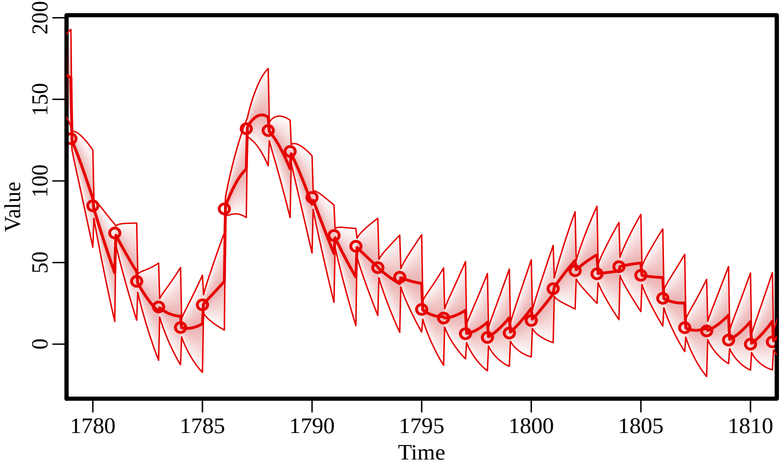
Hierarchical Bayesian Continuous Time Dynamic Models

September 2018

Charles Driver

Max Planck Institute for Human Development

- Model in which the present state depends in some way on the previous state, and an error / uncertainty component



- Theories implicitly (or explicitly) dynamic.

- Theories implicitly (or explicitly) dynamic.
- Mapping fuzzy theory to dynamic system highlights gaps, helps ensure coherent, testable, incrementally improvable theory.

- Pre register

- Pre register
- Test

- Pre register
- Test

- Pre register
- Test
- Replicate

- Pre register
- Test
- Replicate
- Theory supported

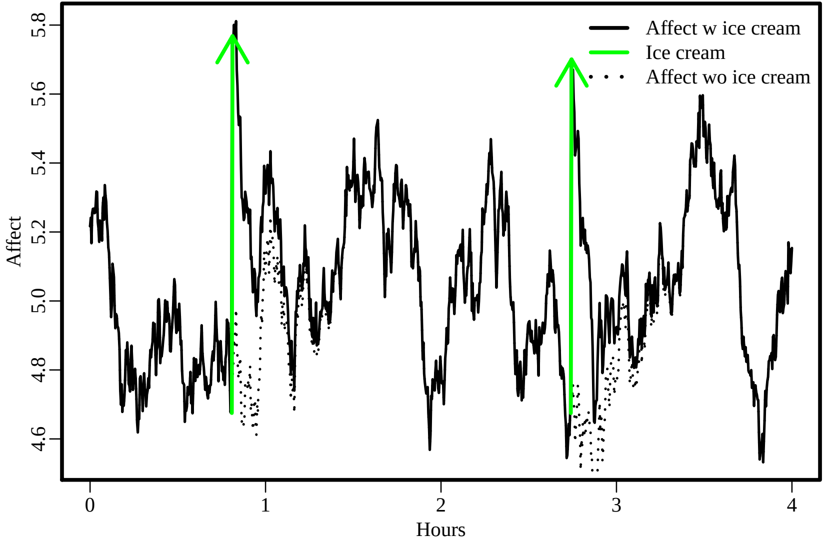
What went wrong?

- Fuzzy theory – very vague predictions

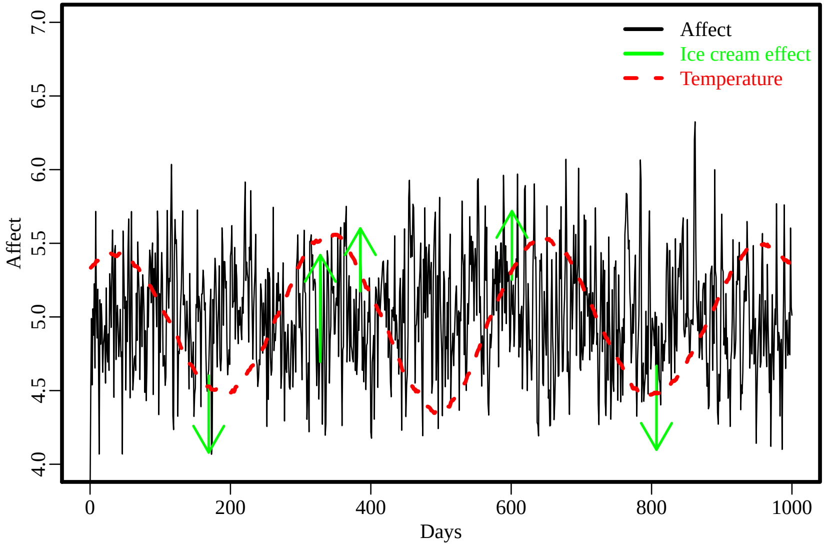
- Fuzzy theory – very vague predictions
- Tenuous and highly specific measurement of theory

