.04582  0.2141
 Residual             0.05524  0.2350
Number of obs: 1200, groups:  student, 200

Fixed effects:
              Estimate Std. Error         df t value Pr(>|t|)
(Intercept)  2.640e+00  9.817e-02  2.843e+02  26.897  < 2e-16 ***
time         1.025e-01  3.993e-03  9.932e+02  25.657  < 2e-16 ***
job         -1.718e-01  1.809e-02  1.097e+03  -9.497  < 2e-16 ***
highgpa      8.471e-02  2.798e-02  1.913e+02   3.028   0.0028 **
sex          1.473e-01  3.331e-02  1.913e+02   4.421 1.65e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Correlation of Fixed Effects:
        (Intr) time   job    highgp
        -0.144
        -0.428  0.102
    gpa -0.876  0.003  0.029
        -0.251  0.003  0.027  0.073
```

Both partial effects are positive and significant

Step 3 in R

```
> VarianceLV12 <- as.data.frame(VarCorr(ML_level1_2))
>
> VarianceLv1
       grp        var1 var2       vcov     sdcor
1  student (Intercept) <NA> 0.05274081 0.2296537
2 Residual        <NA> <NA> 0.05523791 0.2350275
> VarianceLV12
       grp        var1 var2       vcov     sdcor
1  student (Intercept) <NA> 0.04582142 0.2140594
2 Residual        <NA> <NA> 0.05524169 0.2350355
>
> # Explained Variance on Level 1 (compared to model with just level-1 predictors)
> (VarianceLv1[2, 4] - VarianceLV12[2, 4]) / VarianceLv1[2, 4]
[1] -6.84343e-05
> # 0. No explained variance since you only added a level 2 predictor
>
> # Explained Variance on Level 2 (compared to model with just level 1 predictors)
> (VarianceLv1[1, 4] - VarianceLV12[1, 4]) / VarianceLv1[1, 4]
[1] 0.1311962
> # 0.1311962
```

No explained variance on level 1: a higher-level predictor can never explain variance on the lower level(s)!
(e.g., sex cannot explain differences within a person).

If desired, you could do this change in explained variance computation separately!

Step 4: Add random slopes

Step 4: Add random slopes

- We want to check if the relationship between level 1 predictors and the outcome varies across individuals (level 2).
- We allow the regression coefficients for the level 1 predictors to differ across the higher units.
  - We check whether the variance in these regression coefficients is significantly different from zero.

- Hypotheses:
  - H_0: The relation between the explanatory variable and the dependent variable is the same within all level-2 units (the variance is = 0)
  - H_1: The relation between the explanatory variable and the dependent variable is not the same within all level-2 units (the variance is > 0)

Step 4 in R

```
> # Step 4: Add random slopes
>
> # # you do this per lower level variable;
> ML_level1_2_RE_a <- lmer(gpa ~ 1 + time + job + (1 + time | student) + highgpa + sex, gpa)
Warning message:
In checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv,  :
  Model failed to converge with max|grad| = 0.00407156 (tol = 0.002, component 1)
> anova(ML_level1_2_RE_a, ML_level1_2)
refitting model(s) with ML (instead of REML)
Data: gpa
Models:
ML_level1_2: gpa ~ 1 + time + job + (1 | student) + highgpa + sex
ML_level1_2_RE_a: gpa ~ 1 + time + job + (1 + time | student) + highgpa + sex
                 npar    AIC    BIC   logLik deviance  Chisq Df Pr(>Chisq)
ML_level1_2         7 296.76 332.39 -141.380   282.76
ML_level1_2_RE_a    9 188.12 233.93  -85.059   170.12 112.64  2  < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
>
> ML_level1_2_RE_b <- lmer(gpa ~ 1 + time + job + (1 +    job| student) + highgpa + sex, gpa)
boundary (singular) fit: see help('isSingular')
> anova(ML_level1_2_RE_b, ML_level1_2)
refitting model(s) with ML (instead of REML)
Data: gpa
Models:
ML_level1_2: gpa ~ 1 + time + job + (1 | student) + highgpa + sex
ML_level1_2_RE_b: gpa ~ 1 + time + job + (1 + job | student) + highgpa + sex
                 npar    AIC    BIC  logLik deviance  Chisq Df Pr(>Chisq)
ML_level1_2         7 296.76 332.39 -141.38   282.76
ML_level1_2_RE_b    9 288.40 334.21 -135.20   270.40 12.356  2   0.002075 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Step 4 in R

```
> ML_level1_2_RE <- lmer(gpa ~ 1 + time + job + (1 +  time + job| student) + highgpa + sex, gpa)
boundary (singular) fit: see help('isSingular')
> summary(ML_level1_2_RE)
Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: gpa ~ 1 + time + job + (1 + time + job | student) + highgpa +      sex
   Data: gpa

REML criterion at convergence: 196.2

Scaled residuals:
    Min      1Q  Median      3Q     Max 
-3.0139 -0.5416 -0.0092  0.5466  3.4026 

Random effects:
 Groups   Name        Variance Std.Dev. Corr       
 student  (Intercept) 0.065634 0.25619             
          time        0.003682 0.06068   0.16      
          job         0.002346 0.04844  -0.70 -0.82
 Residual             0.041464 0.20363             
Number of obs: 1200, groups:  student, 200

Fixed effects:
              Estimate Std. Error         df t value Pr(>|t|)    
(Intercept)   2.572847   0.092942 288.665553  27.682  < 2e-16 ***
time          0.101536   0.005511 195.746557  18.424  < 2e-16 ***
job          -0.138601   0.017446 729.748983  -7.945 7.38e-15 ***
highgpa       0.087399   0.026229 195.218476   3.332 0.001031 ** 
sex           0.122321   0.031133 192.789734   3.929 0.000119 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Step 5: Add cross-level interactions

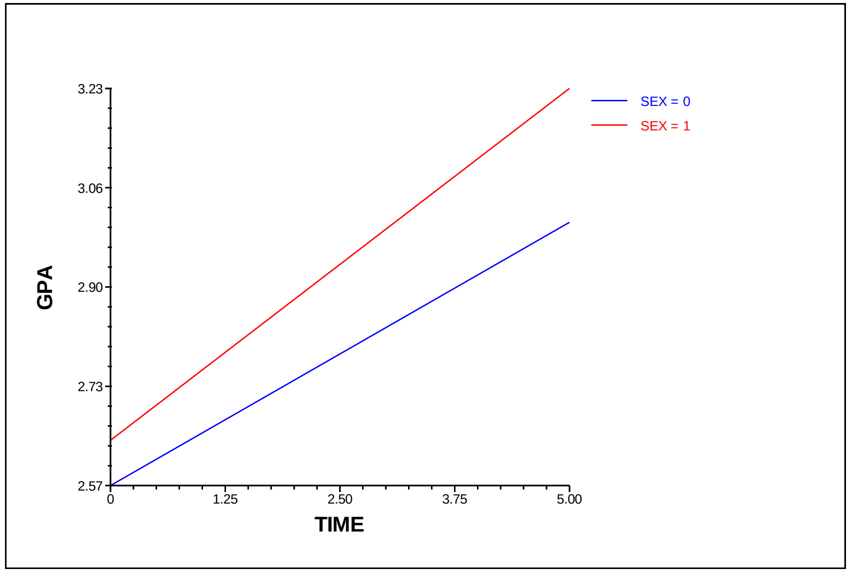
Step 5: Add cross-level interactions

- We want to see if the variance in slopes across level-2 groups can be explained by level-2 predictors.
- Add cross-level interactions between level-1 explanatory variables that had significant slope variation and level-2 explanatory variables

- Hypotheses:
  - H_0: The explanatory variable cannot explain the variance in the relations between the explanatory variable and the dependent variable in different level-2 units
  - H_1: the explanatory variable explains (a part of) the variance in the relations between the explanatory variable and the dependent variable in different level-2 units

Step 5: Add cross-level interactions

```
> # Step 5: Add cross-level interactions (let's only take sex)
> ML_level1_2_RE_CL <- lmer(gpa ~ 1 + time * sex + job * sex +
+                               (1 +  time + job| student) + highgpa + sex, gpa)
boundary (singular) fit: see help('isSingular')
>
> summary(ML_level1_2_RE_CL)
Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: gpa ~ 1 + time * sex + job * sex + (1 + time + job | student) +      highgpa + sex
   Data: gpa

REML criterion at convergence: 198.1

Scaled residuals:
    Min      1Q  Median      3Q     Max
-2.9992 -0.5323 -0.0123  0.5425  3.3969

