# A Dynamic Mixture Biometric Model of Cognitive Development in the NLSYC

Michael Hunter*,**, David Bard**, William Beasley***, Kelly Meredith****, & Joseph Rodgers*****

*University of Oklahoma
**University of Oklahoma Health Sciences Center
***Howard Live Oak, LLC
****Oklahoma City University
*****Vanderbilt University

Behavior Genetics Association
Charlottesville, VA
June 19, 2014

# In search of windows

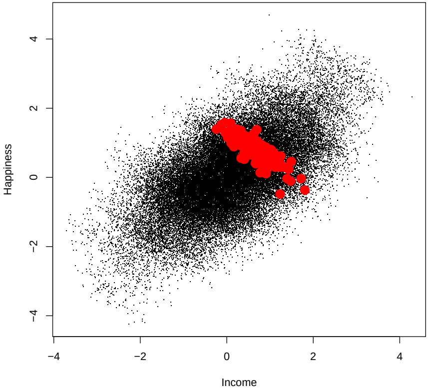
*Gene's, like Leibniz's monads, have no windows; the higher properties of life are emergent. . . . And once assembled, organisms have no windows.*

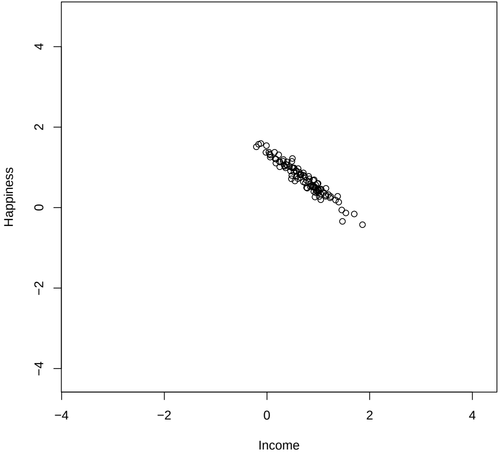
*-Edward O. Wilson "Sociobiology" Chapter 2*

# Outline

- Role of the individual
- Individual Behavior Genetics Modeling
- Examine NLSY79C Cognitive Development
  - NlsyLinks
  - OpenMx
- Dynamic Mixture ACE Modeling

# Perspective




Happiness
Income
4
2
0
−2
−4
−4
−2
0
2
4

# Where to go from here?

## No, it's not dinner time yet.

- Between-person models are valid, but (generally) only between people.
- Model individuals and processes.
- Balance the Idiographic/Nomothetic trade-off
- Options
  1. Assume complete heterogeneity: Separate processes
  2. Assume complete homogeneity: Same process
  3. Assume a mixture of homogeneous groups

# Where to go from here?

## No, it's not dinner time yet.

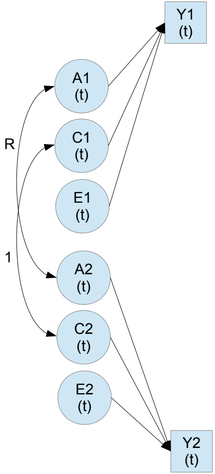
- Between-person models are valid, but (generally) only between people.
- Model individuals and processes.
- Balance the Idiographic/Nomothetic trade-off
- Options
  1. Assume complete heterogeneity: Separate processes
  2. Assume complete homogeneity: Same process
  3. **Assume a mixture of homogeneous groups**

# Where to go from here?

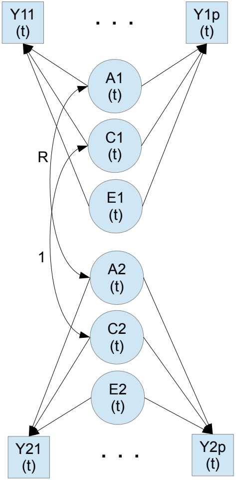
## No, it's not dinner time yet.

- Between-person models are valid, but (generally) only between people.
- Model individuals and processes.
- Balance the Idiographic/Nomothetic trade-off
- Options
  1. Assume complete heterogeneity: Separate processes
  2. Assume complete homogeneity: Same process
  3. **Assume a mixture of homogeneous groups**
- How do you model behavior genetic variability within people?