- Fuzzy theory – very vague predictions
- Tenuous and highly specific measurement of theory
- Generalizability non-existent

- Fuzzy theory – very vague predictions
- Tenuous and highly specific measurement of theory
- Generalizability non-existent
- See Tal Yarkoni, Paul Meehl, no doubt others – long running issues!



Affect w ice cream
Ice cream
Affect wo ice cream
Affect
5.8
5.6
5.4
5.2
5.0
4.8
4.6
Hours
0
1
2
3
4


Affect
Ice cream effect
Temperature
7.0
6.5
6.0
5.5
5.0
4.5
4.0
Affect
0
200
400
600
800
1000
Days

Summary of why:

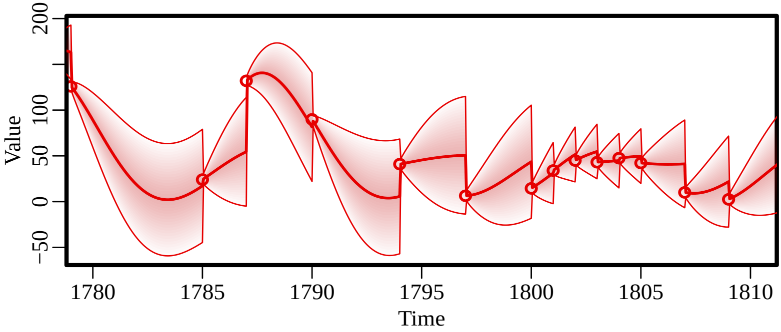
- Match practice with theory.

- Match practice with theory.
- Easily see missing elements of theory / knowledge.

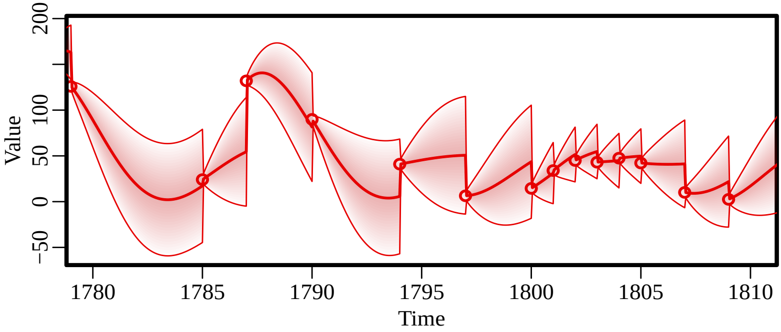
- Match practice with theory.
- Easily see missing elements of theory / knowledge.
- Reduce hand waving and vagueness – away from simple nulls towards numerically grounded theory.

- Match practice with theory.
- Easily see missing elements of theory / knowledge.
- Reduce hand waving and vagueness – away from simple nulls towards numerically grounded theory.
- I'm convinced! How to do it?

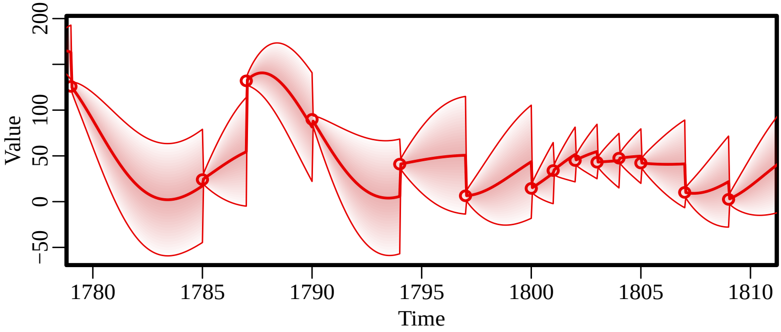
- Typical,discrete time approaches to dynamic modelling (e.g. latent change, autoregressive) directly model the dependence of later states on earlier states. $\eta_t = A\eta_{t-1} + b + \zeta(t)$