Random effects:
 Groups   Name        Variance Std.Dev. Corr
 student  (Intercept) 0.064339 0.25365
          time        0.003475 0.05895   0.22
          job         0.002638 0.05136  -0.67 -0.87
 Residual             0.041344 0.20333
Number of obs: 1200, groups:  student, 200

Fixed effects:
             Estimate Std. Error         df t value Pr(>|t|)
(Intercept)  2.642842   0.101036 353.620232  26.157  < 2e-16 ***
time         0.085139   0.007855 195.051992  10.838  < 2e-16 ***
sex         -0.016621   0.084899 466.484172  -0.196  0.84487
job         -0.157632   0.023791 633.471172  -6.626 7.39e-11 ***
highgpa      0.086563   0.026291 195.051234   3.292  0.00118 **
time:sex     0.031136   0.010840 194.917726   2.872  0.00453 **
sex:job      0.041768   0.035015 673.830567   1.193  0.23334
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```



Reversible Change

Reversible Change

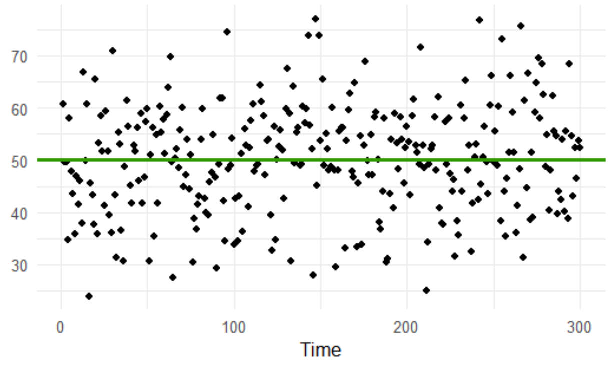
- In a previous video we mentioned that ILD is “Ideal for studying short-term, reversible changes and variability”
  - E.g., mood fluctuations throughout the day.

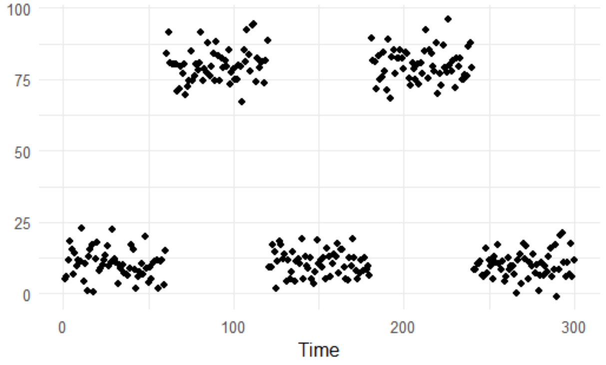
- What do we mean again with short-term reversible change?

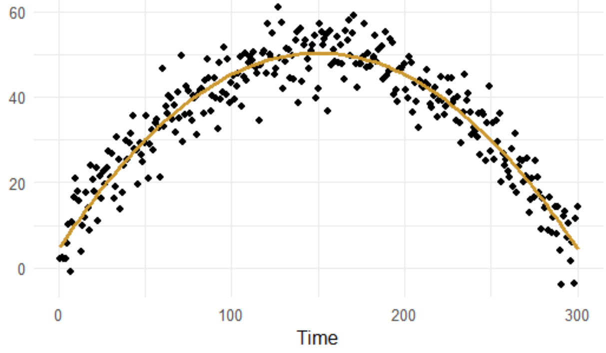
- How is it different from the mean level change discussed in a previous video?

Reversible Change

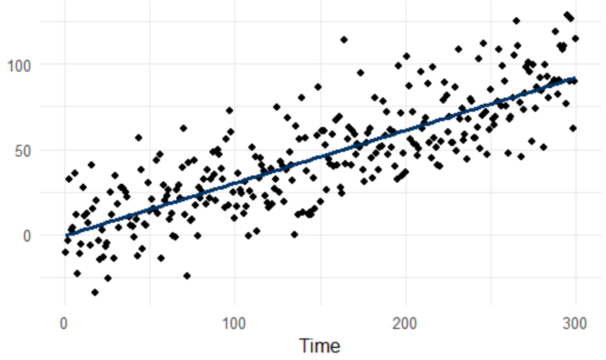
- Reversible changes can return to previous states





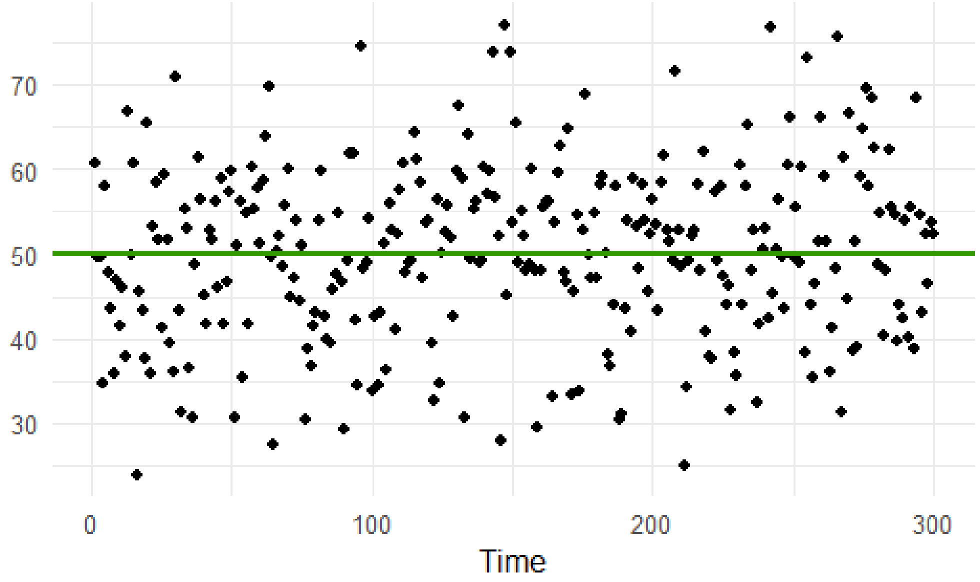


- Non-reversible changes or trends show consistent movement in one direction



Reversible Change

- Often, ILD studies study change in which the mean does not change over time at all (or in which this type of change is of secondary interest).
- So here, *time* is not a (focal) predictor.
- But how do we predict fluctuations in scores then?

