Examining contingent discrete change over time with associative latent transition analysis

Brian P. Flaherty
Department of Psychology
University of Washington
bxf4@u.washington.edu

May 15, 2006

Outline

- Introduction: What sort of questions are we interested in?
- Latent class model
- Latent transition or latent Markov models
- Associative latent transition model
- Empirical illustration
- Conclusions/discussion

Substantive questions?

Phenomena and questions of interest:

- ☐ Focus on unobservable groups or categories in the population.
- ☐ How does change among discrete states proceed over time?
- ☐ How well are states measured?
- ☐ Is change comparable between two groups (e.g., treatment and control)?
- ☐ How are two changing categorical variables associated?

Latent class models

Latent class model

$$P(\boldsymbol{W}) = \sum_{c=1}^{C} \gamma_c \left(\prod_{j=1}^{q} \prod_{k=1}^{r_j} \rho_{jk|c}^{I(w_j=k)} \right)$$

where

- W is a response pattern;
- γ_c is the proportion of the population in latent class c;
- $\rho_{jk|c}^{I(w_j=k}$ is the probability of response k to item j for latent class c.
- $I(\cdot)$ is the indicator function, here used to select the appropriate response probabilities.

I will abbreviate latent class as LC.

Estimation

- EM algorithm
- identification problems/multiple modes of likelihood
- use many sets of start values to evaluate the problem
- parameter restrictions to simplify the model

Fit and model selection

- Typically use G^2 or Pearson's X^2.
- But, fit assessment is difficult due to sparse tables.
  - Sparse means a low n relative to the number of cells in the observed data contingency table.
  - Sparseness reduces the expectation of the test statistic.
- Typical practice with sparse data:
  - Ideally G^2 lower than df
  - Interpretable solution
- Model comparisons (i.e., nested comparisons) perform better than absolute fit tests with sparse data

LC results

	Probability of a Yes response		
LC (proportion)	Ever try	Past 30 day	Ever regular
LC 1 (.45)	0.0	0.01	0.03[1]
LC 2 (.38)	1.0[2]	0.28	0.03[1]
LC 3 (.17)	1.0[2]	0.99	0.97

Latent transition model

Latent transition model

$$P(\boldsymbol{W}) = \sum_{s_1=1}^{S_1} \cdots \sum_{s_T=1}^{S_T} \delta_{s_1} \left(\prod_{j=1}^{q} \prod_{k=1}^{r_j} \rho_{1jk|s_1}^{I(w_{1j}=k)} \right) \times \prod_{t=2}^{T} \left(\tau_{s_t|s_{t-1}}^{(t-1)} \prod_{j=1}^{q} \prod_{k=1}^{r_j} \rho_{tjk|s_t}^{I(w_{tj}=k)} \right).$$

where

- δ_{s_1} is the proportion of the population in LC s_1 at Time 1;
- $\tau_{s_t|s_{t-1}}^{(t-1)}$ is the transition probability from LC s_{t-1} at Time $t-1$ to LC s_t at Time t.