- Typical,discrete time approaches to dynamic modelling (e.g. latent change, autoregressive) directly model the dependence of later states on earlier states. $\eta_t = A\eta_{t-1} + b + \zeta(t)$
- Continuous time (differential equation) approaches model the *direction of change* given the current state. $\mathrm{d}\eta(t) = A\eta(t) + b + \zeta(t)$



- Typical,discrete time approaches to dynamic modelling (e.g. latent change, autoregressive) directly model the dependence of later states on earlier states. $\eta_t = A\eta_{t-1} + b + \zeta(t)$
- Continuous time (differential equation) approaches model the *direction of change* given the current state. $\mathrm{d}\eta(t) = A\eta(t) + b + \zeta(t)$
- Expectations regarding future states are then a deterministic function of underlying continuous time parameters and the time interval between measurements.
  $\eta_t = A^*\eta_{t-1} + b^* + \zeta^*(t)$



- More accuracy by incorporating time explicitly.

- More accuracy by incorporating time explicitly.
- More theoretically justifiable - things happen between measurement occasions.

- More accuracy by incorporating time explicitly.
- More theoretically justifiable - things happen between measurement occasions.
- Causality and explicit connection to theory possible / easier – but can be explorative also.

- More accuracy by incorporating time explicitly.
- More theoretically justifiable - things happen between measurement occasions.
- Causality and explicit connection to theory possible / easier – but can be explorative also.
- Tangible benefits:

- More accuracy by incorporating time explicitly.
- More theoretically justifiable - things happen between measurement occasions.
- Causality and explicit connection to theory possible / easier – but can be explorative also.
- Tangible benefits:
  - More flexible data collection - no longer a need for consistent measurement intervals within or between individuals.

- More accuracy by incorporating time explicitly.
- More theoretically justifiable - things happen between measurement occasions.
- Causality and explicit connection to theory possible / easier – but can be explorative also.
- Tangible benefits:
    - More flexible data collection - no longer a need for consistent measurement intervals within or between individuals.
    - Consistent parameter comparisons between different data collection schedules - between waves, studies, or both.

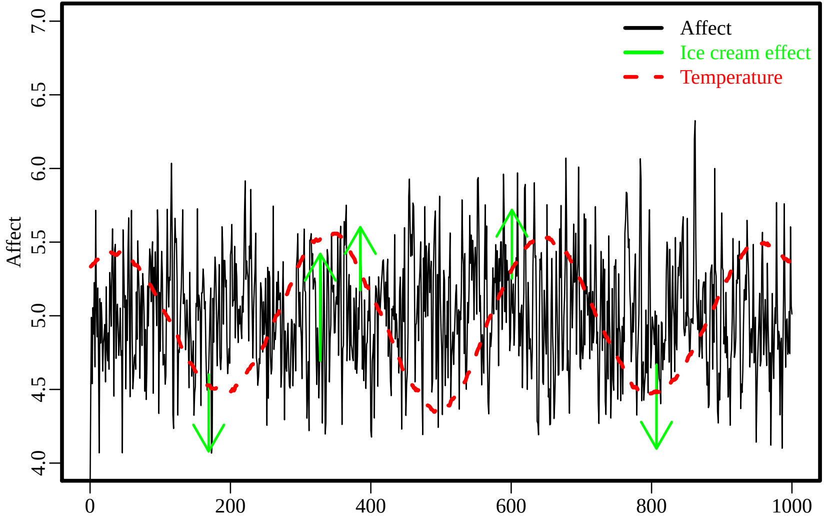
- More accuracy by incorporating time explicitly.
- More theoretically justifiable - things happen between measurement occasions.
- Causality and explicit connection to theory possible / easier – but can be explorative also.
- Tangible benefits:
  - More flexible data collection - no longer a need for consistent measurement intervals within or between individuals.
  - Consistent parameter comparisons between different data collection schedules - between waves, studies, or both.
  - Differential equations allow parsimonious specification of higher order dynamics such as oscillations and slow changing trends.