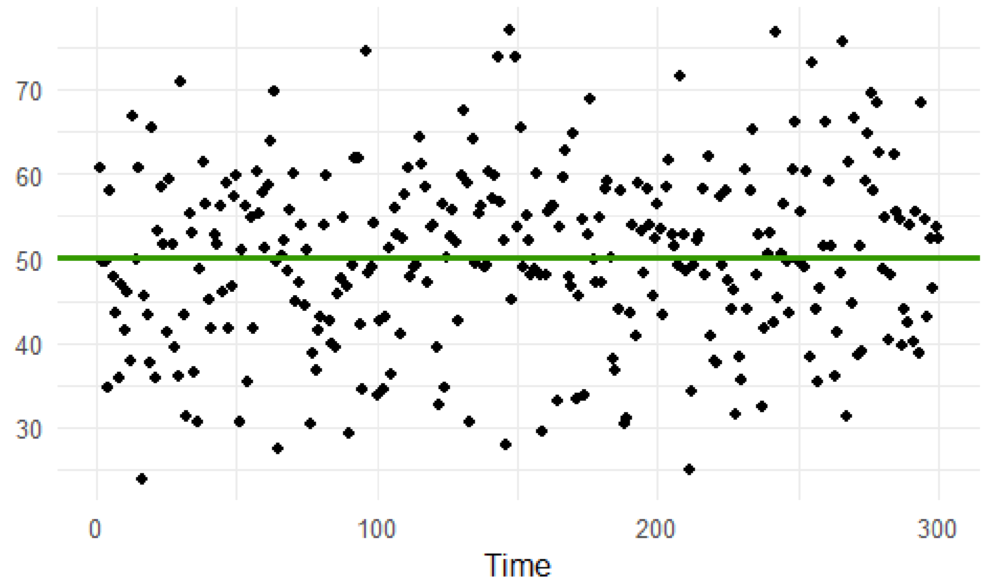


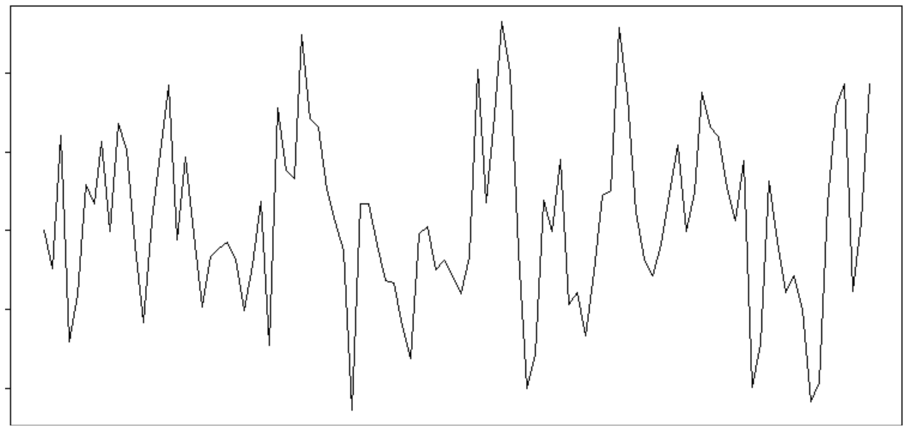
TILBURG UNIVERSITY

tesc

Reversible Change

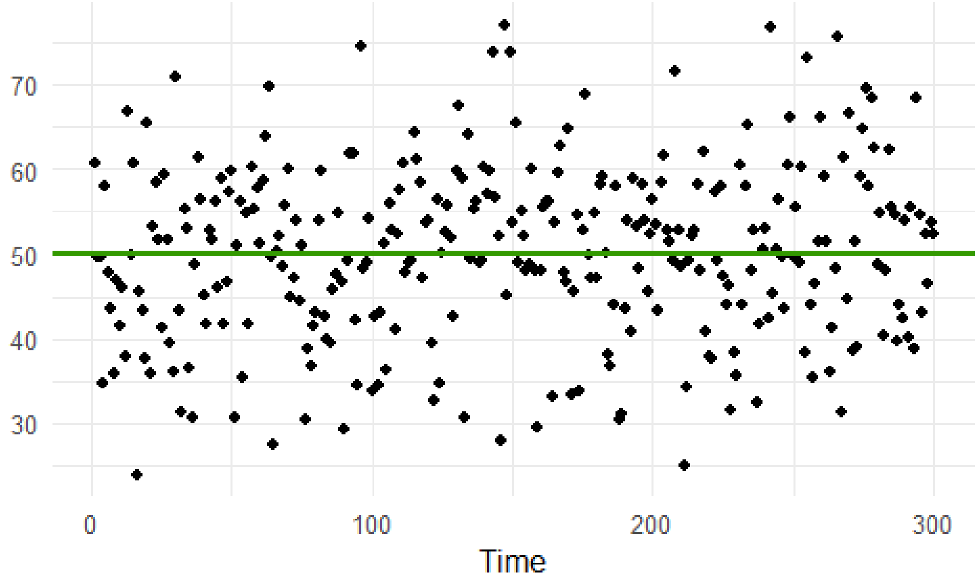
- Most often, ILD studies study change in which the mean does not change over time at all.
- So here, *time* is not a predictor.
- But how do we predict fluctuations in scores then?
- Maybe easier to see if we visualize the data slightly differently.

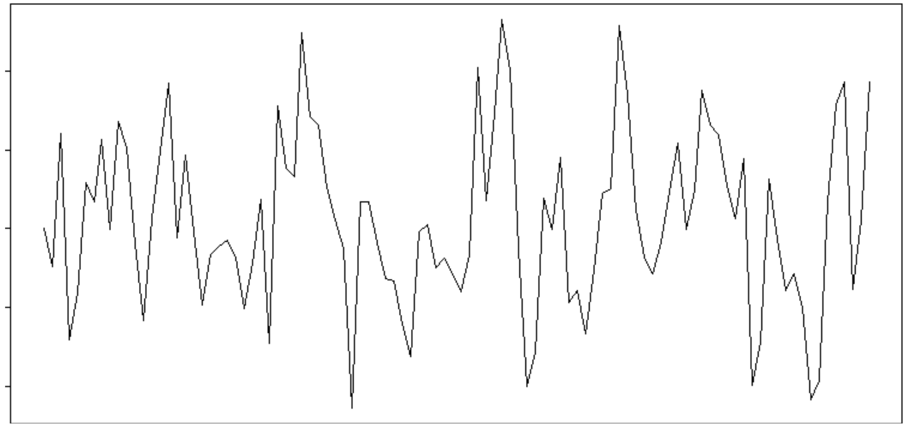




Reversible Change

- In essence, we try to predict the current score from the previous scores.
- Most often from the immediate preceding score.





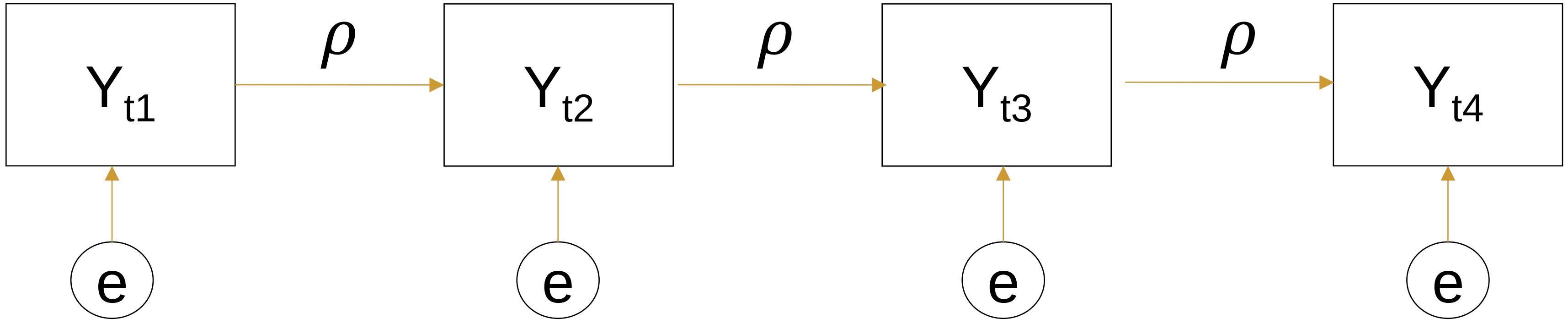
TILBURG UNIVERSITY

tesc

Reversible Change

- A model in which we predict the current score from the immediate preceding one is called a first-order autoregressive (AR(1)) model.

- Visually, the AR(1) model can be depicted as:

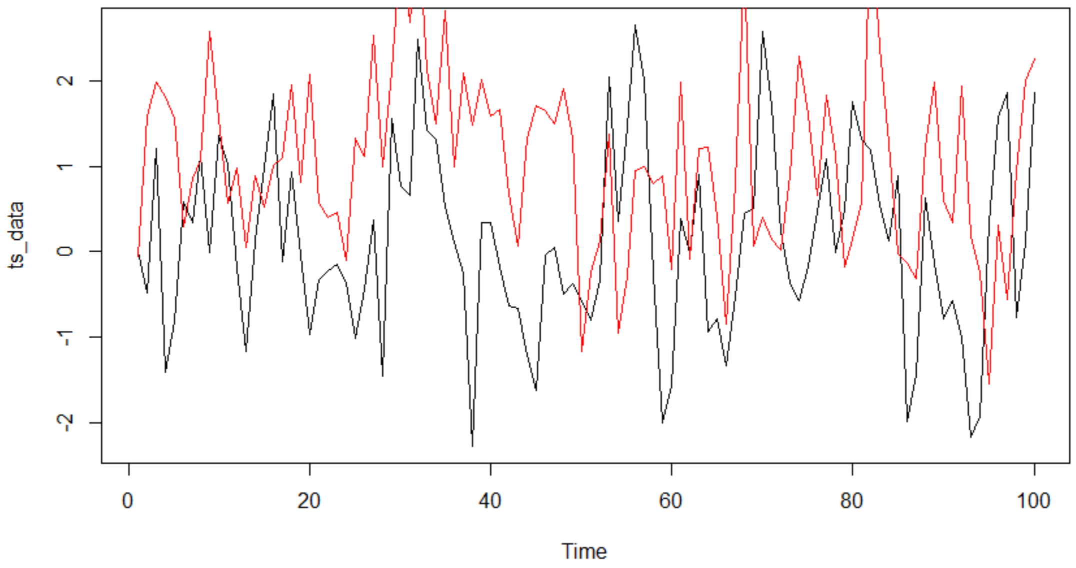


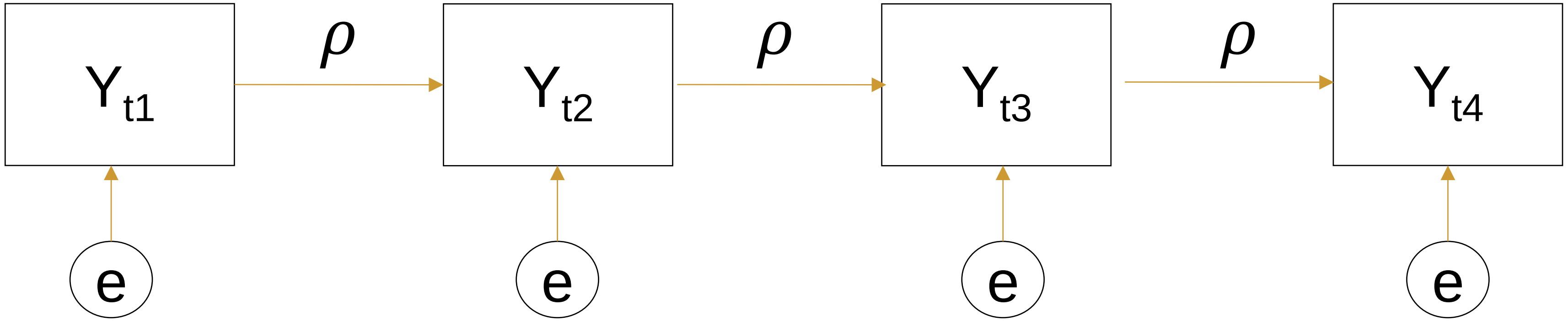
Reversible Change

How are differences in change between people reflected in the AR model?

