TILBURG ◆ UNIVERSITY

Multilevel Latent Markov Modeling in Continuous Time with an Application to the Analysis of Ambulatory Mood Assessment Data

Jeroen K. Vermunt

Department of Methodology and Statistics,
Tilburg University

http://members.home.nl/jeroenvermunt/

j.k.vermunt@uvt.nl

Social Statistics Workshop, Recent advances in the analysis of categorical and count data, LSE, 12 December 2013

Introduction

- Latent Markov modeling, also referred to hidden Markov, Markov switching, regime switching, latent transition modeling, etc …

- Becoming a quite popular tool for longitudinal data analysis. Various textbooks:
  - Bartolucci, Farcomeni & Pennoni
  - Collins & Lanza
  - Frühwirth-Schnatter
  - Zucchini, & MacDonald

- In this talk
  - illustrate recent contributions and ongoing research of myself and my group using a single rather complex but very modern application
  - show some developments in the Latent GOLD software version 5.0, which was released a few weeks ago

- Issues such as:
  - multiple ordinal observed responses, and measurement error
  - unobserved heterogeneity
  - (longitudinal) nesting structure
  - dealing with the unequal spacing of the measurements
  - model fit assessment
  - samples size, number of measurements, and power issues

- What I am presenting is not really new, it is the (easiness of) combination of separate pieces that is new

Application: Ambulatory Mood Assessment Study

- Described in Crayen et al. (2012), Psychosomatic Medicine

- Data collection using a mobile device (kind of mobile phone)
  - 164 students from the Free University of Berlin participated during 1 week: 8 assessments per day at random time intervals (m=100, sd=25)
  - 56 measurements in total per person (sometimes missing)
  - data collection method also referred to as “experience sampling”

- The pleasant-unpleasant mood dimension was assessed by the two ordered categorical items, unwell-well and bad-good (4 categories collapsed into 3)

- We are interested in
  - Mood regulation in time in a general/healthy population
  - Heterogeneity: individual differences in mood regulation

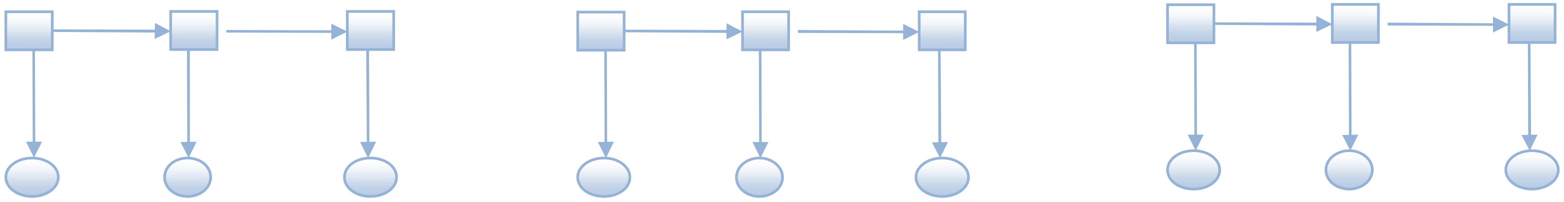
- Natural modeling choice:
  - Markov/transition model
  - measurement error: latent Markov model (LMM)
  - heterogeneity: mixed or mixture LMM

- Specific issue in this application: measurements within days are probably much more strongly associated than measurements between days

- Various modeling options
  - Latent Markov model with within and between day transition probabilities
  - Days nested within persons:
    - Static heterogeneity: two-level LMM
    - Dynamic heterogeneity I: two-level mixture LMM
    - Dynamic heterogeneity II: nested within- and between-day LMMs
    - Dynamic heterogeneity III: two-level nested LMMs