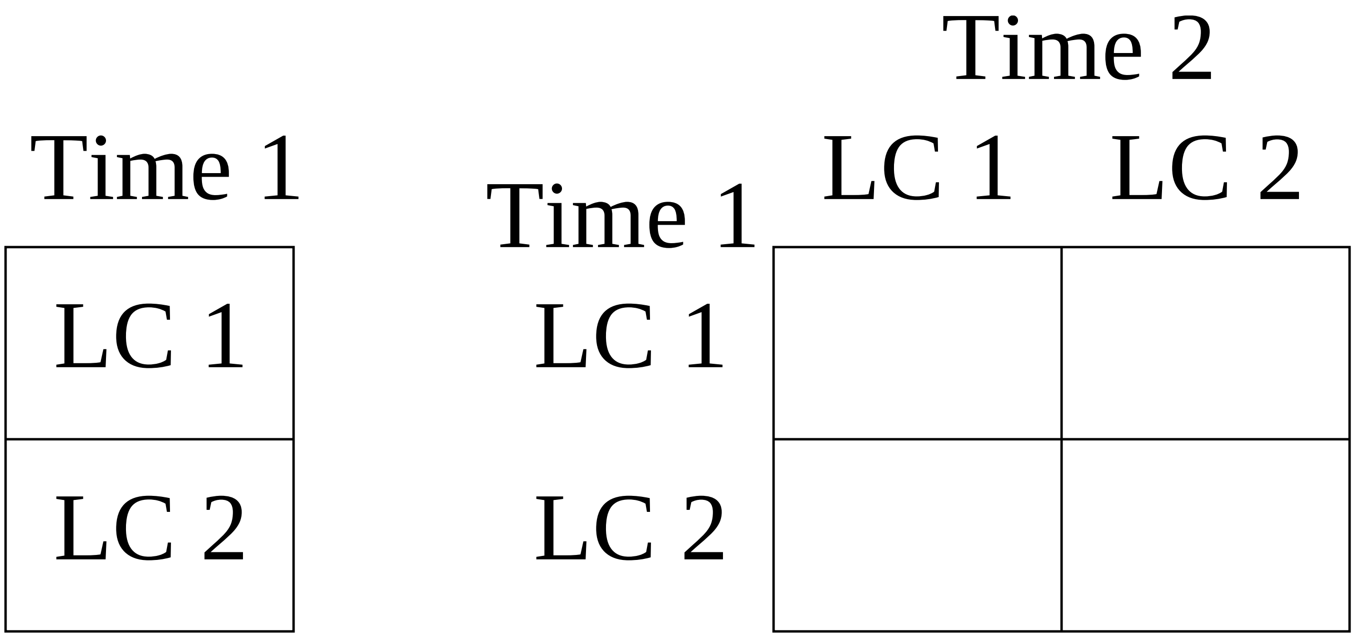
Diagram of LC approach for two times

Interested in two latent classes at two time points.
Could fit a four class LC model (2 classes $\times$ 2 times):

LC 1 LC 1	LC 1 LC 2	LC 2 LC 1	LC 2 LC 2

Diagram of LTA for two times

Instead, one can model the starting point and then change between time points.



Transition probabilities are row conditional.

Same number of estimates in these two approaches, hence reparameterization.

Associative latent transition model

$$P(\boldsymbol{W}) = \sum_{c_1=1}^{C_1} \sum_{d_1=1}^{D_1} \cdots \sum_{c_T=1}^{C_T} \sum_{d_T=1}^{D_T} \alpha_{c_1} \beta_{d_1|c_1} \left(\prod_{j=1}^{q} \prod_{k=1}^{r_j} \rho_{1jk|c_1 d_1}^{I(w_{1j}=k)} \right) \times \prod_{t=2}^{T} \left(\epsilon_{c_t|c_{t-1} d_{t-1}}^{(t-1)} \eta_{d_t|c_{t-1} c_t d_{t-1}}^{(t-1)} \prod_{j=1}^{q} \prod_{k=1}^{r_j} \rho_{tjk|c_t d_t}^{I(w_{tj}=k)} \right)$$

where

- α_{c_1} is the probability of membership in X LC c_1 at Time 1;
- $\beta_{d_1|c_1}$ is the probability of membership in Y LC d_1 given membership in X LC c_1 at Time 1;

$$P(\boldsymbol{W}) = \sum_{c_1=1}^{C_1} \sum_{d_1=1}^{D_1} \cdots \sum_{c_T=1}^{C_T} \sum_{d_T=1}^{D_T} \alpha_{c_1} \beta_{d_1|c_1} \left(\prod_{j=1}^{q} \prod_{k=1}^{r_j} \rho_{1jk|c_1 d_1}^{I(w_{1j}=k)} \right) \times$$
$$\prod_{t=2}^{T} \left(\epsilon_{c_t|c_{t-1}d_{t-1}}^{(t-1)} \eta_{d_t|c_{t-1}c_t d_{t-1}}^{(t-1)} \prod_{j=1}^{q} \prod_{k=1}^{r_j} \rho_{tjk|c_t d_t}^{I(w_{tj}=k)} \right)$$

where

- $\epsilon_{c_t|c_{t-1}d_{t-1}}^{(t-1)}$ is the probability of X LC membership c_t at Time t conditional on membership in both X LC c_{t-1} and Y LC d_{t-1} at Time $t-1$;

$$P(\boldsymbol{W}) = \sum_{c_1=1}^{C_1} \sum_{d_1=1}^{D_1} \cdots \sum_{c_T=1}^{C_T} \sum_{d_T=1}^{D_T} \alpha_{c_1} \beta_{d_1|c_1} \left(\prod_{j=1}^{q} \prod_{k=1}^{r_j} \rho_{1jk|c_1 d_1}^{I(w_{1j}=k)} \right) \times$$
$$\prod_{t=2}^{T} \left(\epsilon_{c_t|c_{t-1}d_{t-1}}^{(t-1)} \eta_{d_t|c_{t-1}c_t d_{t-1}}^{(t-1)} \prod_{j=1}^{q} \prod_{k=1}^{r_j} \rho_{tjk|c_t d_t}^{I(w_{tj}=k)} \right)$$

where

- $\eta_{d_t|c_{t-1}c_t d_{t-1}}^{(t-1)}$ is the probability of Y LC membership d_t at Time t conditional on membership in X LC's c_{t-1} and c_t at Times $t-1$ and t, respectively, and Y LC membership d_{t-1} at Time $t-1$.

