A generalized finite mixture model

Jang SCHILTZ (University of Luxembourg)

July 28, 2015
Outline

1 Nagin’s Finite Mixture Model
Outline

1. Nagin’s Finite Mixture Model

2. Generalizations of Nagin’s model
Outline

1. Nagin’s Finite Mixture Model
2. Generalizations of Nagin’s model
3. Our model
Outline

1. Nagin’s Finite Mixture Model
2. Generalizations of Nagin’s model
3. Our model
General description of Nagin’s model

We have a collection of individual trajectories.
General description of Nagin’s model

We have a collection of individual trajectories.

We try to divide the population into a number of homogenous sub-populations and to estimate, at the same time, a typical trajectory for each sub-population.
General description of Nagin’s model

We have a collection of individual trajectories.

We try to divide the population into a number of homogenous sub-populations and to estimate, at the same time, a typical trajectory for each sub-population.

Hence, this model can be interpreted as functional fuzzy cluster analysis.
General description of Nagin’s model

We have a collection of individual trajectories.

We try to divide the population into a number of homogenous sub-populations and to estimate, at the same time, a typical trajectory for each sub-population.

Hence, this model can be interpreted as functional fuzzy cluster analysis.

Finite mixture model (Daniel S. Nagin (Carnegie Mellon University))
General description of Nagin’s model

We have a collection of individual trajectories.

We try to divide the population into a number of homogenous sub-populations and to estimate, at the same time, a typical trajectory for each sub-population.

Hence, this model can be interpreted as functional fuzzy cluster analysis.

Finite mixture model (Daniel S. Nagin (Carnegie Mellon University))
- mixture: population composed of a mixture of unobserved groups
General description of Nagin’s model

We have a collection of individual trajectories.

We try to divide the population into a number of homogenous sub-populations and to estimate, at the same time, a typical trajectory for each sub-population.

Hence, this model can be interpreted as functional fuzzy cluster analysis.

Finite mixture model (Daniel S. Nagin (Carnegie Mellon University))

- mixture: population composed of a mixture of unobserved groups
- finite: sums across a finite number of groups
Consider a population of size $N$ and a variable of interest $Y$. 

Let $Y_i = y_{i1}, y_{i2}, \ldots, y_{iT}$ be $T$ measures of the variable, taken at times $t_1, \ldots, t_T$ for subject number $i$. 

$\pi_j$: probability of a given subject to belong to group number $j$. 

$\Rightarrow \pi_j$ is the size of group $j$. 

$P(Y_i) = \sum_{j=1}^{\pi_j} P_j(Y_i)$, (1) 

where $P_j(Y_i)$ is probability of $Y_i$ if subject $i$ belongs to group $j$. 

Jang SCHILTZ (University of Luxembourg)
The Likelihood Function (1)

Consider a population of size $N$ and a variable of interest $Y$.

Let $Y_i = y_{i1}, y_{i2}, ..., y_{iT}$ be $T$ measures of the variable, taken at times $t_1, ..., t_T$ for subject number $i$. 
The Likelihood Function (1)

Consider a population of size $N$ and a variable of interest $Y$.

Let $Y_i = y_{i1}, y_{i2}, \ldots, y_{iT}$ be $T$ measures of the variable, taken at times $t_1, \ldots t_T$ for subject number $i$.

$\pi_j$ : probability of a given subject to belong to group number $j$
The Likelihood Function (1)

Consider a population of size \( N \) and a variable of interest \( Y \).

Let \( Y_i = y_{i1}, y_{i2}, \ldots, y_{iT} \) be \( T \) measures of the variable, taken at times \( t_1, \ldots, t_T \) for subject number \( i \).

\( \pi_j \): probability of a given subject to belong to group number \( j \)

\[ \Rightarrow \pi_j \text{ is the size of group } j. \]
The Likelihood Function (1)

Consider a population of size $N$ and a variable of interest $Y$.

Let $Y_i = y_{i1}, y_{i2}, \ldots, y_{iT}$ be $T$ measures of the variable, taken at times $t_1, \ldots, t_T$ for subject number $i$.

$\pi_j$ : probability of a given subject to belong to group number $j$

$\Rightarrow \pi_j$ is the size of group $j$.

$\Rightarrow P(Y_i) = \sum_{j=1}^{r} \pi_j P^j(Y_i), \quad (1)$
The Likelihood Function (1)

Consider a population of size \( N \) and a variable of interest \( Y \).

Let \( Y_i = y_{i1}, y_{i2}, \ldots, y_{iT} \) be \( T \) measures of the variable, taken at times \( t_1, \ldots, t_T \) for subject number \( i \).