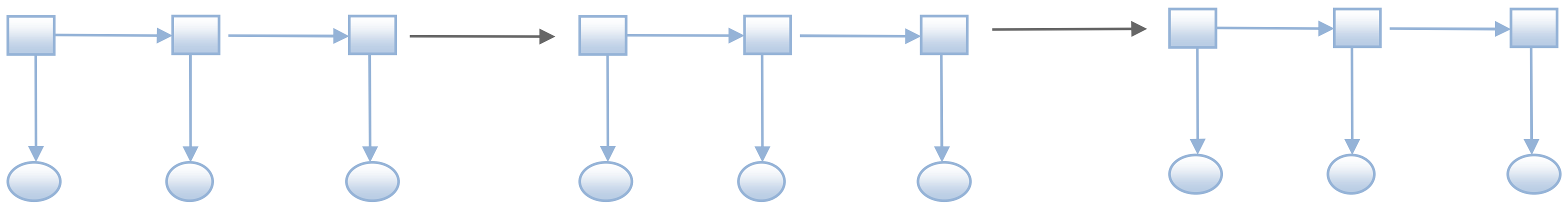
- Quite easy to specify with Latent GOLD Syntax language

- LMM assuming independence between days



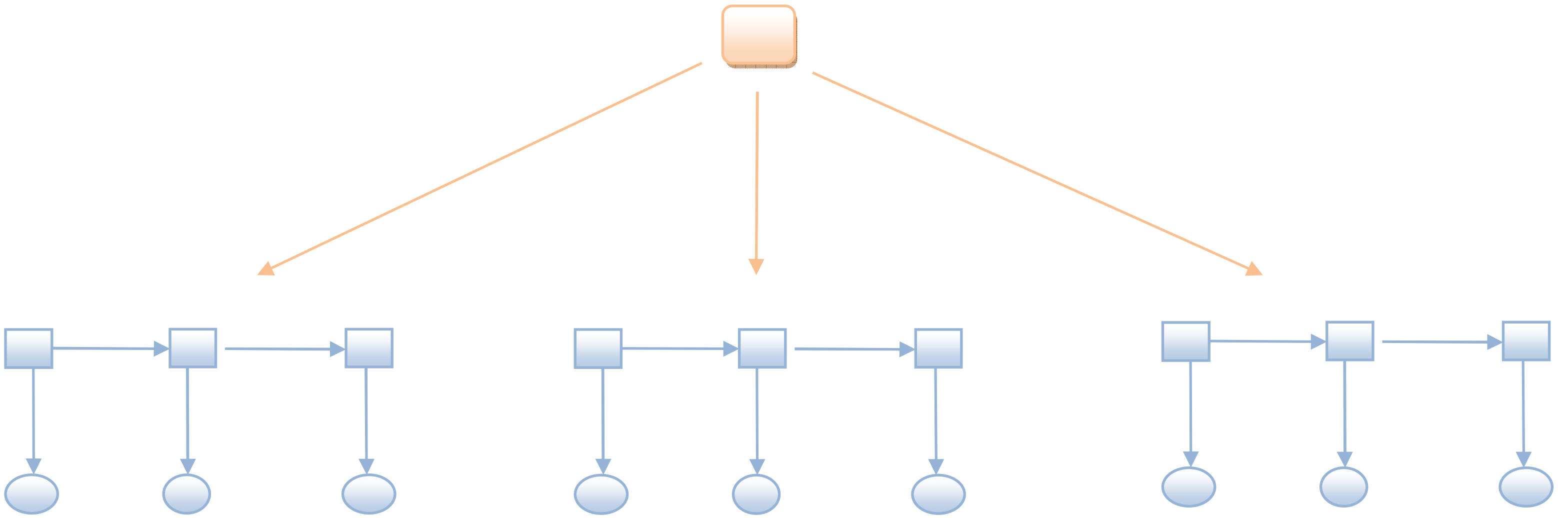
```
variables
    caseid DayID;
    dependent well, good;
    latent State nominal dynamic 3;
equations
    State[=0] <- 1;
    State <- 1 | State[-1];
    well <- 1 + State;
    good <- 1 + State;
```