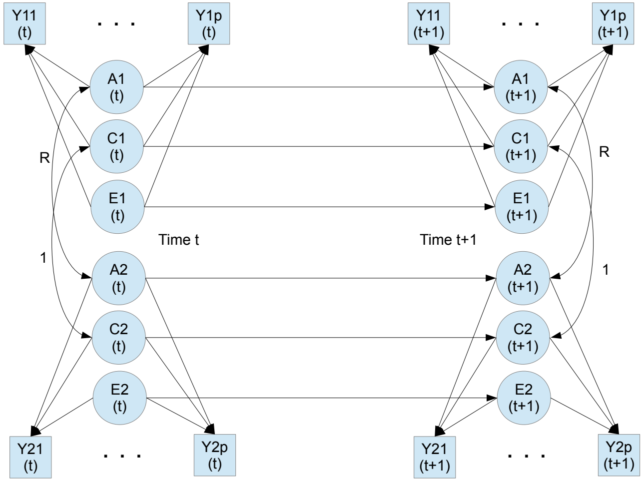
# Neale & Cardon (1992)



# Neale & Cardon (1992)



# Molenaar (2010) “On the limits . . . ”



# State Space Model

## Measurement

- Structural Equation Measurement Model

$$\boldsymbol{y}_i = \Lambda \boldsymbol{\eta}_i + K \boldsymbol{x}_i + \boldsymbol{\varepsilon}_i \quad \text{with} \quad \boldsymbol{\varepsilon}_i \sim \mathcal{N}(\mathbf{0}, \Theta) \tag{1}$$

- State Space Measurement Model

$$\boldsymbol{y}_i = \Lambda \boldsymbol{\eta}_i + K \boldsymbol{x}_i + \boldsymbol{\varepsilon}_i \quad \text{with} \quad \boldsymbol{\varepsilon}_i \sim \mathcal{N}(\mathbf{0}, \Theta) \tag{2}$$

# State Space Model

## Transition/Structural

- Structural Equation Structural Model

$$\boldsymbol{\eta}_i = B\boldsymbol{\eta}_i + \Gamma\boldsymbol{x}_i + \boldsymbol{\zeta}_i \quad \text{with} \quad \boldsymbol{\zeta}_i \sim \mathcal{N}\left(\mathbf{0}, \Psi\right) \tag{3}$$

- State Space Structural Model

$$\boldsymbol{\eta}_{i+1} = B\boldsymbol{\eta}_i + \Gamma\boldsymbol{x}_i + \boldsymbol{\zeta}_i \quad \text{with} \quad \boldsymbol{\zeta}_i \sim \mathcal{N}\left(\mathbf{0}, \Psi\right) \tag{4}$$

# State Space Model

## Transition/Structural

- Structural Equation Structural Model

$$\boldsymbol{\eta}_i = B\boldsymbol{\eta}_i + \Gamma\boldsymbol{x}_i + \boldsymbol{\zeta}_i \quad \text{with} \quad \boldsymbol{\zeta}_i \sim \mathcal{N}\left(\boldsymbol{0}, \Psi\right) \tag{3}$$

- State Space Structural Model

$$\boldsymbol{\eta}_{i+1} = B\boldsymbol{\eta}_i + \Gamma\boldsymbol{x}_i + \boldsymbol{\zeta}_i \quad \text{with} \quad \boldsymbol{\zeta}_i \sim \mathcal{N}\left(\boldsymbol{0}, \Psi\right) \tag{4}$$

- Differential Equation in Discrete Time

# State Space Model

## Transition/Structural

- Structural Equation Structural Model

$$\boldsymbol{\eta}_i = B\boldsymbol{\eta}_i + \Gamma \boldsymbol{x}_i + \boldsymbol{\zeta}_i \quad \text{with} \quad \boldsymbol{\zeta}_i \sim \mathcal{N}\left(\mathbf{0}, \Psi\right) \tag{3}$$

- State Space Structural Model

$$\boldsymbol{\eta}_{i+1} = B\boldsymbol{\eta}_i + \Gamma \boldsymbol{x}_i + \boldsymbol{\zeta}_i \quad \text{with} \quad \boldsymbol{\zeta}_i \sim \mathcal{N}\left(\mathbf{0}, \Psi\right) \tag{4}$$

- Differential Equation in Discrete Time
- Implemented by me in OpenMx 2.0 Beta Release
- Uses Kalman filter

# Check out the 2.0 Beta Release

## and State Space Models

```
source('http://openmx.psyc.virginia.edu/getOpenMxBeta.R')
require(OpenMx)
?mxExpectationStateSpace
```