\( \pi_j \): probability of a given subject to belong to group number \( j \)

\[ \Rightarrow \pi_j \text{ is the size of group } j. \]

\[ \Rightarrow P(Y_i) = \sum_{j=1}^{r} \pi_j P^j(Y_i), \quad (1) \]

where \( P^j(Y_i) \) is probability of \( Y_i \) if subject \( i \) belongs to group \( j \).
The Likelihood Function (2)

Aim of the analysis: Find $r$ groups of trajectories of a given kind (for instance polynomials of degree 4, $P(t) = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \beta_4 t^4.$
Aim of the analysis: Find $r$ groups of trajectories of a given kind (for instance polynomials of degree 4, $P(t) = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \beta_4 t^4$).

Statistical Model:

$$y_{it} = \beta^j_0 + \beta_1^j t + \beta_2^j t^2 + \beta_3^j t^3 + \beta_4^j t^4 + \varepsilon_{it}, \quad (2)$$

where $\varepsilon_{it} \sim N(0, \sigma)$, $\sigma$ being a constant standard deviation.

We try to estimate a set of parameters $\Omega = \{\beta^j_0, \beta_1^j, \beta_2^j, \beta_3^j, \beta_4^j, \pi_j, \sigma\}$ which allow to maximize the probability of the measured data.
Possible data distributions

- Count data $\Rightarrow$ Poisson distribution
- Binary data $\Rightarrow$ Binary logit distribution
- Censored data $\Rightarrow$ Censored normal distribution
Possible data distributions

- count data $\Rightarrow$ Poisson distribution
Possible data distributions

- count data $\Rightarrow$ Poisson distribution
- binary data $\Rightarrow$ Binary logit distribution
Possible data distributions

- count data $\Rightarrow$ Poisson distribution
- binary data $\Rightarrow$ Binary logit distribution
- censored data $\Rightarrow$ Censored normal distribution
The case of a normal distribution (1)

Notations:

\[ \beta_j = \beta_{j0} + \beta_{j1}t + \beta_{j2}t^2 + \beta_{j3}t^3 + \beta_{j4}t^4. \]

\( \phi \): density of standard centered normal law.

Then,

\[ L = \frac{1}{\sigma^N} \prod_{i=1}^{r} \sum_{j=1}^{\pi} T \prod_{t=1}^{\phi} (y_{it} - \beta_j t) . \]

(3)

It is too complicated to get closed-forms equations.
The case of a normal distribution (1)

Notations:

\[ \beta^j t = \beta^j_0 + \beta^j_1 t + \beta^j_2 t^2 + \beta^j_3 t^3 + \beta^j_4 t^4. \]

\( \phi \): density of standard centered normal law.

Then,

\[ L = 1\sigma N \prod_{i=1}^{r} \sum_{j=1}^{\pi} T \prod_{t=1}^{\phi} \left( y_i t - \beta_j t \sigma \right). \]
The case of a normal distribution (1)

Notations :
- \( \beta^j t = \beta^j_0 + \beta^j_1 t + \beta^j_2 t^2 + \beta^j_3 t^3 + \beta^j_4 t^4. \)
- \( \phi: \) density of standard centered normal law.
The case of a normal distribution (1)

Notations:

- $\beta^j t = \beta^j_0 + \beta^j_1 t + \beta^j_2 t^2 + \beta^j_3 t^3 + \beta^j_4 t^4$.
- $\phi$: density of standard centered normal law.

Then,
The case of a normal distribution (1)

Notations:
- \( \beta_i t = \beta^i_0 + \beta^i_1 t + \beta^i_2 t^2 + \beta^i_3 t^3 + \beta^i_4 t^4 \).
- \( \phi \): density of standard centered normal law.

Then,

\[
L = \frac{1}{\sigma} \prod_{i=1}^{N} \left\{ \prod_{j=1}^{r} \pi_j \prod_{t=1}^{T} \phi \left( \frac{y_{it} - \beta^j t}{\sigma} \right) \right\}.
\] (3)
The case of a normal distribution (1)

Notations:
- \( \beta^j_t = \beta^j_0 + \beta^j_1 t + \beta^j_2 t^2 + \beta^j_3 t^3 + \beta^j_4 t^4 \).
- \( \phi \): density of standard centered normal law.

Then,
\[
L = \frac{1}{\sigma} \prod_{i=1}^{N} \sum_{j=1}^{r} \pi_j \prod_{t=1}^{T} \phi \left( \frac{y_{it} - \beta^j_t}{\sigma} \right).
\]

It is too complicated to get closed-forms equations.
An application example

The data:
Salaries of workers in the private sector in Luxembourg from 1987 to 2006.
About 1.3 million salary lines corresponding to 85,049 workers.