- LMM with different transition probabilities between and within days



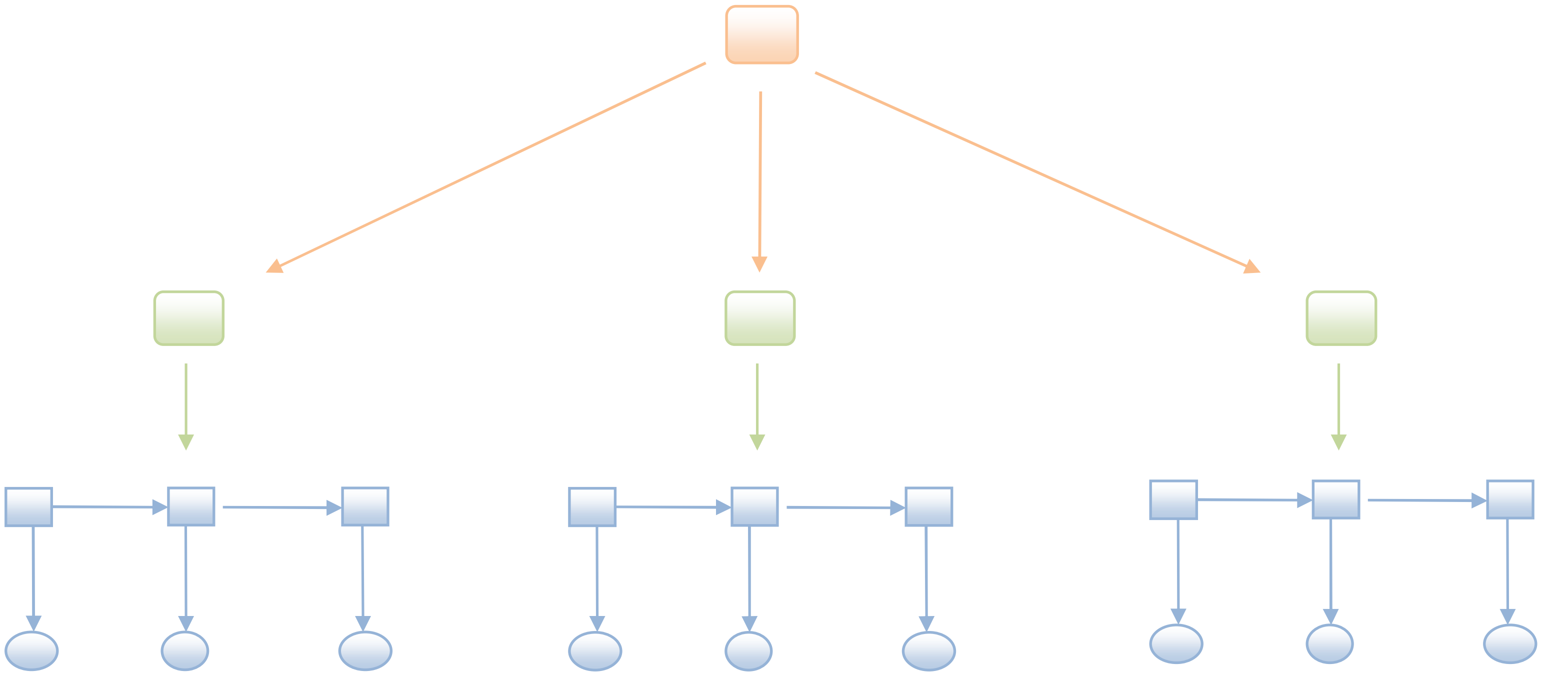
```
variables
    caseid UserID;
    dependent well, good;
    independent firsthour, laterhour;
    latent State nominal dynamic 3;
equations
    State[=0] <- 1;
    State <- firsthour | State[-1] + laterhour | State[-1];
    well <- 1 + State;
    good <- 1 + State;
```