# Longitudinal Behavior Genetics Models

## Short Story

- Does this work? . . . Yes!
- Moreover
  - Can be idiographic: different heritabilities for Sibling 1 versus Sibling 2
  - Able to estimate additive genetics *or* common environments correlation parameter
  - Extendable to time-varying heritabilities with extended Kalman filter time-varying loadings

# NLSY

## National Longitudinal Survey of Youth

- Original 1979 Generation 1 Sample
  - Nationally representative household probability sample
  - 12,686 young men and women
  - 14-22 years old
- Children (NLSYC, Generation 2)
  - Children of Generation 1 Females
  - Beginning 1986
  - 3276 different mothers
  - 11,075 NLSY-Children kinship links

# NlsyLinks

## The R Package

- Uses responses by children and parents to lots of questions about
  - parentage
  - relatedness
  - cohabitation
  - so on
- Makes a deterministic, accurate classification of
  - siblings (full, half)
  - cousins
  - aunt/niece
  - parent/child
  - so on

# Links79Pair

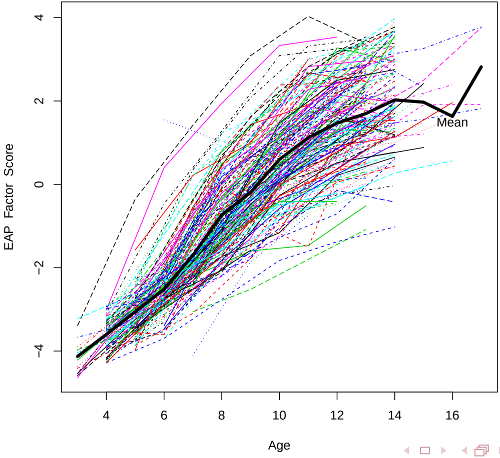
```
> require(NlsyLinks)
Loading required package: NlsyLinks
> data(Links79Pair)
> head(Links79Pair)
  ExtendedID Subject1Tag Subject2Tag   R RelationshipPath
1          2         200         201 0.5      ParentChild
2          2         200         202 0.5      ParentChild
3          2         201         202 0.5     Gen2Siblings
4          3         300         301 0.5      ParentChild
5          3         301         302 0.5     Gen2Siblings
6          3         300         302 0.5      ParentChild
> 
```

# Present Analyses

- Goal: look at longitudinal variation in cognitive ability using biometrically informed models
- 5 cognitive variables
  - PIAT Math, Reading Comp, Reading Recog (Ages 5-14)
  - PPVT (Ages 3-14)
  - WISC Memory for Digit Span (Ages 7-11)
- Modal Number of Occasions: 2-4

# Longitudinal Factor Scores

## method of Estabrook & Neale (2013)

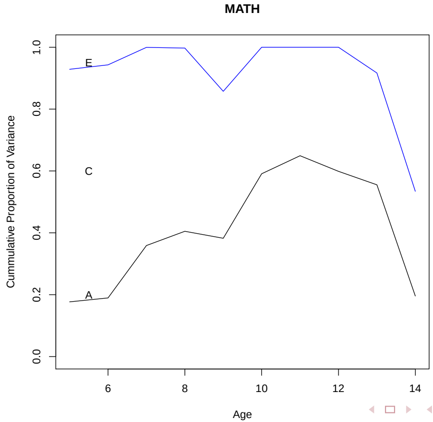


# Repeated Cross Sections

## ACE Slice at Different Ages

Haworth et al., 2010

$h^2$(childhood) $= .41$, $h^2$(adolescence) $= .55$, $h^2$(YA)$=.66$



# Multisubject State Space ACE

## Random Subsample 1

| | A | C | E |
|---|---|---|---|
| AREG | 0.82 | 0.89 | 0.85 |
| DIGIT | 0.14 | 0.66 | 0.19 |
| MATH | 0.27 | 0.71 | 0.02 |
| RECOG | 0.03 | 0.74 | 0.24 |
| COMP | 0.07 | 0.79 | 0.15 |
| PPVT | 0.00 | 1.00 | 0.00 |

# Multisubject State Space ACE

## Random Subsample 2

| | A | C | E |
|---|---|---|---|
| AREG | 0.87 | 0.88 | 0.84 |
| DIGIT | 0.19 | 0.57 | 0.24 |
| MATH | 0.15 | 0.80 | 0.05 |
| RECOG | 0.00 | 0.68 | 0.32 |
| COMP | 0.05 | 0.78 | 0.17 |
| PPVT | 0.00 | 1.00 | 0.00 |

