

Table of contents

- 1 Overview of models
- 2 General notation
- 3 Model 1
- 4 Model 2
- 5 Model 3
- 6 References

Overview of models

Overview of models

The following tables provide an overview of the specified models, with model extensions/changes highlighted in **bold**. Please note that the input data are assumed to be complete for each model (i.e., no missing data).

Model	Dependent variable	Constant & trend component
1	continuous, Normal, iid err	pooled, linear
2	continuous, Normal, iid err	individual effects , linear
3	continuous, Normal, iid err	individual effects, quadratic
4	count, Poisson , iid err	individual effects, quadratic

Table 1: Overview of specified models

Overview of models — continued

Placeholder

Model	Dependent variable	Constant & trend component
5	placeholder	placeholder

Table 2: Overview of specified models — continued

General notation

General notation

Individual n , for $n = 1, \dots, N$, where N is the number of individuals

Time period t , for $t = 1, \dots, T$, where T is the number of time periods

\mathbf{Y}^{obs} is a $N \times T$ matrix representing the observed dependent variable, where *obs* refers to *sim* (simulated) or *act* (actual) data

\mathbf{Y}^{pred} is a $N \times T$ matrix representing the predicted dependent variable

\mathbf{X} is a matrix of size $N \times T$ representing the explanatory variable (e.g., time periods, starting at zero)

General notation — continued

Latent class (aka, sub-population) c , for $c = 1, \dots, C$, where C is the number of classes

Π is a $N \times C$ matrix representing the mixture proportion (e.g., the probability that individual $n = 1$ belongs to class $c = 1$); thus,

$$0 \leq \pi_{n,c} \leq 1 \quad \text{and} \quad \sum_{c=1}^C \pi_{n,c} = 1 \quad (1)$$

Model 1

Model 1 — dependent variable

Continuous

Normal (aka, Gaussian)

iid err: independent (i.e., uncorrelated over individuals and time periods)
and identically distributed errors

No missing data

Model 1 — constant and trend component

Constant **without** interindividual differences within-class (i.e., the constant is allowed to vary between classes but not within classes)

Linear (i.e., non-stationary, deterministic) trend component **without** interindividual differences within-class (i.e., the trend component is allowed to vary between classes but not within classes)

Based on Koop (2003), the assumptions of **no** interindividual differences within-class regarding the constant and trend component correspond to a pooled model (i.e., a pooled model for each class)

According to Ram and Grimm (2009), latent class growth analysis (or LCGA for short) refers to a subset of GMMs assuming **no** interindividual differences within-class regarding the constant and trend component

Model 1 — likelihood

$$y_{n,t}^{obs} \sim \sum_{c=1}^C \pi_{n,c} \text{Normal}(\mu_{c,n,t}, \sigma_c) \quad (2)$$

where \mathbf{M} is a C -tuple containing $N \times T$ matrices, and $\boldsymbol{\sigma}$ is a row vector of size C .

$$\mu_{c,n,t} = \beta_{0,c} + \beta_{1,c} x_{n,t} \quad (3)$$

where β_0 is a row vector of size C representing the constant, and β_1 is a row vector of size C with

$$\beta_{1,c} < \beta_{1,c+1} \quad (4)$$

for identification. In combination with \mathbf{X} , β_1 represents the linear trend component.

Model 1 — prior

$$\boldsymbol{\pi}_n \sim \textit{Dirichlet}(\boldsymbol{\alpha}) \quad (5)$$

where $\boldsymbol{\alpha}$ is a row vector of size C with $\alpha_c = 1$ (i.e., $\boldsymbol{\pi}_n$ is assigned a flat, proper prior).

$$\beta_{0,c} \sim \textit{Normal}(0, 5) \quad (6)$$

$$\beta_{1,c} \sim \textit{Normal}(0, 1) \quad (7)$$

$$\sigma_c \sim \textit{Normal}(0, 1) \mathbf{1}(\sigma_c > 0) \quad (8)$$

where the Normal distribution is truncated to the left at zero, modeled via the indicator function $\mathbf{1}(\sigma_c > 0)$.

Model 2

Model 2 — dependent variable

Continuous

Normal (aka, Gaussian)

iid err: independent (i.e., uncorrelated over individuals and time periods)
and identically distributed errors

No missing data

Model 2 — constant and trend component

Constant **with** interindividual differences within-class (i.e., the constant is allowed to vary between classes and within classes)

Linear (i.e., non-stationary, deterministic) trend component **without** interindividual differences within-class (i.e., the trend component is allowed to vary between classes but not within classes)

Based on Koop (2003), the assumptions of **interindividual differences** within-class regarding the constant and **no interindividual differences** within-class regarding the trend component correspond to an individual effects model (i.e., an individual effects model for each class)

Model 2 — likelihood

$$y_{n,t}^{obs} \sim \sum_{c=1}^C \pi_{n,c} \text{Normal}(\mu_{c,n,t}, \sigma_c) \quad (2)$$

where \mathbf{M} is a C -tuple containing $N \times T$ matrices, and $\boldsymbol{\sigma}$ is a row vector of size C .

$$\mu_{c,n,t} = \beta_{0,n,c} + \beta_{1,c} x_{n,t} \quad (9)$$

where \mathbf{B}_0 is a $N \times C$ matrix representing the constant, and β_1 is a row vector of size C with

$$\beta_{1,c} < \beta_{1,c+1} \quad (4)$$

for identification. In combination with \mathbf{X} , β_1 represents the linear trend component.

Model 2 — prior

$$\boldsymbol{\pi}_n \sim \textit{Dirichlet}(\boldsymbol{\alpha}) \quad (5)$$

where $\boldsymbol{\alpha}$ is a row vector of size C with $\alpha_c = 1$ (i.e., $\boldsymbol{\pi}_n$ is assigned a flat, proper prior).

$$\beta_{0,n,c} \sim \textit{Normal}(0, 10) \quad (10)$$

$$\beta_{1,c} \sim \textit{Normal}(0, 1) \quad (7)$$

$$\sigma_c \sim \textit{Normal}(0, 1) \mathbf{1}(\sigma_c > 0) \quad (8)$$

where the Normal distribution is truncated to the left at zero, modeled via the indicator function $\mathbf{1}(\sigma_c > 0)$.

Model 3

Koop, G. (2003). *Bayesian Econometrics*. Wiley.

Ram, N. & Grimm, K. J. (2009). Growth Mixture Modeling: A Method for Identifying Differences in Longitudinal Change Among Unobserved Groups. *International Journal of Behavioral Development*, 33(6), 565-576.