Some sociological variables:
- gender (male, female)
- nationality and residentship
- working sector
- year of birth
- year of birth of children
- age in the first year of professional activity
An application example

The data: Salaries of workers in the private sector in Luxembourg from 1987 to 2006.
An application example

**The data**: Salaries of workers in the private sector in Luxembourg from 1987 to 2006.

About 1.3 million salary lines corresponding to 85,049 workers.
An application example

**The data**: Salaries of workers in the private sector in Luxembourg from 1987 to 2006.

About 1.3 million salary lines corresponding to 85,049 workers.

Some sociological variables:
- gender (male, female)
An application example

**The data**: Salaries of workers in the private sector in Luxembourg from 1987 to 2006.

About 1.3 million salary lines corresponding to 85,049 workers.

Some sociological variables:
- gender (male, female)
- nationality and residentship
An application example

**The data:** Salaries of workers in the private sector in Luxembourg from 1987 to 2006.

About 1.3 million salary lines corresponding to 85,049 workers.

Some sociological variables:
- gender (male, female)
- nationality and residentship
- working sector
An application example

The data: Salaries of workers in the private sector in Luxembourg from 1987 to 2006.

About 1.3 million salary lines corresponding to 85,049 workers.

Some sociological variables:

- gender (male, female)
- nationality and residentship
- working sector
- year of birth
An application example

The data: Salaries of workers in the private sector in Luxembourg from 1987 to 2006.

About 1.3 million salary lines corresponding to 85,049 workers.

Some sociological variables:
- gender (male, female)
- nationality and residentship
- working sector
- year of birth
- year of birth of children
An application example

The data: Salaries of workers in the private sector in Luxembourg from 1987 to 2006.

About 1.3 million salary lines corresponding to 85,049 workers.

Some sociological variables:
- gender (male, female)
- nationality and residentship
- working sector
- year of birth
- year of birth of children
- age in the first year of professional activity
Result for 9 groups (dataset 1)
Result for 9 groups (dataset 1)
Outline

1. Nagin’s Finite Mixture Model
2. Generalizations of Nagin’s model
3. Our model
Predictors of trajectory group membership

\[ \pi_j(x_i) = e^{x_i \theta_j r \sum k=1 e^{x_i \theta_k r \sum t=1 \phi(y_{it} - \beta_j t \sigma)}}. \]
Predictors of trajectory group membership

\( x \): vector of variables potentially associated with group membership (measured before \( t_1 \)).
Predictors of trajectory group membership

x: vector of variables potentially associated with group membership (measured before $t_1$).

Multinomial logit model:

$$
\pi_j(x_i) = \frac{e^{x_i \theta_j}}{r \sum_{k=1} e^{x_i \theta_k}},
$$

(4)

where $\theta_j$ denotes the effect of $x_i$ on the probability of group membership.
Predictors of trajectory group membership

$x$ : vector of variables potentially associated with group membership (measured before $t_1$).

Multinomial logit model:

$$
\pi_j(x_i) = \frac{e^{x_i \theta_j}}{r \sum_{k=1} e^{x_i \theta_k}},
$$

(4)

where $\theta_j$ denotes the effect of $x_i$ on the probability of group membership.

$$
L = \frac{1}{\sigma} \prod_{i=1}^{N} \sum_{j=1}^{r} \frac{e^{x_i \theta_j}}{r \sum_{k=1} e^{x_i \theta_k}} \prod_{t=1}^{T} \phi \left( \frac{y_{it} - \beta^j t}{\sigma} \right).
$$

(5)
Adding covariates to the trajectories (1)

Let $z_1, ..., z_M$ be covariates potentially influencing $Y$. We are then looking for trajectories

$$y_{it} = \beta_{j0} + \beta_{j1}t + \beta_{j2}t^2 + \beta_{j3}t^3 + \beta_{j4}t^4 + \alpha_{j1}z_1 + ... + \alpha_{jM}z_M + \epsilon_{it},$$

(6)

where $\epsilon_{it} \sim N(0, \sigma)$, $\sigma$ being a constant standard deviation and $z_l$ are covariates that may depend or not upon time $t$.

Unfortunately the influence of the covariates in this model is limited to the intercept of the trajectory.
Adding covariates to the trajectories (1)

Let $z_1...z_M$ be covariates potentially influencing $Y$. 
Adding covariates to the trajectories (1)

Let $z_1...z_M$ be covariates potentially influencing $Y$.