- More accuracy by incorporating time explicitly.
- More theoretically justifiable - things happen between measurement occasions.
- Causality and explicit connection to theory possible / easier – but can be explorative also.
- Tangible benefits:
  - More flexible data collection - no longer a need for consistent measurement intervals within or between individuals.
  - Consistent parameter comparisons between different data collection schedules - between waves, studies, or both.
  - Differential equations allow parsimonious specification of higher order dynamics such as oscillations and slow changing trends.
- Downsides:

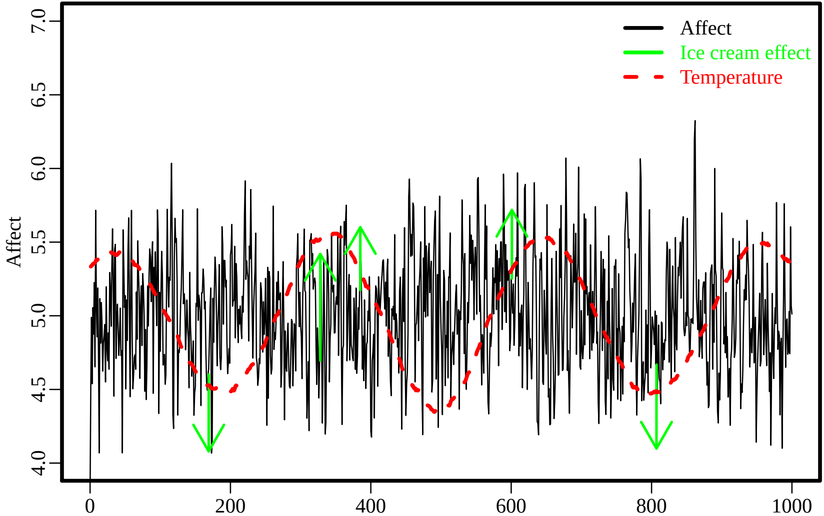
- More accuracy by incorporating time explicitly.
- More theoretically justifiable - things happen between measurement occasions.
- Causality and explicit connection to theory possible / easier – but can be explorative also.
- Tangible benefits:
  - More flexible data collection - no longer a need for consistent measurement intervals within or between individuals.
  - Consistent parameter comparisons between different data collection schedules - between waves, studies, or both.
  - Differential equations allow parsimonious specification of higher order dynamics such as oscillations and slow changing trends.
- Downsides:
  - More mathematically demanding than autoregressive / latent change approaches, more computationally demanding in general.

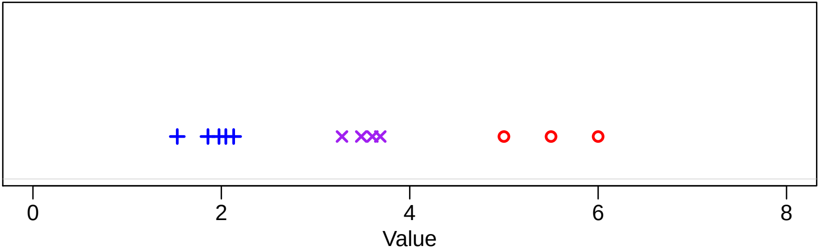
- They typically exist, ignoring them can be very problematic.

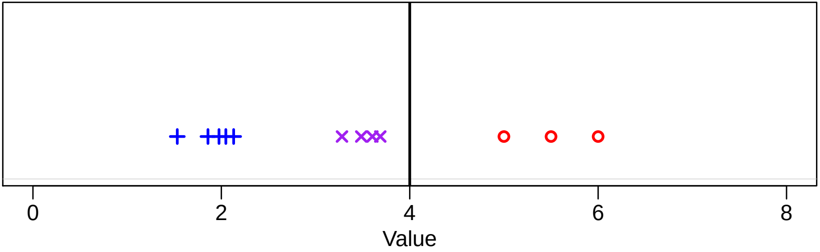
- They typically exist, ignoring them can be very problematic.
- They can vary depending on time scale.



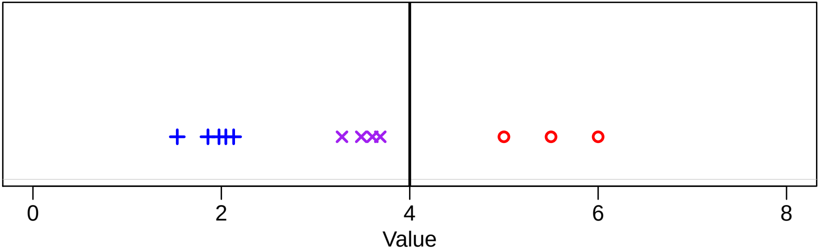
- They typically exist, ignoring them can be very problematic.
- They can vary depending on time scale.
- They make estimation trickier.