Reversible Change

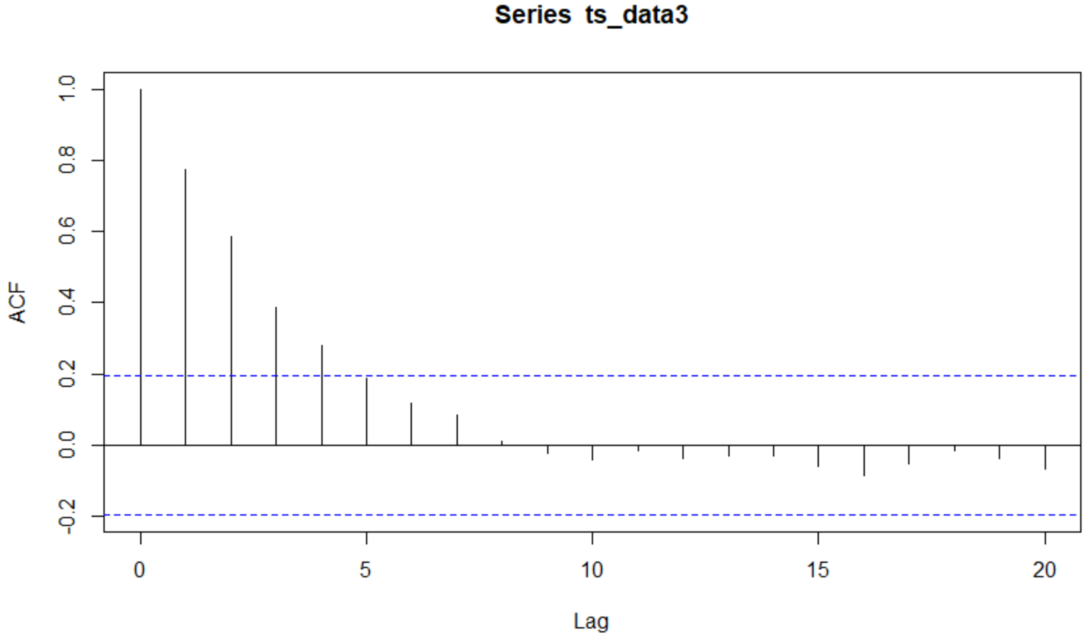
- How is a difference in means reflected?

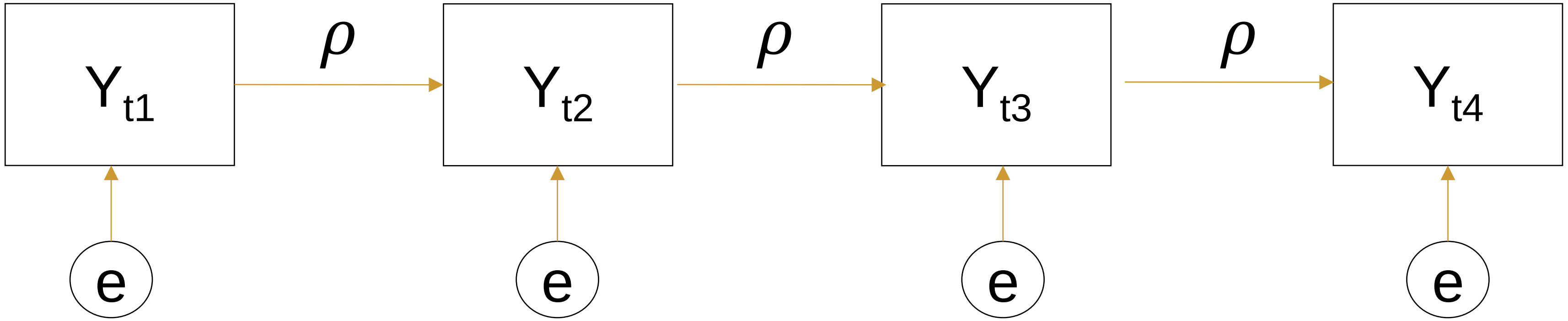




Reversible Change

- And a difference in autocorrelation (i.e., how much the current score depends on the previous one)?





Reversible Change

- Now how do we fit this AR(1) model?
  - And allow for these differences in means and autocorrelation between people

- Turns out we can do that using the methods we already learned.

Reversible Change

Level 1: $$Y_{it} = b_{intercept\_i} + b_{X\_i}X_{it} + e_{it}$$

Level 2: $$b_{intercept\_i} = \gamma_{intercept\_intercept} + u_{intercept\_i}$$

$$b_{X\_i} = \gamma_{intercept\_X} + u_{X\_i}$$

Reversible Change

Level 1: $$Y_{it} = b_{intercept\_i} + \rho_i Y_{i,t-1} + e_{it}$$

Level 2: $$b_{intercept\_i} = \gamma_{intercept\_intercept} + u_{intercept\_i}$$

$$\rho_i = \gamma_{intercept\_\rho} + u_{\rho\_i}$$

Reversible Change

File Edit View Data Transform Analyze G

81 : Y

	Y	X	var
1	2,18	.	
2	3,93	2,18	
3	3,45	3,93	
4	3,28	3,45	
5	,29	3,28	
6	,47	,29	
7	3,37	,47	
8	7,27	3,37	
9	3,63	7,27	
10	3,37	3,63	
11	3,32	3,37	
12	1,68	3,32	
13	6,85	1,68	
14	2,26	6,85	
15	3,11	2,26	
16	,43	3,11	
17	,95	,43	
18	6,03	,95	
19	5,47	6,03	
20	5,17	5,47	
21	4,03	5,17	
22	2,24	4,03	

Data View Variable View

Reversible Change

File Edit View Data Transform Analyze G

81 : Y

	Y	X	var
1	2,18	.	
2	3,93	2,18	
3	3,45	3,93	
4	3,28	3,45	
5	,29	3,28	
6	,47	,29	
7	3,37	,47	
8	7,27	3,37	
9	3,63	7,27	
10	3,37	3,63	
11	3,32	3,37	
12	1,68	3,32	
13	6,85	1,68	
14	2,26	6,85	
15	3,11	2,26	
16	,43	3,11	
17	,95	,43	
18	6,03	,95	
19	5,47	6,03	
20	5,17	5,47	
21	4,03	5,17	
22	2,24	4,03	

Data View Variable View

Reversible Change

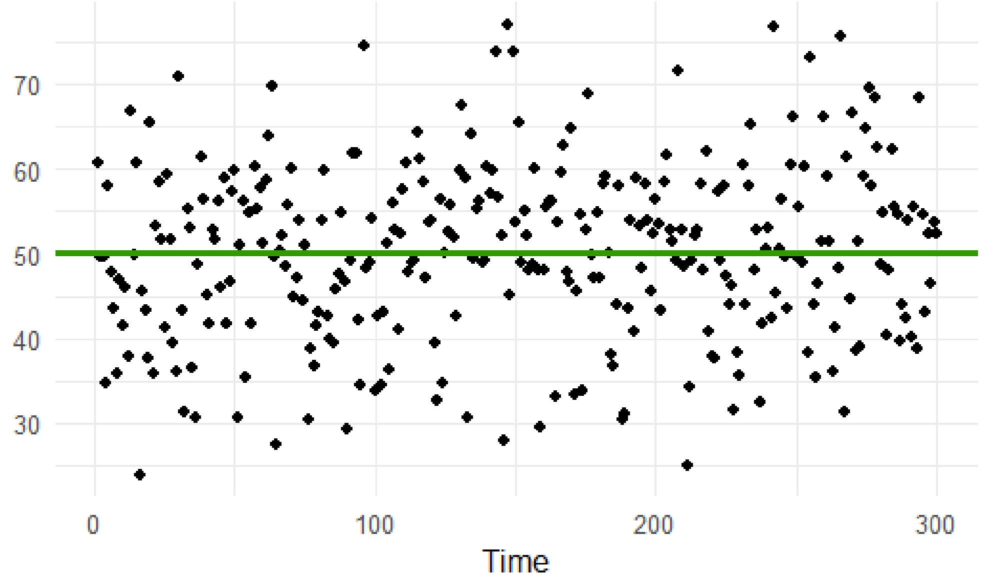
Level 1: $$Y_{it} = b_{intercept\_i} + \rho_i Y_{i,t-1} + e_{it}$$