We are then looking for trajectories

$$y_{it} = \beta_0^j + \beta_1^j t + \beta_2^j t^2 + \beta_3^j t^3 + \beta_4^j t^4 + \alpha_1^j z_1 + \ldots + \alpha_M^j z_M + \varepsilon_{it}, \quad (6)$$

where $\varepsilon_{it} \sim \mathcal{N}(0, \sigma)$, $\sigma$ being a constant standard deviation and $z_l$ are covariates that may depend or not upon time $t$. 

Adding covariates to the trajectories (1)

Let $z_1 \ldots z_M$ be covariates potentially influencing $Y$.  

We are then looking for trajectories 

$$y_{it} = \beta_0^j + \beta_1^j t + \beta_2^j t^2 + \beta_3^j t^3 + \beta_4^j t^4 + \alpha_1^j z_1 + \ldots + \alpha_M^j z_M + \varepsilon_{it},$$

(6)

where $\varepsilon_{it} \sim \mathcal{N}(0, \sigma)$, $\sigma$ being a constant standard deviation and $z_l$ are covariates that may depend or not upon time $t$.

Unfortunately the influence of the covariates in this model is limited to the intercept of the trajectory.
Adding covariates to the trajectories (2)
Adding covariates to the trajectories (2)
Outline

1 Nagin’s Finite Mixture Model

2 Generalizations of Nagin’s model

3 Our model
Our model
Our model

Let $x_1 \ldots x_M$ and $z_t$ be covariates potentially influencing $Y$. 

Our model

Let $x_1 \ldots x_M$ and $z_t$ be covariates potentially influencing $Y$.

We propose the following model:

$$
y_{it} = \left( \beta_0^j + \sum_{l=1}^{M} \alpha_{0l}^j x_{il} + \gamma_0^j z_{it} \right) + \left( \beta_1^j + \sum_{l=1}^{M} \alpha_{1l}^j x_{il} + \gamma_1^j z_{it} \right) t + \left( \beta_2^j + \sum_{l=1}^{M} \alpha_{2l}^j x_{il} + \gamma_2^j z_{it} \right) t^2 + \left( \beta_3^j + \sum_{l=1}^{M} \alpha_{3l}^j x_{il} + \gamma_3^j z_{it} \right) t^3 + \left( \beta_4^j + \sum_{l=1}^{M} \alpha_{4l}^j x_{il} + \gamma_4^j z_{it} \right) t^4 + \varepsilon_{it}^j,$$

where $\varepsilon_{it}^j \sim \mathcal{N}(0, \sigma^j)$, $\sigma^j$ being the standard deviation, constant in group $j$. 
Men versus women
Statistical Properties

The model's estimated parameters are the result of maximum likelihood estimation. As such, they are consistent and asymptotically normally distributed.

Confidence intervals of level $\alpha$ for the parameters $\beta_{jk}$:

$$CI_{\alpha}(\beta_{jk}) = \left[ \hat{\beta}_{jk} - t_{1-\alpha/2; N-(2+M)s_{\text{ASE}}(\hat{\beta}_{jk})}; \hat{\beta}_{jk} + t_{1-\alpha/2; N-(2+M)s_{\text{ASE}}(\hat{\beta}_{jk})} \right].$$

(7)

Confidence intervals of level $\alpha$ for the disturbance factor $\sigma_j$:

$$CI_{\alpha}(\sigma_j) = \left[ \sqrt{(N-(2+M)s-1)\hat{\sigma}_j^2 \chi^2_{1-\alpha/2}; N-(2+M)s-1}; \sqrt{(N-(2+M)s-1)\hat{\sigma}_j^2 \chi^2_{\alpha/2}; N-(2+M)s-1} \right].$$

(8)
Statistical Properties

The model’s estimated parameters are the result of maximum likelihood estimation. As such, they are consistent and asymptotically normally distributed.
Statistical Properties

The model’s estimated parameters are the result of maximum likelihood estimation. As such, they are consistent and asymptotically normally distributed.