# Multisubject Mixture State Space ACE

## Random Subsample 1

| | Model 1 | | | Model 2 | | |
|---|---|---|---|---|---|---|
| | $\pi = .11$ | | | $\pi = .89$ | | |
| | A1 | C1 | E1 | A2 | C2 | E2 |
| AREG | 0.80 | 0.89 | 0.84 | - | - | - |
| DIGIT | 0.00 | 0.16 | 0.84 | 0.03 | 0.83 | 0.14 |
| MATH | 0.02 | 0.98 | 0.00 | 0.28 | 0.69 | 0.03 |
| RECOG | 0.26 | 0.39 | 0.35 | 0.04 | 0.76 | 0.14 |
| COMP | 0.63 | 0.37 | 0.00 | 0.07 | 0.79 | 0.14 |
| PPVT | 0.00 | 1.00 | 0.00 | 0.00 | 1.00 | 0.00 |

# Multisubject Mixture State Space ACE

## Random Subsample 1

| | Model 1 | | | Model 2 | | |
|---|---|---|---|---|---|---|
| | $\pi = .11$ | | | $\pi = .89$ | | |
| | A1 | C1 | E1 | A2 | C2 | E2 |
| AREG | 0.80 | 0.89 | 0.84 | - | - | - |
| DIGIT | **0.00** | 0.16 | 0.84 | **0.03** | 0.83 | 0.14 |
| MATH | **0.02** | 0.98 | 0.00 | **0.28** | 0.69 | 0.03 |
| RECOG | **0.26** | 0.39 | 0.35 | **0.04** | 0.76 | 0.14 |
| COMP | **0.63** | 0.37 | 0.00 | **0.07** | 0.79 | 0.14 |
| PPVT | **0.00** | 1.00 | 0.00 | **0.00** | 1.00 | 0.00 |

# Multisubject Mixture State Space ACE

## Random Subsample 2

| | Model 1 | | | Model 2 | | |
|---|---|---|---|---|---|---|
| | $\pi = .001$ | | | $\pi = .999$ | | |
| | A1 | C1 | E1 | A2 | C2 | E2 |
| AREG | 0.84 | 0.89 | 0.85 | - | - | - |
| DIGIT | 0.28 | 0.01 | 0.71 | 0.33 | 0.54 | 0.13 |
| MATH | 0.04 | 0.25 | 0.72 | 0.27 | 0.71 | 0.02 |
| RECOG | 0.25 | 0.15 | 0.60 | 0.04 | 0.70 | 0.26 |
| COMP | 0.30 | 0.21 | 0.49 | 0.08 | 0.77 | 0.15 |
| PPVT | 0.30 | 0.37 | 0.33 | 0.00 | 1.00 | 0.00 |

# Multisubject Mixture State Space ACE

## Random Subsample 2

| | Model 1 | | | Model 2 | | |
|---|---|---|---|---|---|---|
| | $\pi = .001$ | | | $\pi = .999$ | | |
| | A1 | C1 | E1 | A2 | C2 | E2 |
| AREG | 0.84 | 0.89 | 0.85 | - | - | - |
| DIGIT | **0.28** | 0.01 | 0.71 | **0.33** | 0.54 | 0.13 |
| MATH | **0.04** | 0.25 | 0.72 | **0.27** | 0.71 | 0.02 |
| RECOG | **0.25** | 0.15 | 0.60 | **0.04** | 0.70 | 0.26 |
| COMP | **0.30** | 0.21 | 0.49 | **0.08** | 0.77 | 0.15 |
| PPVT | **0.30** | 0.37 | 0.33 | **0.00** | 1.00 | 0.00 |

# Summary

- Within- & Between-person moments are often different
- Within-person behavior genetic models are possible
- Cross-sectional ACE results do *not* resemble homogeneous longitudinal ACE results
- Population might not be homogeneous
- Dynamic Mixture ACE modeling could find unknown homogeneous groups

# Future Work

- Incorporation of SES
- Mixture on larger sample: memory issues
- Combinations of growth trends and autoregressive dynamics in state space ACE models
- Feedback between environment and phenotype?
- Further autoregressive relationships between ACE components?