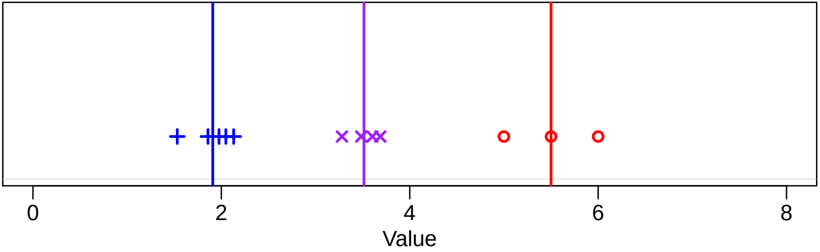

0
2
4
6
8
Value



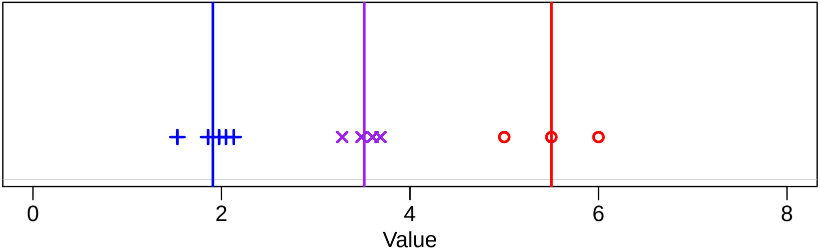
- Complete pooling - estimate single fixed effect parameter for entire sample.



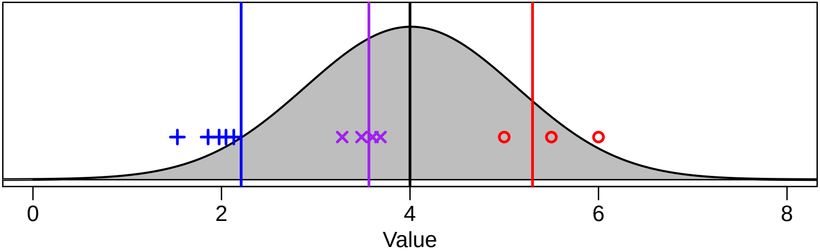
- Complete pooling - estimate single fixed effect parameter for entire sample.
    - Contaminates within subject model with between subjects differences, and not a great way to understand individual differences...



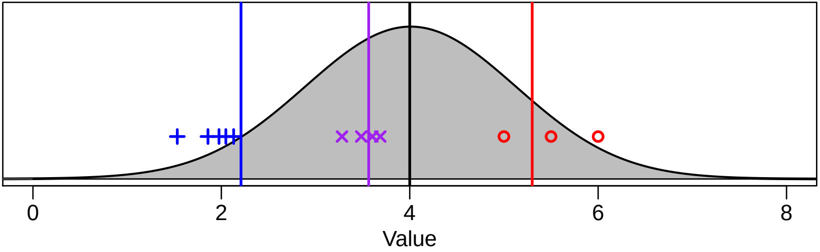
- Complete pooling - estimate single fixed effect parameter for entire sample.
    - Contaminates within subject model with between subjects differences, and not a great way to understand individual differences...
- No pooling - assume individuals have no similarities, estimate seperate model for each.



- Complete pooling - estimate single fixed effect parameter for entire sample.
  - Contaminates within subject model with between subjects differences, and not a great way to understand individual differences...
- No pooling - assume individuals have no similarities, estimate seperate model for each.
  - Simple and perfect if sufficient data exists – 'sufficient' may be extremely large – otherwise prone to finite sample biases and high variance.



- Complete pooling - estimate single fixed effect parameter for entire sample.
  - Contaminates within subject model with between subjects differences, and not a great way to understand individual differences...
- No pooling - assume individuals have no similarities, estimate seperate model for each.
  - Simple and perfect if sufficient data exists – 'sufficient' may be extremely large – otherwise prone to finite sample biases and high variance.
- Partial pooling - estimate population distribution for individual models.