Empirical illustration - Methods

Data

- National Longitudinal Study of Adolescent Health (Add Health);
- U.S. students in grades 7 through 12;
- 6,504 initially surveyed in 1994;
- 4,834 re-interviewed in 1995.

Analyses use data from all 6,504 respondents.

Three Yes/No items measuring recent tobacco use:

- ☐ Have you ever tried cigarette smoking, even just 1 or 2 puffs?
- ☐ Have you smoked cigarettes on any of the last 30 days?
- ☐ Have you ever smoked cigarettes regularly (at least 1 cigarette every day for 30 days)?

Three Yes/No items measuring recent alcohol use:

- ☐ Have you had a drink of beer, wine, or liquor—not just a sip or a taste of someone else's drink—more than 2 or 3 times in your life?
- ☐ Have you drank alcohol in the past 12 months?
- ☐ Have you ever been drunk or drank 5 or more drinks in a row in the past 12 months?

Analyses

- Fit single LTA models first:
  - Choose a number of LC's;
  - Examine measurement structure.
- Fit multiple sets of start values to base ALTA model to assess identification (i.e., multi-modality).
- Fit ALTA model corresponding to various patterns of association between tobacco and alcohol classes.

Empirical illustration - Results

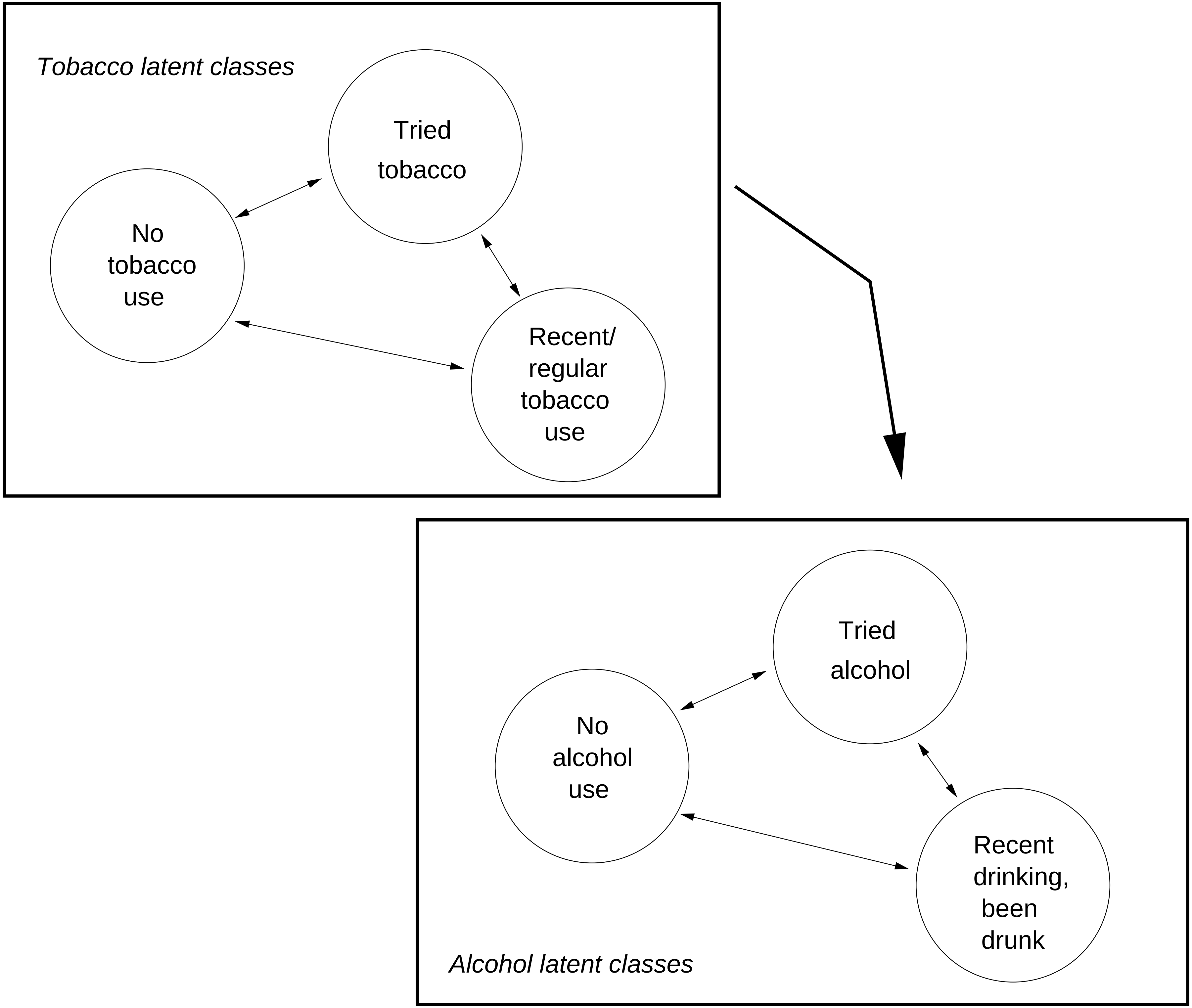
Measurement estimates

Tobacco response probabilities

	Probability of a Yes response		
LC (proportion)	Ever try	Past 30 day	Ever regular
LC 1	0.0	0.01	0.03[1]
LC 2	1.0[2]	0.28	0.03[1]
LC 3	1.0[2]	0.99	0.97