*Development and Psychopathology* **25** (2013), 7–16

doi:10.1017/S0954579412000867

***SPECIAL SECTION ARTICLE***

# Phenotype–environment correlations in longitudinal twin models

CHRISTOPHER R. BEAM AND ERIC TURKHEIMER
*University of Virginia*

- Simulations that suggest feedback loop between environments and phenotype
- Changing the Gene-Environment correlation within families

# Acknowledgments


- "NLSY Kinship Links: Reliable and Valid Sibling Identification" NIH R01 HD065865
- Advisors: Joseph L. Rodgers & Hairong Song
- NlsyLinks Development Team: Will Beasley, Joe Rodgers, David Bard, Kelly Meredith
- OpenMx Development Team: Steven Boker, Michael Neale, Hermine Maes, Timothy Brick, Ryne Estabrook, Timo von Oertzen, Joshua Pritikin

# Thank You

mhunter@ou.edu

# iFACE

Molenaar (2010) “On the limits ...”

A Single Simulated Twin Pair with an Idiographic Filter ACE (iFACE) Model Fit as a State Space Model in OpenMx

Table : Autoregressive Parameters

| Name | True Value | Est Value | True Prop | Est Prop |
|---|---|---|---|---|
| $areg_1$ | 0.7 | 0.646 | - | - |
| $areg_2$ | 0.7 | 0.670 | - | - |
| $creg_1$ | 0.7 | 0.750 | - | - |
| $creg_2$ | 0.7 | 0.781 | - | - |
| $ereg_1$ | 0.7 | 0.699 | - | - |
| $ereg_2$ | 0.7 | 0.727 | - | - |
| addcor: $R$ | 0.5 | 0.507 | - | - |

# iFACE

## Molenaar (2010) “On the limits ...”

### Additive Genetics Parameters

| Name | True Value | Est Value | True Prop | Est Prop |
|---|---|---|---|---|
| $\alpha_{tw1,\ v1}$ | 1.73 | 1.32 | 0.33 | 0.19 |
| $\alpha_{tw1,\ v2}$ | 3.64 | 3.46 | 0.86 | 0.73 |
| $\alpha_{tw1,\ v3}$ | 1.30 | 1.01 | 0.35 | 0.21 |
| $\alpha_{tw1,\ v4}$ | 2.65 | 2.30 | 0.57 | 0.40 |
| $\alpha_{tw2,\ v1}$ | 2.61 | 2.82 | 0.74 | 0.80 |
| $\alpha_{tw2,\ v2}$ | 1.03 | 1.18 | 0.20 | 0.26 |
| $\alpha_{tw2,\ v3}$ | 1.68 | 1.81 | 0.54 | 0.62 |
| $\alpha_{tw2,\ v4}$ | 1.57 | 1.74 | 0.50 | 0.57 |

# iFACE

## Molenaar (2010) "On the limits . . ."

### Common Environments Parameters

| Name | True Value | Est Value | True Prop | Estimated Prop |
|---|---|---|---|---|
| $\beta_{v1}$ | 1.00 | 0.96 | 0.11 | 0.10 |
| $\beta_{v2}$ | 1.00 | 0.95 | 0.06 | 0.06 |
| $\beta_{v3}$ | 1.00 | 0.91 | 0.21 | 0.17 |
| $\beta_{v4}$ | 1.00 | 0.93 | 0.08 | 0.06 |
| $\beta_{v1}$ | 1.00 | 0.96 | 0.11 | 0.09 |
| $\beta_{v2}$ | 1.00 | 0.95 | 0.19 | 0.17 |
| $\beta_{v3}$ | 1.00 | 0.91 | 0.19 | 0.16 |
| $\beta_{v4}$ | 1.00 | 0.93 | 0.20 | 0.16 |

# iFACE

## Molenaar (2010) "On the limits ..."

Common Environments Parameters

| Name | True Value | Est Value | True Prop | Est Prop |
|---|---|---|---|---|
| $\gamma_{tw1,\ v1}$ | 2.27 | 2.59 | 0.56 | 0.71 |
| $\gamma_{tw1,\ v2}$ | 1.10 | 1.84 | 0.08 | 0.21 |
| $\gamma_{tw1,\ v3}$ | 1.46 | 1.72 | 0.44 | 0.62 |
| $\gamma_{tw1,\ v4}$ | 2.08 | 2.66 | 0.35 | 0.53 |
| $\gamma_{tw2,\ v1}$ | 1.20 | 1.05 | 0.16 | 0.11 |
| $\gamma_{tw2,\ v2}$ | 1.79 | 1.75 | 0.61 | 0.57 |
| $\gamma_{tw2,\ v3}$ | 1.17 | 1.08 | 0.26 | 0.22 |
| $\gamma_{tw2,\ v4}$ | 1.22 | 1.17 | 0.30 | 0.26 |