- Complete pooling - estimate single fixed effect parameter for entire sample.
    - Contaminates within subject model with between subjects differences, and not a great way to understand individual differences...
- No pooling - assume individuals have no similarities, estimate seperate model for each.
    - Simple and perfect if sufficient data exists – 'sufficient' may be extremely large – otherwise prone to finite sample biases and high variance.
- Partial pooling - estimate population distribution for individual models.
    - More complex models but most flexible - parameters are not either 'freely varying' or 'not varying at all' but the extent of allowed variation is estimated.

$$p(\boldsymbol{\Phi}, \boldsymbol{\mu}, \mathbf{R}, \beta | \mathbf{Y}, \mathbf{Z}) = \frac{p(\mathbf{Y}|\boldsymbol{\Phi}) p(\boldsymbol{\Phi}|\boldsymbol{\mu}, \mathbf{R}, \boldsymbol{\beta}, \mathbf{Z}) p(\boldsymbol{\mu}, \mathbf{R}, \beta)}{p(\mathbf{Y})} \tag{1}$$

Where subject specific parameters $\boldsymbol{\Phi}_i$ are determined in the following manner:

$$\boldsymbol{\Phi}_i = \text{tform}\left(\boldsymbol{\mu} + \mathbf{R}\mathbf{h}_i + \boldsymbol{\beta}\mathbf{z}_i\right) \tag{2}$$

$$\mathbf{h}_i \sim \mathrm{N}(\mathbf{0}, \mathbf{1}) \tag{3}$$

$$\boldsymbol{\mu} \sim \mathrm{N}(\mathbf{0}, \mathbf{1}) \tag{4}$$

$$\boldsymbol{\beta} \sim \mathrm{N}(\mathbf{0}, \mathbf{1}) \tag{5}$$

- ctsem - open source R software

- ctsem - open source R software
  - Began life supporting mixed effects models, using OpenMx as backend.

- ctsem - open source R software
  - Began life supporting mixed effects models, using OpenMx as backend.
  - Now also supports fully random effects with Bayesian estimation, using Stan as backend.

- ctsem - open source R software
  - Began life supporting mixed effects models, using OpenMx as backend.
  - Now also supports fully random effects with Bayesian estimation, using Stan as backend.
- ctModel function constructs a ctsem model based on specified matrices of free and fixed parameters.

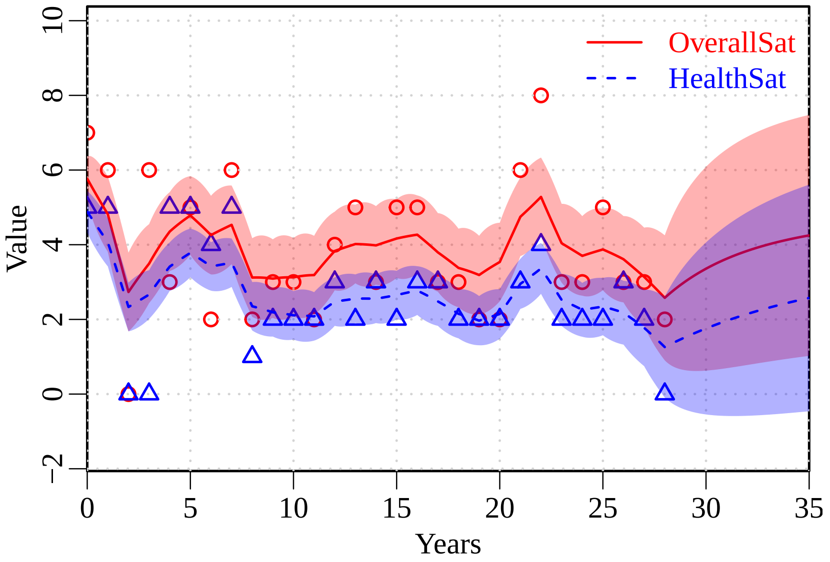
- ctsem - open source R software
  - Began life supporting mixed effects models, using OpenMx as backend.
  - Now also supports fully random effects with Bayesian estimation, using Stan as backend.
- ctModel function constructs a ctsem model based on specified matrices of free and fixed parameters.
- Modify individual differences, priors, and covariate effects as needed.

- ctsem - open source R software
  - Began life supporting mixed effects models, using OpenMx as backend.
  - Now also supports fully random effects with Bayesian estimation, using Stan as backend.
- ctModel function constructs a ctsem model based on specified matrices of free and fixed parameters.
- Modify individual differences, priors, and covariate effects as needed.
- ctStanFit function constructs a Stan model and calls rstan for estimation, using either Kalman filter for continuous variables or direct sampling of states for other measurement models.