Level 2: $$b_{intercept\_i} = \gamma_{intercept\_intercept} + u_{intercept\_i}$$

$$\rho_i = \gamma_{intercept\_\rho} + u_{\rho\_i}$$

- Couple of things:
  - The intercept is not the mean:

$$\mu_i = \frac{b_{intercept\_i}}{1 - \rho_i^2}$$



Reversible Change

```
head(VARData)
mean(VARData$Y1)
```

```
> head(VARData)
# A tibble: 6 x 6
  individual     Y1     Y2  Y1lag   Y2lag  time
       <dbl>  <dbl>  <dbl>  <dbl>   <dbl> <int>
1          1  1.10  -0.124 NA      NA         1
2          1  0.588  1.86   1.10   -0.124     2
3          1  2.99   2.20   0.588   1.86      3
4          1  3.68   2.39   2.99    2.20      4
5          1  2.34   2.58   3.68    2.39      5
6          1  2.67   4.79   2.34    2.58      6
> mean(VARData$Y1)
[1] 3.908995
```

TILBURG UNIVERSITY

tesc

Reversible Change

```
AR1 <- brm(Y1 ~ Y1lag + (1 + Y1lag | individual),
           iter = 5000, data = VARData)
summary(AR1)
```

TILBURG UNIVERSITY

tesc

Reversible Change

```
> summary(AR1)
 Family: gaussian
  Links: mu = identity; sigma = identity
Formula: Y1 ~ Y1lag + (1 + Y1lag | individual)
   Data: VARData (Number of observations: 4900)
  Draws: 4 chains, each with iter = 5000; warmup = 2500; thin = 1;
         total post-warmup draws = 10000

Multilevel Hyperparameters:
~individual (Number of levels: 100)
                     Estimate Est.Error l-95% CI u-95% CI Rh
sd(Intercept)            0.33      0.08     0.17     0.48 1.
sd(Y1lag)                0.07      0.02     0.02     0.11 1.
cor(Intercept,Y1lag)    -0.49      0.33    -0.82     0.53 1.

Regression Coefficients:
          Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS
Intercept     2.15      0.06     2.03     2.28 1.00     8452
Y1lag         0.46      0.01     0.43     0.49 1.00     7258

Further Distributional Parameters:
      Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
sigma     1.02      0.01     1.00     1.04 1.00    15145     6625

Draws were sampled using sampling(NUTS). For each parameter, Bulk_ESS
and Tail_ESS are effective sample size measures, and Rhat is the potential
scale reduction factor on split chains (at convergence, Rhat = 1).
```

```
> head(VARData)
# A tibble: 6 × 6
  individual    Y1     Y2 Y1lag  Y2lag  time
       <dbl> <dbl>  <dbl> <dbl>  <dbl> <int>
1          1 1.10  -0.124 NA     NA        1
2          1 0.588  1.86  1.10   -0.124    2
3          1 2.99   2.20  0.588  1.86     3
4          1 3.68   2.39  2.99   2.20     4
5          1 2.34   2.58  3.68   2.39     5
6          1 2.67   4.79  2.34   2.58     6
> mean(VARData$Y1)
[1] 3.908995
```

TILBURG UNIVERSITY

tesc

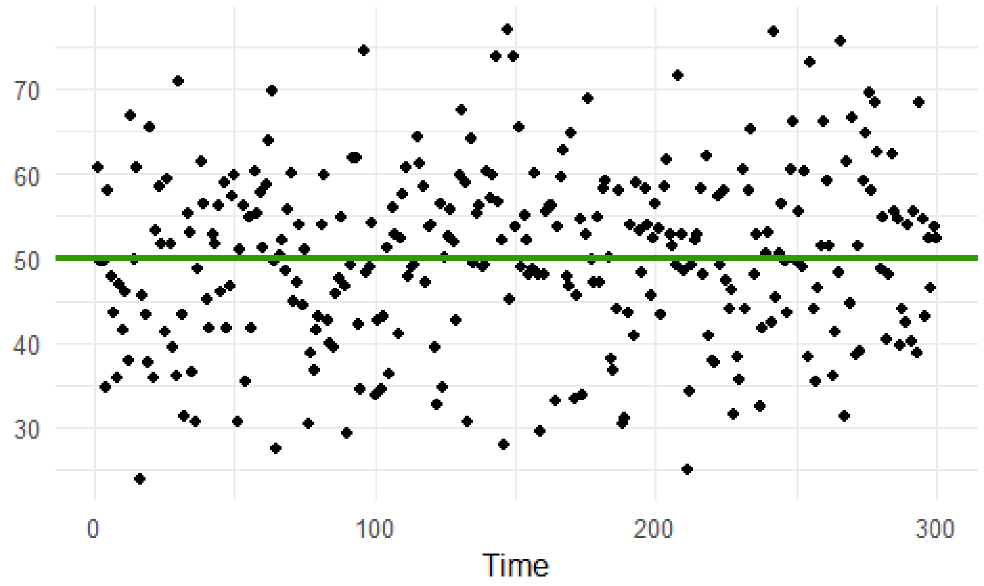
Reversible Change

Level 1: $Y_{it} = b_{intercept\_i} + \rho_i Y_{i,t-1} + e_{it}$

Level 2: $b_{intercept\_i} = \gamma_{intercept\_intercept} + u_{intercept\_i}$

$\rho_i = \gamma_{intercept\_\rho} + u_{\rho\_i}$

- Couple of things:
  - Can get the mean directly by group-mean centering the lagged predictor.



TILBURG UNIVERSITY

tesc

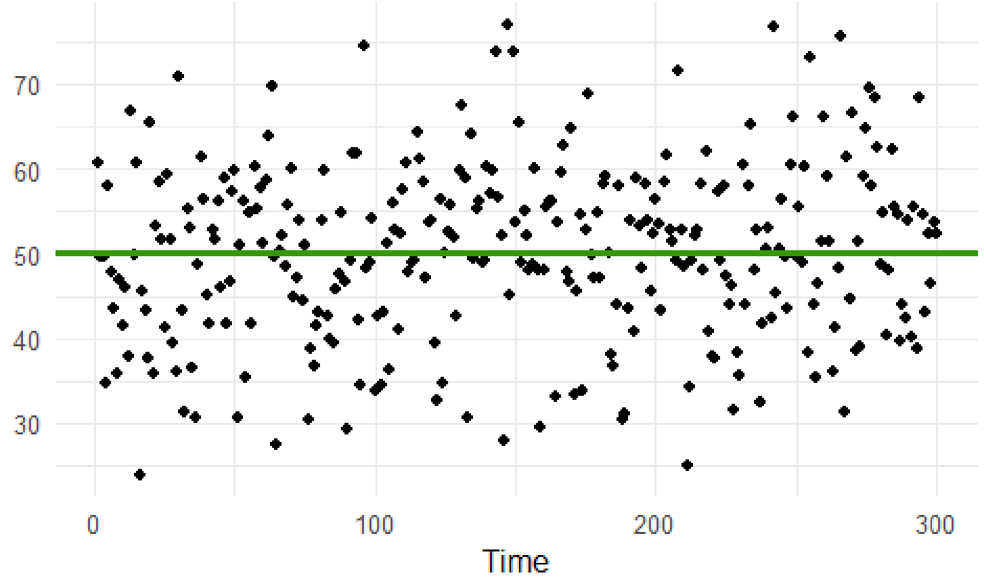
Reversible Change

Level 1: $$Y_{it} = \mu_i + \rho_i(Y_{i,t-1} - \mu_i) + e_{it}$$

Level 2: $$\mu_i = \gamma_{intercept\_\mu} + u_{\mu\_i}$$

$$\rho_i = \gamma_{intercept\_\rho} + u_{\rho\_i}$$

- Couple of things:
  - Can get the mean directly by group-mean centering the lagged predictor.