# Linear Growth

## Explicit and Recursive

- Explicit
- Recursive

# Linear Growth

## Explicit and Recursive

- Explicit
- Linear

$$y_i = b_0 + b_1 t_i + e_i \tag{5}$$

- Recursive
- Linear

$$y_{t+1} = y_t + b_1 + e_t \quad \text{with} \quad y_0 = b_0 \tag{7}$$

# Linear Growth

## Explicit and Recursive

- Explicit
- Linear

$$y_i = b_0 + b_1 t_i + e_i \tag{5}$$

- Quadratic

$$y_i = b_0 + b_1 t_i + b_2 t_i^2 + e_i \tag{6}$$

- Recursive
- Linear

$$y_{t+1} = y_t + b_1 + e_t \quad \text{with} \quad y_0 = b_0 \tag{7}$$

- Quadratic

$$y_{t+1} = y_t + b_1 + b_2 + 2b_2 t + e_t \quad \text{with} \quad y_0 = b_0 \tag{8}$$

# Polynomial Growth

## Explicit and Recursive

- Explicit
-

$$y_i = b_0 + b_1 t_i + e_i \tag{9}$$

$$y_i = b_0 + b_1 t_i + b_2 t_i^2 + e_i \tag{10}$$

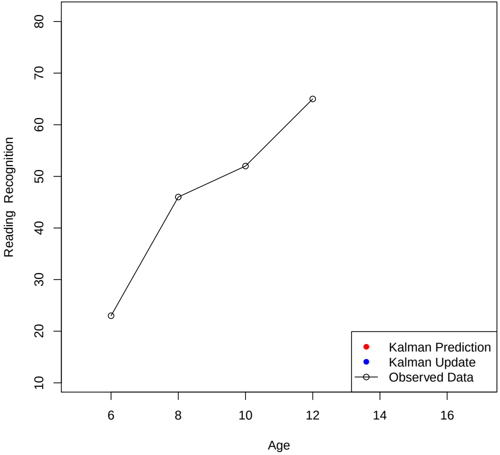
$$y_i = b_0 + b_1 t_i + b_2 t_i^2 + b_3 t_i^3 + e_i \tag{11}$$

- Recursive
-

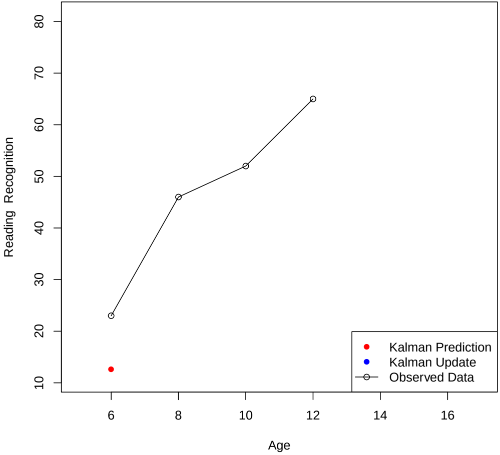
$$y_{t+1} = y_t + b_1 + e_t \quad \text{with} \quad y_0 = b_0 \tag{12}$$

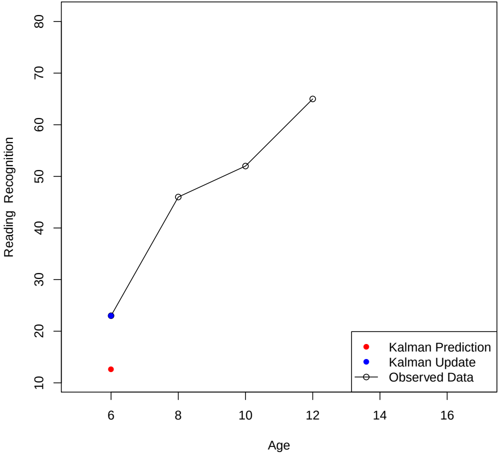
$$y_{t+1} = y_t + b_1 + b_2 + 2b_2 t + e_t \quad \text{with} \quad y_0 = b_0 \tag{13}$$

$$y_{t+1} = y_t + b_1 + b_2 + b_3 + 2b_2 t + 2b_3 t + 3b_3 t^2 + e_t \quad \text{with} \quad y_0 = b_0 \tag{14}$$