- Static heterogeneity: two-level LMM



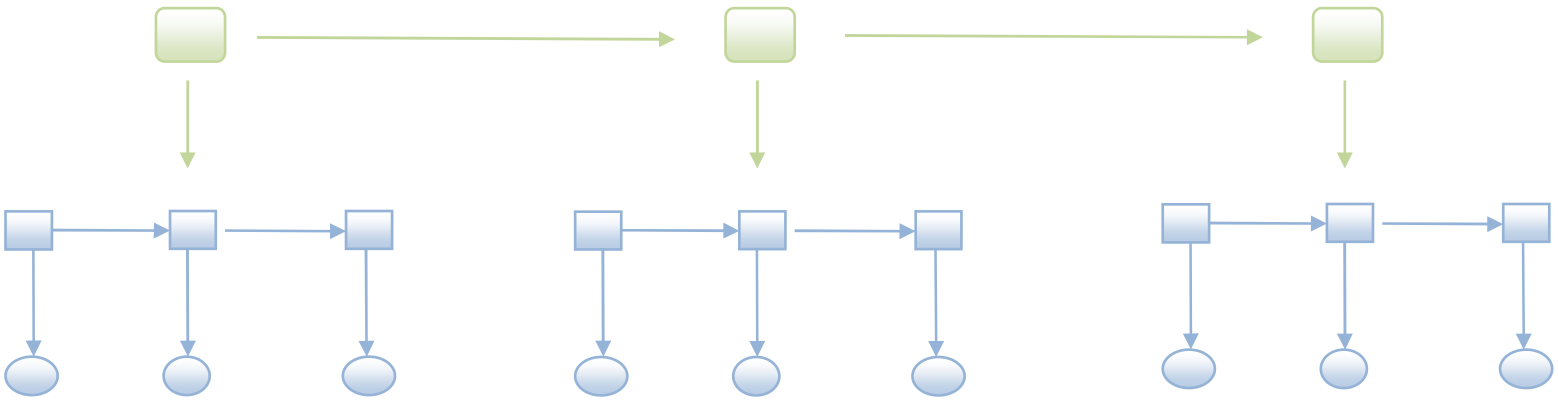
```
variables
   groupid UserId;
   caseid DayID;
   dependent well, good;
   latent Class nominal group 2, State nominal dynamic 3;
equations
   Class <- 1;
   State[=0] <- 1 | Class;
   State <- 1 | State[-1]  Class;
   well <- 1 + State;
   good <- 1 + State;
```