Alcohol response probabilities

	Probability of a Yes response		
LC (proportion)		Past 12 month	Drunk/Binge
LC 1	0.0	0.00	0.0[3]
LC 2	1.0[4]	0.67	0.0[3]
LC 3	1.0[4]	1.00	1.0



Tobacco latent classes
Tried tobacco
No tobacco use
Recent/ regular tobacco use
Tried alcohol
No alcohol use
Recent drinking, been drunk
Alcohol latent classes

Identification assessment

The model of tobacco and alcohol:

- 88 parameter estimates
- 8 are conditional response probabilities (using the same parameter restrictions as earlier)

Ran 1000 sets of random start values.

Fourteen different G^2 values (to 1 decimal place) and 2 NaN values (not a number, e.g., $0/0$).

G^2 value	Number of runs
1335.8	356
5538.2	77
5538.3	67
5737.1	86
6412.4	84
6491.4	73
9019.5	101
9845.0	39
10586.7	27
10600.8	3
10600.9	25
13598.1	15
14189.0	23
14433.1	22

Selected ALTA estimates

Unconditional tobacco class membership (α's)

	No tobacco	Try tobacco	Regular tobacco
Time 1	0.45	0.38	0.17
Time 2	0.43	0.24	0.33

Alcohol class conditional on tobacco class (β's)

Time 1	Time 1 Alcohol class		
Tobacco LC	No alc	Try alcohol	Drunk
No tobacco	0.71	0.19	0.10
Try tobacco	0.30	0.33	0.37
Reg tobacco	0.11	0.16	0.73

Time 2 X's

Where are people likely to be among tobacco classes at Time 2 if they hadn't ever smoked at Time 1? (Selected ϵ's)

Time 1	Time 2 Tobacco class		
Alcohol LC	No tobacco	Try tobacco	Regular tobacco
No alcohol	0.85	0.13	0.02
Try alcohol	0.78	0.20	0.02
Drunk	0.67	0.24	0.09

Time 2 Y's

What happens to people starting in No tobacco and No alcohol? (Selected η's)

	Time 2 LC		
Time 2 Tobacco class	No alcohol	Try alcohol	Drunk
No tobacco	0.87	0.09	0.04
Try tobacco	0.65	0.15	0.20
Reg tobacco	0.44	0.06	0.50

Among those who had not tried alcohol at Time 1, which combinations of tobacco use classes are associated with the highest probability of being in the Drunk class at Time 2?

Tobacco classes	P(Drunk)
No tob, No tob	0.04
No tob, Try tob	0.20
No tob, Reg tob	0.50
Try tob, No tob	0.08
Try tob, Try tob	0.23
Try tob, Reg tob	0.55
Reg tob, No tob	0.11
Reg tob, Try tob	0.19
Reg tob, Reg tob	0.41

Constrained T2 P(drunk)

Constrained model G^2 is 1347.9 with 4013 degrees of freedom.

G^2 difference is 12.1 with 6 degrees of freedom (4013-4007).

The p-value of this nested model comparison is 0.0598.

The constrained $\hat{\eta}$'s are:

Time 2 Tobacco class	P(Drunk)
No tob	0.05
Try tob	0.21
Reg tob	0.46

Moving to less use

What about transitions out of the Drunk/binge LC?

Do the rates that people change from Drunk/binge to No alcohol depend on a reducing to No tobacco use?

Unconstrained estimates:

P(Time 2 No alc | Time 1 Drunk, Try tob $\rightarrow$ No tob)	0.26
P(Time 2 No alc | Time 1 Drunk, Reg tob $\rightarrow$ No tob)	0.45

Constrained estimate is 0.29.

Constrained model $G^2 = 1343.36$.

$\Delta G^2 = 7.55$ on 1 degree of freedom, p $= 0.006$.

Conclusions

Conclusions

- Parameterization to examine contingent relations between two dynamic latent class variables.
- The goal of the parameterization is to make contingent change explicit.
- Can flexibly test simple and complex hypotheses with parameter restrictions.

Limitations

- Big models, a lot of parameters;
- Shares difficulties in assessing fit with other large latent class models;
- Current implementation doesn't provide standard errors.

Not yet clear what the bounds are of this model:

- How many times?
- How many classes?
- How many items?