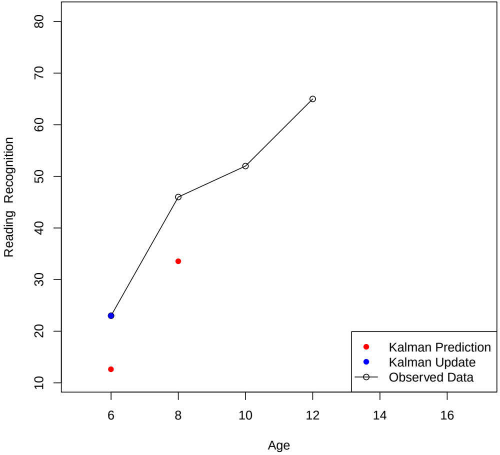


Reading Recognition
10
20
30
40
50
60
70
80
6
8
10
12
14
16
Age
Kalman Prediction
Kalman Update
Observed Data

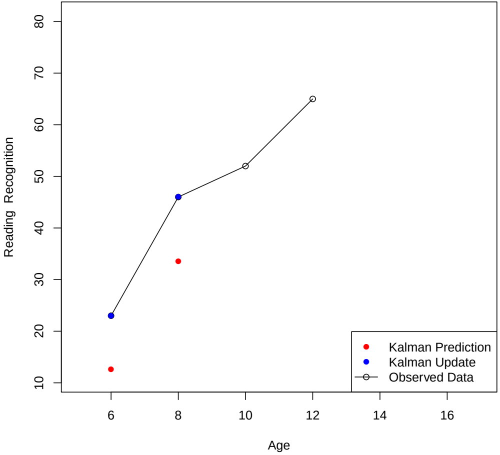

Reading Recognition
80
70
60
50
40
30
20
10
6
8
10
12
14
16
Age
Kalman Prediction
Kalman Update
Observed Data

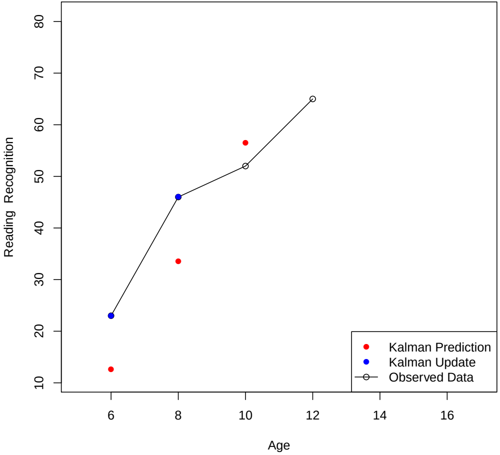


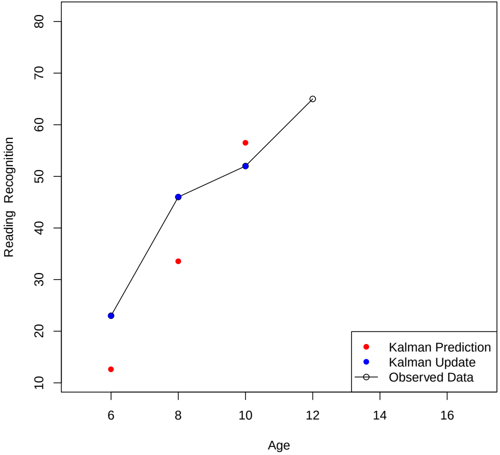
Reading Recognition
10
20
30
40
50
60
70
80
6
8
10
12
14
16
Age
Kalman Prediction
Kalman Update
Observed Data