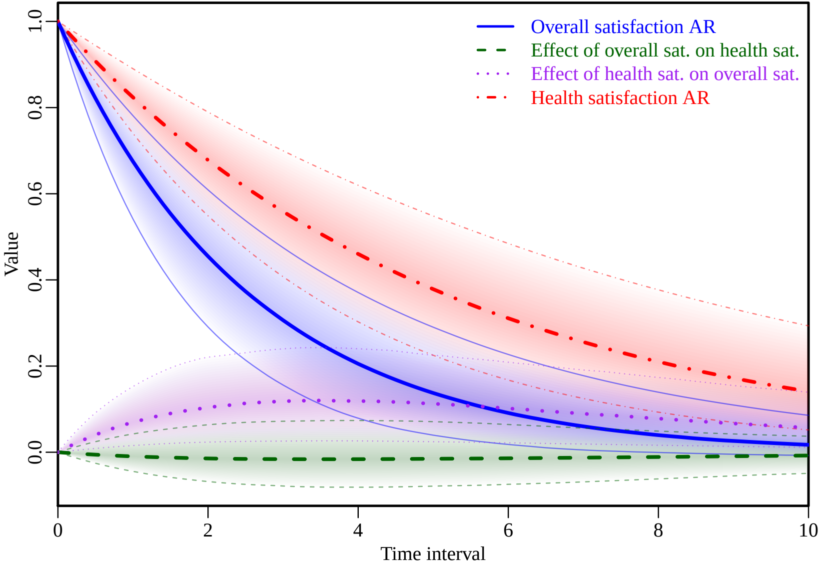
- 200 subjects at random from German socioeconomic panel, 1984-2012. Mean ages 30-70. Maximum of 29 obs. per subject.

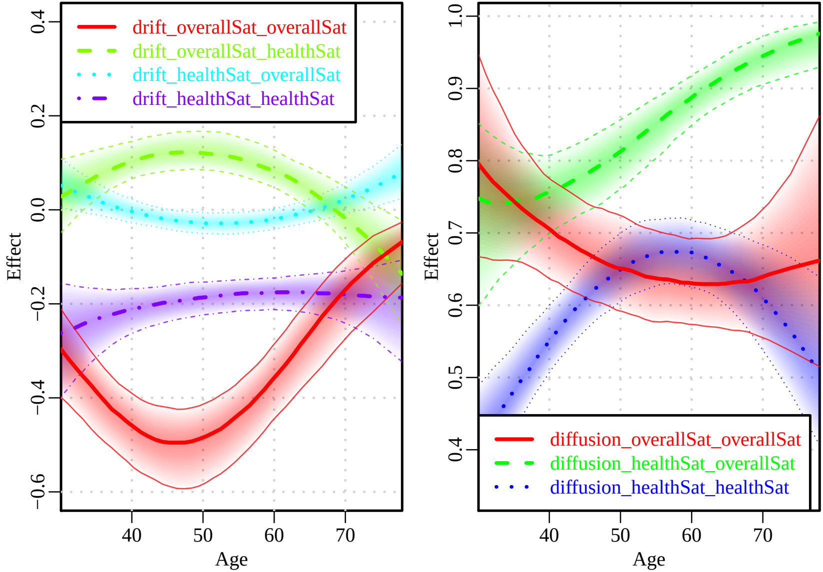
- 200 subjects at random from German socioeconomic panel, 1984-2012. Mean ages 30-70. Maximum of 29 obs. per subject.
- Scores of overall life satisfaction and satisfaction with health, 0-10 point scale.

- 200 subjects at random from German socioeconomic panel, 1984-2012. Mean ages 30-70. Maximum of 29 obs. per subject.
- Scores of overall life satisfaction and satisfaction with health, 0-10 point scale.
- Fit basic bivariate process model to examine individual dynamics, relations between dynamics, and differences in dynamics given age.