- Dynamic heterogeneity I: two-level mixture LMM



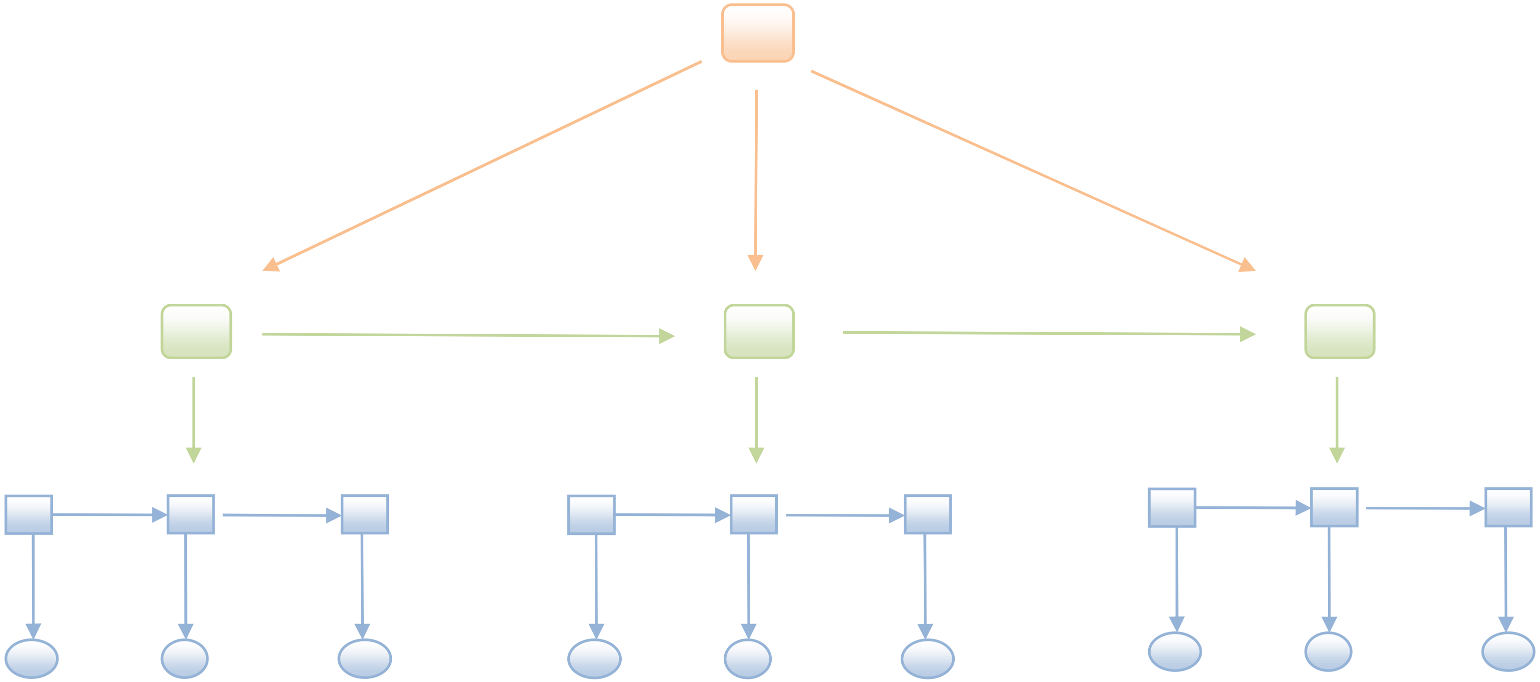
```
variables
    groupid UserId;
    caseid DayID;
    dependent well, good;
    latent Class nominal group 2, DClass nominal 2, State nominal dynamic 3;
equations
    Class <- 1;
    DClass <- 1 | Class;
    State[=0] <- 1 | DClass;
    State <- 1 | State[-1] DClass;
    well <- 1 + State;
    good <- 1 + State;
```

- Dynamic heterogeneity II: nested within-day and between-day LMMs



```
variables
    caseid UserID;
    dependent well, good;
    independent firsthour, laterhour;
    latent DClass nominal dynamic 2, State nominal dynamic 3;
equations
    DClass[=0] <- 1;
    DClass <- firsthour | DClass[-1] + (a~wei) laterhour | DClass[-1];
    State[=0] <- (b) 1 | DClass[=0];
    State <- (b) firsthour | DClass + laterhour | State[-1] DClass;
    well <- 1 + State;
    good <- 1 + State;
    a = {1 1 1 1 1 0 0 1};
```

- Dynamic heterogeneity III: two-level nested LMMs



```
variables
   caseid UserID;
   dependent well, good;
   independent firsthour, laterhour;
   latent Class nominal 2, DClass nominal dynamic 2, State nominal dynamic 3;
equations
   Class <- 1;
   DClass[=0] <- 1 | Class;
   DClass <- firsthour | DClass[-1] Class + (a~wei) laterhour | DClass[-1];
   State[=0] <- (b) 1 | DClass[=0];
   State <- (b) firsthour | DClass + laterhour | State[-1] DClass;
   well <- 1 + State;
   good <- 1 + State;
   a = {1 1 1 1 1 0 0 1};
```