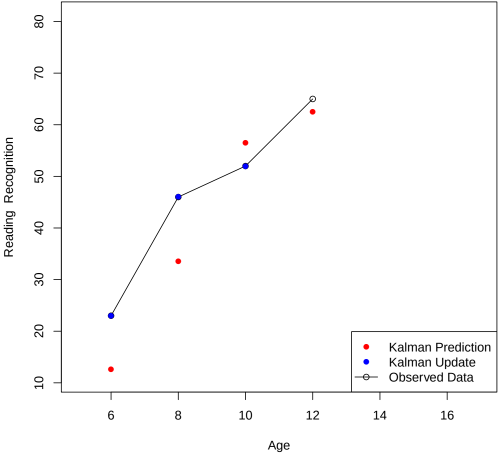


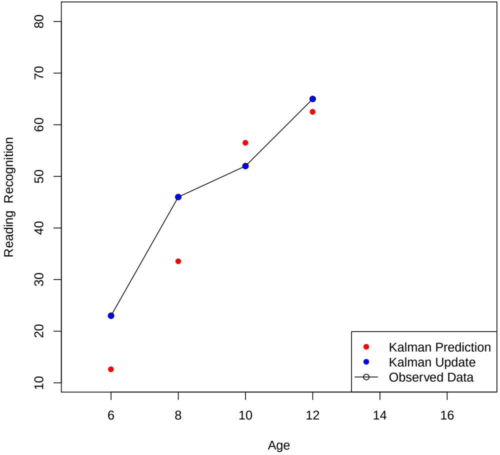
Reading Recognition
10
20
30
40
50
60
70
80
6
8
10
12
14
16
Age
Kalman Prediction
Kalman Update
Observed Data

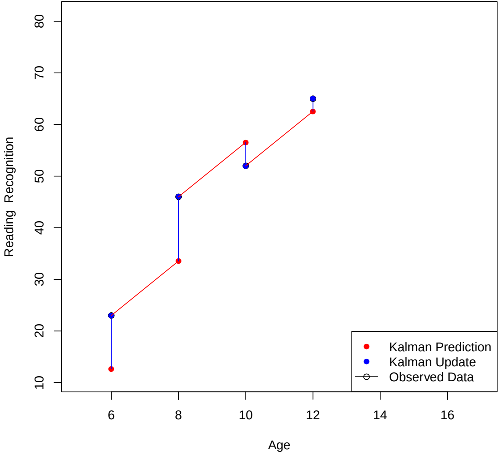

Reading Recognition
Age
10
20
30
40
50
60
70
80
6
8
10
12
14
16
Kalman Prediction
Kalman Update
Observed Data

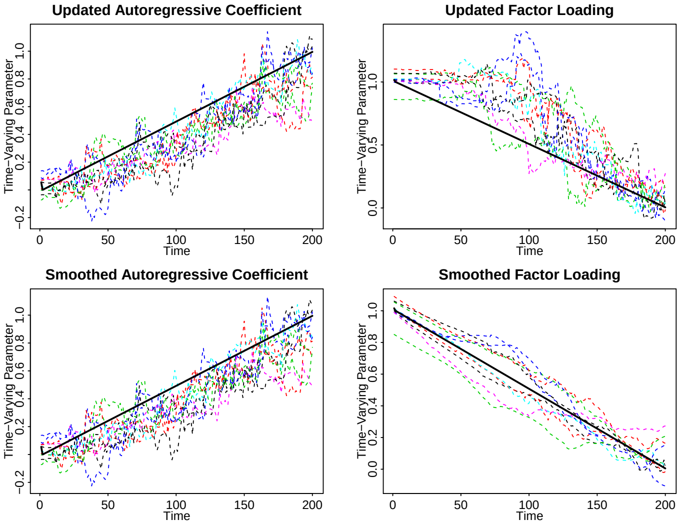

Reading Recognition
10
20
30
40
50
60
70
80
Age
6
8
10
12
14
16
Kalman Prediction
Kalman Update
Observed Data


Reading Recognition
10
20
30
40
50
60
70
80
6
8
10
12
14
16
Age
Kalman Prediction
Kalman Update
Observed Data


Reading Recognition
10
20
30
40
50
60
70
80
6
8
10
12
14
16
Age
Kalman Prediction
Kalman Update
Observed Data


Reading Recognition
10
20
30
40
50
60
70
80
6
8
10
12
14
16
Age
Kalman Prediction
Kalman Update
Observed Data


Reading Recognition
10
20
30
40
50
60
70
80
6
8
10
12
14
16
Age
Kalman Prediction
Kalman Update
Observed Data


Updated Autoregressive Coefficient
Updated Factor Loading
Smoothed Autoregressive Coefficient
Smoothed Factor Loading
Time−Varying Parameter
Time
0
50
100
150
200
−0.2
0.2
0.4
0.6
0.8
1.0
0.0
0.5