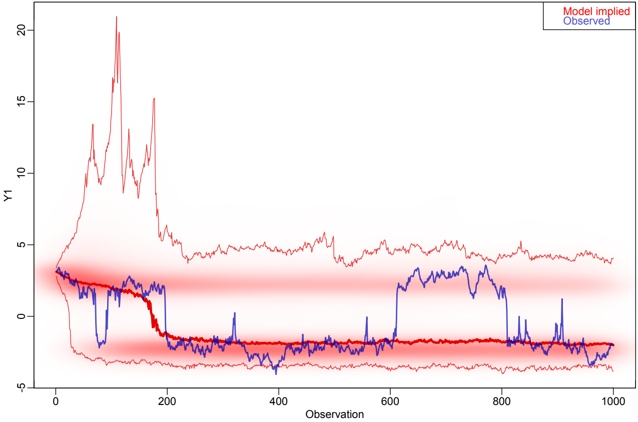


OverallSat
HealthSat
Value
Years
10
8
6
4
2
0
−2
0
5
10
15
20
25
30
35

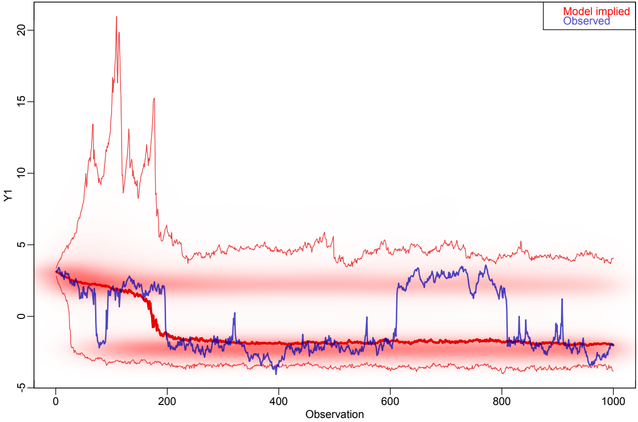

Overall satisfaction AR
Effect of overall sat. on health sat.
Effect of health sat. on overall sat.
Health satisfaction AR
Value
Time interval
0.0
0.2
0.4
0.6
0.8
1.0
0
2
4
6
8
10


drift_overallSat_overallSat
drift_overallSat_healthSat
drift_healthSat_overallSat
drift_healthSat_healthSat
Effect
Age
0.4
0.2
0.0
−0.2
−0.4
−0.6
40
50
60
70
diffusion_overallSat_overallSat
diffusion_healthSat_overallSat
diffusion_healthSat_healthSat
1.0
0.9
0.8
0.7
0.6
0.5
0.4

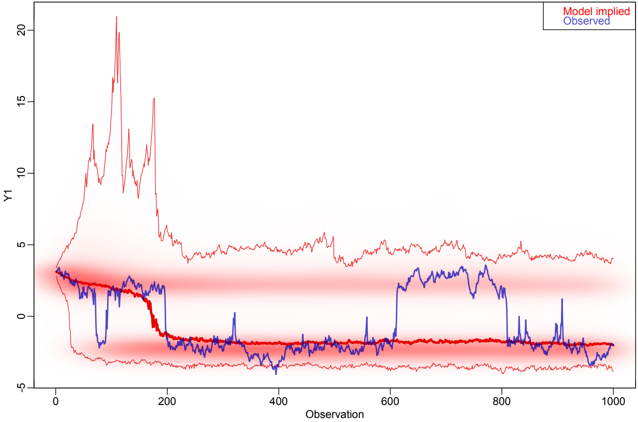
- Nonlinear / time varying models



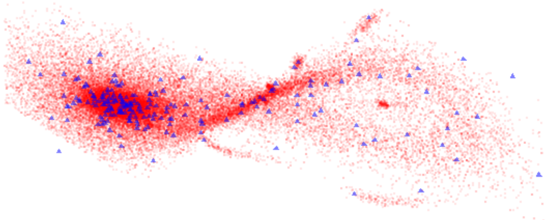
- Nonlinear / time varying models
- Optimization + importance sampling approach – much quicker than full Bayes



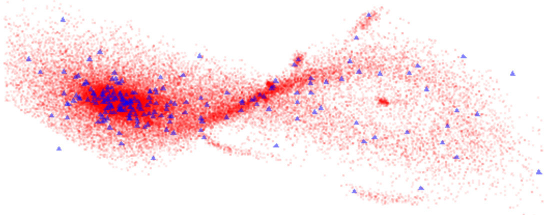
- Nonlinear / time varying models
- Optimization + importance sampling approach – much quicker than full Bayes
- Various posterior predictive plots



- ctsem and vignettes
  https://cran.r-project.org/web/packages/ctsem/index.html



- ctsem and vignettes
  https://cran.r-project.org/web/packages/ctsem/index.html
- Relevant articles:
  https://www.researchgate.net/profile/Charles_Driver



- ctsem and vignettes
  https://cran.r-project.org/web/packages/ctsem/index.html
- Relevant articles:
  https://www.researchgate.net/profile/Charles_Driver
- Thanks!