Tilburg University

- Global fit measures

Model	LL	#par	BIC
Independence between days	-10469	16	21018
Between-day transitions	-10419	22	20948
Two-level	-10237	25	20600
Two-level mixture	-10232	27	20601
Nested models	-10223	27	20582
Two-level nested models	-10219	31	20595

- Local fit measures (bivariate residuals)

Null model

	well	good
good	6907.56	
Pairs	122.61	188.70
Lag1	61.28	76.79
Lag2	30.29	44.26

Final model

	well	Good
good	0.01	
Pairs	10.47	19.47
Lag1	0.07	0.02
Lag2	0.23	0.40

Continuous-Time Framework

- One thing we ignored so far is that measurements are unequally spaced
- In a discrete-time framework one can deal with this by defining a latent state per time point (per minute, per 15 minutes), and having missing values when no measurement is available
- More natural is to switch to a continuous time framework: Jackson et al. (2002, 2003) and Böckenholt (2005) proposed continuous-time latent Markov models
- Measurements are snapshots of a continuous time process. Measurements occur at irregular time points and the between information is missing
- Work on continuous-time (non-latent) Markov models is quite old (Cox & Miller, 1965; Coleman, 1964)

- As in discrete-time models, we model the transitions between time points.

  In fact, model structure, likelihood function, and estimation algorithm remains the same

- But:

  Transition probabilities are functions of transition intensities and the length of the time interval

  A log-linear model is defined for the transition intensities. The parameters are thus not logits of probabilities but log intensities

- With 3 latent states, the transition intensity matrix $\mathbf{Q}_i$ for person i equals

$$\mathbf{Q}_i = \begin{bmatrix} -q_{2|1,\mathbf{z}_i} - q_{3|1,\mathbf{z}_i} & q_{2|1,\mathbf{z}_i} & q_{3|1,\mathbf{z}_i} \\ q_{1|2,\mathbf{z}_i} & -q_{1|2,\mathbf{z}_i} - q_{3|2,\mathbf{z}_i} & q_{3|2,\mathbf{z}_i} \\ q_{1|3,\mathbf{z}_i} & q_{2|3,\mathbf{z}_i} & -q_{1|3,\mathbf{z}_i} - q_{2|3,\mathbf{z}_i} \end{bmatrix}$$

- The log-linear model for $q_{r|s,\mathbf{z}}$ equals

$$q_{r|s,\mathbf{z}_i} = \exp(\eta_{r|s\mathbf{z}_i}) = \exp(\gamma_{rs0} + \sum_{p=1}^{P} \gamma_{rsp} z_{pi}).$$

- The matrix $\mathbf{P}_{it}$ with the transition probabilities for time interval t are obtain as

$$\mathbf{P}_{it} = \mathbf{Exp}(\delta_{it}\mathbf{Q}_i),$$

where δ_{it} is the length of the time interval and $\mathbf{Exp}(\cdot)$ the matrix exponential.

- Software implementation requires (see, e.g., Kalbfleish and Lawless, JASA, 1985)

  1. computation of the transition probabilities from the intensities

     $\mathbf{P} = \mathbf{A}\ \mathrm{diag}(\exp(\mathbf{v}\ \delta))\ \mathbf{A}^{-1}$

     $\mathbf{A}$ and $\mathbf{v}$ are eigenvectors and eigenvalues of $\mathbf{Q}$

  2. computation of first derivatives of **P** towards the log-intensity parameters

     $\partial\mathbf{P}/\partial\gamma = \mathbf{A}\ \mathbf{W}\ \mathbf{A}^{-1}$
     $w_{rs} = b_{r|s}\ [\exp(v_r\ \delta) - \exp(v_s\ \delta)\ ]\ /\ [v_r - v_s]$ for $r \neq s$
     $w_{rr} = b_{r|r}\ \delta\ \exp(v_r\ \delta)$
     $\mathbf{B} = \mathbf{A}^{-1}\ (\partial\mathbf{Q}/\partial\gamma)\ \mathbf{A}$

  3. definition of time intervals, fine tuning starting values, specific output (e.g. transition probabilities for specific time interval, sojourn times, etc.)