Reversible Change

```
VARData <- VARData %>%
  group_by(individual) %>%
  mutate(Y1lag_c = Y1lag - mean(Y1lag, na.rm=T))


AR1_c <- brm(Y1 ~ Y1lag_c + (1 + Y1lag_c | individual ),
             iter = 5000, data = VARData)
summary(AR1_c)
```

TILBURG UNIVERSITY

tesc

Reversible Change

```
> summary(AR1_c)
 Family: gaussian
  Links: mu = identity; sigma = identity
Formula: Y1 ~ Y1lag_c + (1 + Y1lag_c | individual)
   Data: VARData (Number of observations: 4900)
  Draws: 4 chains, each with iter = 5000; warmup = 2500; thin = 1;
         total post-warmup draws = 10000

Multilevel Hyperparameters:
~individual (Number of levels: 100)
                       Estimate Est.Error l-95% CI u-95% CI
sd(Intercept)              0.56      0.04     0.48     0.65
sd(Y1lag_c)                0.09      0.02     0.05     0.13
cor(Intercept,Y1lag_c)     0.14      0.18    -0.20     0.49

Regression Coefficients:
          Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS
Intercept     3.96      0.06     3.85     4.07 1.00       79
Y1lag_c       0.43      0.02     0.40     0.46 1.00      752

Further Distributional Parameters:
      Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
sigma     1.01      0.01     0.99     1.03 1.00    12167     7269

Draws were sampled using sampling(NUTS). For each parameter, Bulk_ESS
and Tail_ESS are effective sample size measures, and Rhat is the potential
scale reduction factor on split chains (at convergence, Rhat = 1).
>
```

```
> head(VARData)
# A tibble: 6 × 6
  individual    Y1     Y2  Y1lag  Y2lag  time
       <dbl> <dbl>  <dbl>  <dbl>  <dbl> <int>
1          1 1.10  -0.124 NA     NA         1
2          1 0.588  1.86   1.10  -0.124     2
3          1 2.99   2.20   0.588  1.86      3
4          1 3.68   2.39   2.99   2.20      4
5          1 2.34   2.58   3.68   2.39      5
6          1 2.67   4.79   2.34   2.58      6
> mean(VARData$Y1)
[1] 3.908995
```

Tilburg University

tesc

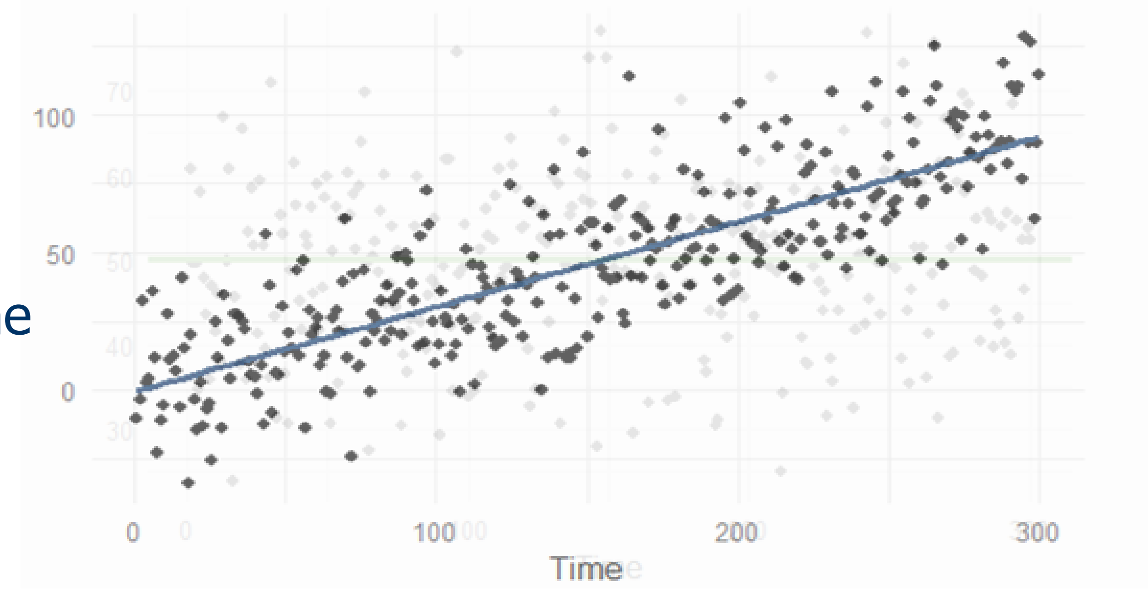
Reversible Change

Level 1: $Y_{it} = b_{intercept\_i} + \rho_i Y_{i,t-1} + e_{it}$

Level 2: $b_{intercept\_i} = \gamma_{intercept\_intercept} + u_{intercept\_i}$

$\rho_i = \gamma_{intercept\_\rho} + u_{\rho\_i}$

- Couple of things:
  - The model assumes there is no mean level change, if there is, you need to correct for the trend.



Reversible Change

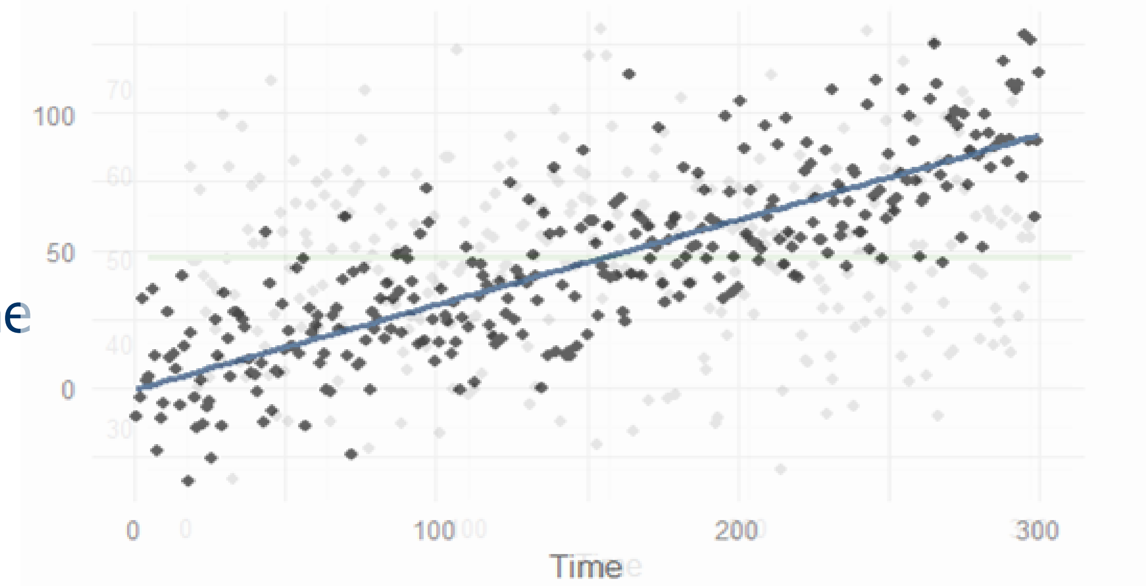
Level 1: $Y_{it} = b_{intercept\_i} + b_{time\_i} Time_{it} + e_{it}$

Level 2: $b_{intercept\_i} = \gamma_{intercept\_intercept} + u_{intercept\_i}$

$b_{time\_i} = \gamma_{intercept\_time} + u_{time\_i}$

$e_{it} = \rho_i e_{i,t-1}$

- Couple of things:
  - The model assumes there is no mean level change, if there is, you need to correct for the trend.