- Estimation of the various continuous-time LMMs with the ambulatory assessment data

		Discrete Time		Continuous Time	
Model	#par	LL	BIC	LL	BIC
Independence between days	16	-10469	21018	-10477	21035
Between-day transitions	22	-10419	20948	-10431	20973
Two-level	25	-10237	20600	-10244	20615
Two-level mixture	27	-10232	20601	-10238	20613
Nested models	27	-10223	20582	-10229	20595
Two-level nested models	31	-10219	20595	-10226	20607

Continuous time model fit slightly worse, but same heterogeneity model is selected

- Measurement probabilities (Left=DT, Right=CT)

well (Left=DT)

State	not	quite	very
1	0.89	0.11	0.00
2	0.05	0.91	0.04
3	0.00	0.23	0.77

well (Right=CT)

State	not	quite	very
1	0.89	0.11	0.00
2	0.05	0.91	0.04
3	0.00	0.23	0.77

good (Left=DT)

State	not	quite	very
1	0.91	0.09	0.00
2	0.02	0.95	0.03
3	0.00	0.17	0.83

good (Right=CT)

State	not	quite	very
1	0.92	0.08	0.00
2	0.02	0.95	0.03
3	0.00	0.18	0.82

- Initial and transition probabilities for States (Left=DT, Right=CT/100 minutes)

	State[=0]		
DClass[=0]	1	2	3
1	0.23	0.76	0.01
2	0.13	0.62	0.25

	State[=0]		
DClass[=0]	1	2	3
1	0.22	0.78	0.01
2	0.10	0.77	0.13

		State		
Dclass	State[-1]	1	2	3
1	1	0.53	0.45	0.02
1	2	0.09	0.86	0.06
1	3	0.05	0.55	0.40
2	1	0.28	0.49	0.22
2	2	0.08	0.58	0.34
2	3	0.03	0.23	0.74

		State		
DClass	State[-1]	1	2	3
1	1	0.53	0.45	0.02
1	2	0.09	0.86	0.06
1	3	0.06	0.55	0.40
2	1	0.28	0.49	0.23
2	2	0.08	0.57	0.35
2	3	0.03	0.23	0.74

Class 1: mood regulates towards State 2; Class 2: tends to move go to State 3

- Marginal State probabilities given DayClass (DT model)

DayClass=1

	States		
Time	1	2	3
0	0.23	0.76	0.01
1	0.19	0.76	0.05
2	0.17	0.76	0.07
3	0.16	0.77	0.07
4	0.15	0.77	0.08
5	0.15	0.77	0.08
6	0.15	0.77	0.08
7	0.15	0.77	0.08
8	0.15	0.77	0.08

DayClass=2

	States		
Time	1	2	3
0	0.13	0.62	0.25
1	0.10	0.48	0.43
2	0.08	0.42	0.50
3	0.07	0.39	0.53
4	0.07	0.38	0.55
5	0.07	0.38	0.55
6	0.07	0.38	0.56
7	0.07	0.38	0.56
8	0.07	0.38	0.56

- Initial and transition probabilities for Day Classes (Left=DT, Right=CT)

DClass[=0]

1	2
0.67	0.33

DClass[=0]

1	2
0.67	0.33

	DClass	
DClass[-1]	1	2
1	0.98	0.02
2	0.09	0.91

	DClass	
DClass[-1]	1	2
1	0.98	0.02
2	0.08	0.92

Work in progress

- What I presented today
- Local fit measures that look at the most relevant aspects in the data that the model is assumed to pick up
- Sample size requirements, power computations
- Parameter recovery with small samples
- Three-step latent Markov modeling

TILBURG UNIVERSITY

Thank you for your attention!