TILBURG UNIVERSITY

tesc

Reversible Change

```
Mean_and_Reversible_Change_Model <- brm(gpa ~ 1 + time + ar(p = 1) +
                                        (1 + time | student),
                                         iter = 2000,
                                         control = list(adapt_delta = 0.8), gpa)

summary(Mean_and_Reversible_Change_Model)
```

Reversible Change

```
> summary(Mean_and_Reversible_Change_Model)
 Family: gaussian
  Links: mu = identity; sigma = identity
Formula: gpa ~ 1 + time + ar(p = 1) + (1 + time | student)
   Data: gpa (Number of observations: 1200)
  Draws: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
         total post-warmup draws = 4000

Correlation Structures:
      Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
ar[1]     0.08      0.05    -0.02     0.18 1.00      909     1908

Multilevel Hyperparameters:
~student (Number of levels: 200)
                    Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
sd(Intercept)           0.20      0.02     0.17     0.24 1.00     1146     2123
sd(time)                0.07      0.01     0.05     0.08 1.00     1000     2274
cor(Intercept,time)    -0.04      0.13    -0.28     0.23 1.00      922     1721

Regression Coefficients:
          Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
Intercept     2.60      0.02     2.56     2.63 1.00     3336     3471
time          0.11      0.01     0.10     0.12 1.00     3171     3256

Further Distributional Parameters:
      Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
sigma     0.21      0.01     0.20     0.22 1.00      952     1923

Draws were sampled using sampling(NUTS). For each parameter, Bulk_ESS
and Tail_ESS are effective sample size measures, and Rhat is the potential
scale reduction factor on split chains (at convergence, Rhat = 1).
```

Tilburg University

tesc

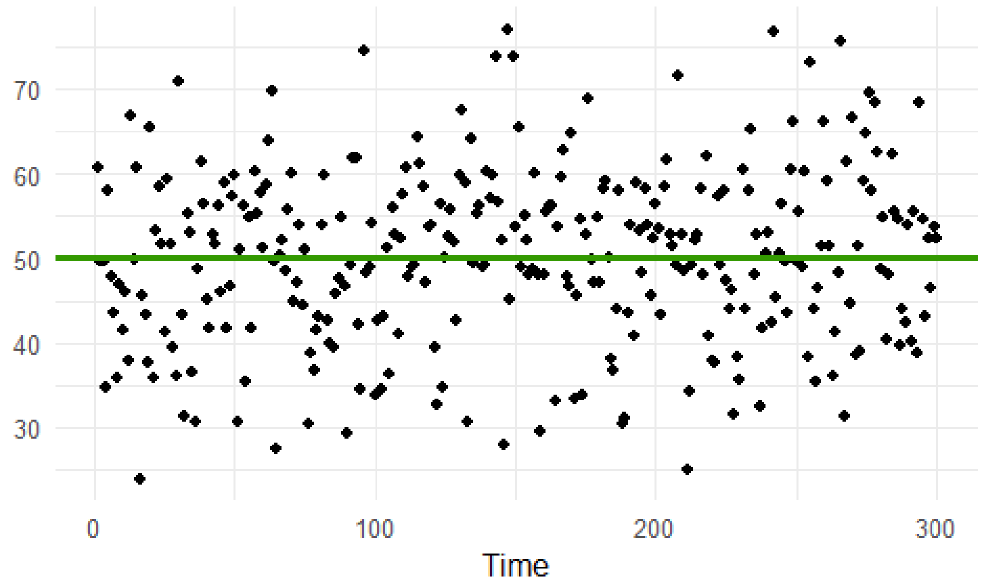
Reversible Change

Level 1: $Y_{it} = b_{intercept\_i} + \rho_i Y_{i,t-1} + e_{it}$

Level 2: $b_{intercept\_i} = \gamma_{intercept\_intercept} + u_{intercept\_i}$

$\rho_i = \gamma_{intercept\_\rho} + u_{\rho\_i}$

- Couple of things:
  - Model assumed that time between measurements is the same!
  - If that's not the case you need to account for that (e.g., including "missing data" between scores).



Reversible Change

ID	TIME	Y_{it}	$Y_{i,\ t-1}$
1	1	5.8	
1	2	6.4	5.8
1	4	5.2	
2	1	6.2	
2	2	6.7	6.2
2	3	5.2	6.7
2	4	5.8	5.3

Reversible Change

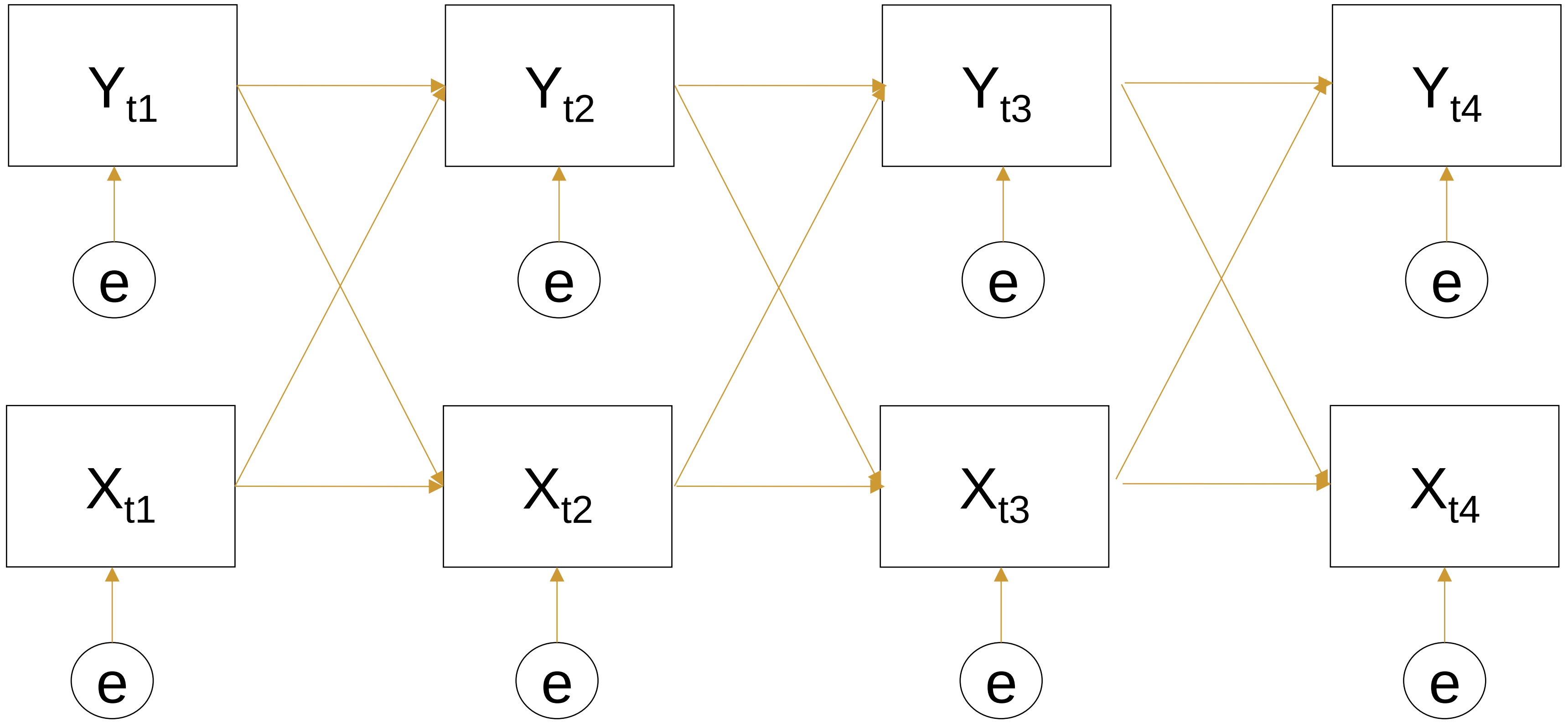
ID	TIME	Y_{it}	$Y_{i, t-1}$
1	1	5.8	
1	2	6.4	5.8
1	3		6.4
1	4	5.2	
2	1	6.2	
2	2	6.7	6.2
2	3	5.2	6.7
2	4	5.8	5.3

Reversible Change

- So far looked at the standard multilevel AR(1) model.
- There are some common extensions.

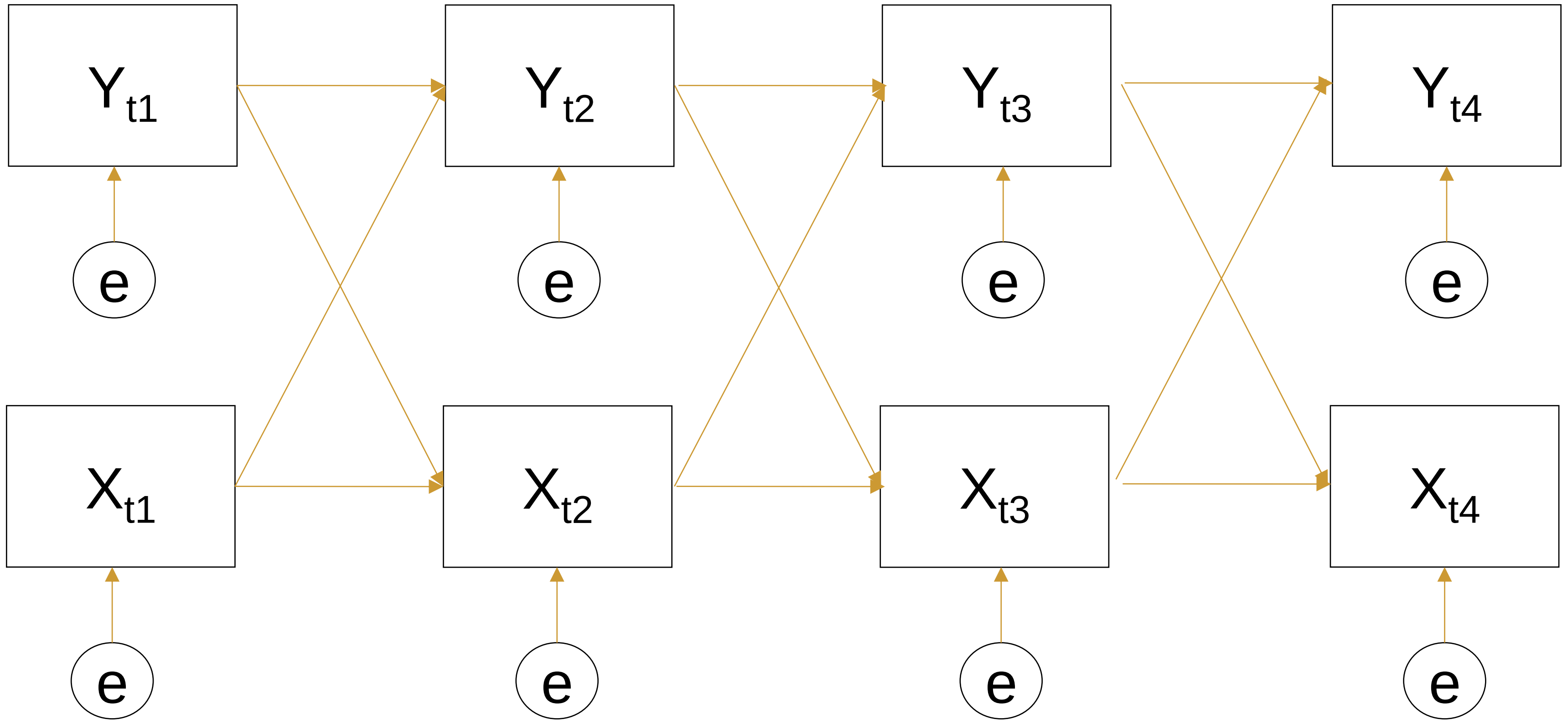
Reversible Change

- Will often be interested in (longitudinal) relation between two or more variables



Reversible Change

- This is called a VAR(1) model



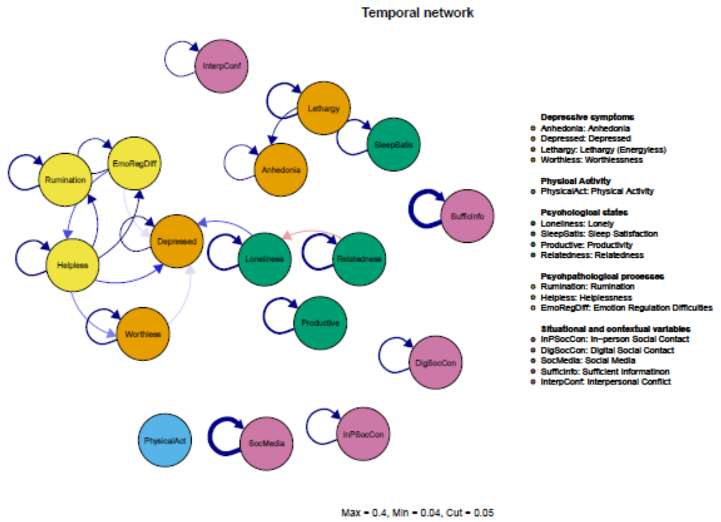
Reversible Change

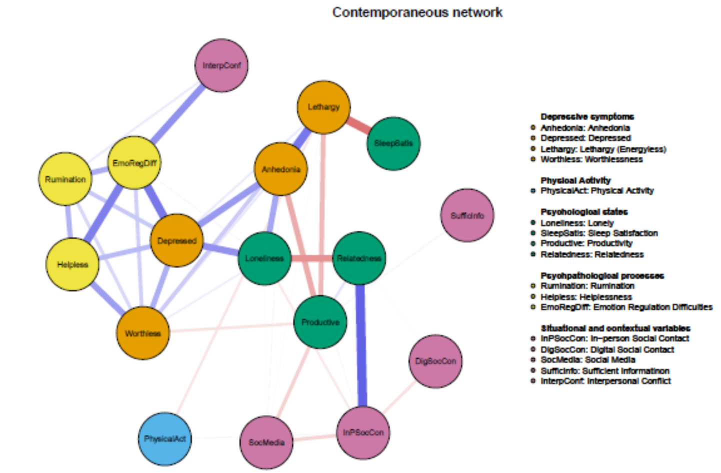
```
bform1 <-
  bf(mvbind(Y1, Y2) ~ Y1lag_c + Y2lag_c + (1|p|individual))


VAR <- brm(bform1, data = VARData, iter = 5000, chains = 2, cores = 2)
summary(VAR)
```

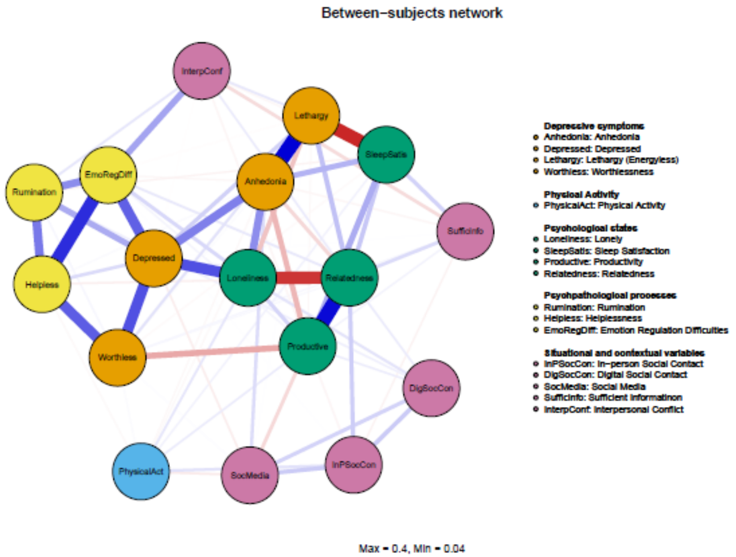
Reversible Change

```
> summary(VAR)
 Family: MV(gaussian, gaussian)
  Links: mu = identity; sigma = identity
         mu = identity; sigma = identity
Formula: Y1 ~ Y1lag_c + Y2lag_c + (1 | p | individual)
         Y2 ~ Y1lag_c + Y2lag_c + (1 | p | individual)
   Data: VARData (Number of observations: 4900)
  Draws: 2 chains, each with iter = 5000; warmup = 2500; thin = 1;
         total post-warmup draws = 5000

Multilevel Hyperparameters:
~individual (Number of levels: 100)
                              Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
sd(Y1_Intercept)                  0.56      0.04     0.48     0.65 1.00     1226     1983
sd(Y2_Intercept)                  0.54      0.04     0.47     0.63 1.00     1324     1883
cor(Y1_Intercept,Y2_Intercept)    0.37      0.09     0.18     0.54 1.00      892     1427

Regression Coefficients:
             Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
Y1_Intercept     3.96      0.06     3.85     4.08 1.00      792     1398
Y2_Intercept     4.36      0.06     4.24     4.47 1.00      841     1317
Y1_Y1lag_c       0.39      0.01     0.36     0.42 1.00     6420     4156
Y1_Y2lag_c       0.10      0.01     0.08     0.13 1.00     6869     3702
Y2_Y1lag_c       0.20      0.01     0.17     0.23 1.00     6128     4101
Y2_Y2lag_c       0.30      0.01     0.27     0.32 1.00     6384     3814

Further Distributional Parameters:
         Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
sigma_Y1     1.01      0.01     0.99     1.03 1.00     8868     4010
sigma_Y2     1.04      0.01     1.02     1.06 1.00     8544     3915

Residual Correlations:
              Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
rescor(Y1,Y2)     0.29      0.01     0.26     0.31 1.00     8533     3896

Draws were sampled using sampling(NUTS). For each parameter, Bulk_ESS
and Tail_ESS are effective sample size measures, and Rhat is the potential
scale reduction factor on split chains (at convergence, Rhat = 1).
```

Reversible Change







- Very close to network models
  - Main difference is the estimation method
- Networks:
  - Not truly multilevel
  - Inference more complicated
  - Scales better to large data

TILBURG UNIVERSITY

t e s c

Reversible Change

- Can also let residual variances be random across people
  - This implies that people exposure or sensitivity to unmodelled factors differ.
  - Also implies that reliability is person-specific.

- Finally, parameters can also be allowed to differ across time and not just individuals.