Confidence intervals of level $\alpha$ for the parameters $\beta_j^k$:

$$\text{CI}_{\alpha}(\beta_j^k) = \left[ \hat{\beta}_j^k - t_{1-\alpha/2; N-(2+M)s_{\text{ASE}}(\hat{\beta}_j^k)}; \hat{\beta}_j^k + t_{1-\alpha/2; N-(2+M)s_{\text{ASE}}(\hat{\beta}_j^k)} \right].$$  

(7)

Confidence intervals of level $\alpha$ for the disturbance factor $\sigma_j$:

$$\text{CI}_{\alpha}(\sigma_j) = \left[ \sqrt{\frac{N-(2+M)s_{\text{ASE}}}{\chi_2^2(1-\alpha/2; N-(2+M)s_{\text{ASE}})}} \hat{\sigma}_j^2; \sqrt{\frac{N-(2+M)s_{\text{ASE}}}{\chi_2^2(\alpha/2; N-(2+M)s_{\text{ASE}})}} \hat{\sigma}_j^2 \right].$$  

(8)
Statistical Properties

The model’s estimated parameters are the result of maximum likelihood estimation. As such, they are consistent and asymptotically normally distributed.

Confidence intervals of level $\alpha$ for the parameters $\beta^j_k$:

$$CI_\alpha(\beta^j_k) = \left[ \hat{\beta}^j_k - t_{1-\alpha/2;N-(2+M)s} \text{ASE}(\hat{\beta}^j_k); \hat{\beta}^j_k + t_{1-\alpha/2;N-(2+M)s} \text{ASE}(\hat{\beta}^j_k) \right].$$

(7)
Statistical Properties

The model’s estimated parameters are the result of maximum likelihood estimation. As such, they are consistent and asymptotically normally distributed.

Confidence intervals of level $\alpha$ for the parameters $\beta^j_k$:

$$CI_\alpha(\beta^j_k) = \left[ \hat{\beta}^j_k - t_{1-\alpha/2;N-(2+M)s} ASE(\hat{\beta}^j_k); \hat{\beta}^j_k + t_{1-\alpha/2;N-(2+M)s} ASE(\hat{\beta}^j_k) \right].$$

Confidence intervals of level $\alpha$ for the disturbance factor $\sigma_j$:
Statistical Properties

The model’s estimated parameters are the result of maximum likelihood estimation. As such, they are consistent and asymptotically normally distributed.

Confidence intervals of level $\alpha$ for the parameters $\beta^j_k$:

$$CI_{\alpha}(\beta^j_k) = \left[ \hat{\beta}^j_k - t_{1-\alpha/2; N-(2+M)s} ASE(\hat{\beta}^j_k); \hat{\beta}^j_k + t_{1-\alpha/2; N-(2+M)s} ASE(\hat{\beta}^j_k) \right].$$

(7)

Confidence intervals of level $\alpha$ for the disturbance factor $\sigma_j$:

$$CI_{\alpha}(\sigma_j) = \left[ \sqrt{\frac{(N-(2+M)s-1)\hat{\sigma}_j^2}{\chi^2_{1-\alpha/2; N-(2+M)s-1}}}, \sqrt{\frac{(N-(2+M)s-1)\hat{\sigma}_j^2}{\chi^2_{\alpha/2; N-(2+M)s-1}}} \right].$$

(8)
Attention to multicollinearity issues!

We analyze the influence of the consumer price index (CPI) on the salary. CPI and time have a correlation of 0.995. Hence a model like

$$S_t = (\beta_0 + \gamma_0 z_t) + (\beta_1 + \gamma_1 z_t) t + (\beta_2 + \gamma_2 z_t) t^2 + (\beta_3 + \gamma_3 z_t) t^3,$$

(9)

where $S_t$ denotes the salary and $z_t$ is Luxembourg's CPI in year $t$ of the study, makes no sense. Because of obvious multicollinearity problems, almost none of the parameters would be significant. Therefore, we simplify the model and calibrate

$$S_t = (\beta_0 + \gamma_0 z_t) + \gamma_1 z_t t + \gamma_2 z_t t^2 + \gamma_3 z_t t^3.$$

(10)
Attention to multicolinearity issues!

We analyze the influence of the consumer price index (CPI) on the salary.
Attention to multicollinearity issues!

We analyze the influence of the consumer price index (CPI) on the salary. CPI and time have a correlation of 0.995.
Attention to multicolinearity issues!

We analyze the influence of the consumer price index (CPI) on the salary. CPI and time have a correlation of 0.995. Hence a model like

\[ S_{it} = (\beta_0^j + \gamma_0^j z_t) + (\beta_1^j + \gamma_1^j z_t) t + (\beta_2^j + \gamma_2^j z_t) t^2 + (\beta_3^j + \gamma_3^j z_t) t^3, \quad (9) \]

where \( S \) denotes the salary and \( z_t \) is Luxembourg’s CPI in year \( t \) of the study, makes no sense.

Because of obvious multicolinearity problems, almost none of the parameters would be significant. Therefore, we simplify the model and calibrate

\[ S_{it} = (\beta_0^j + \gamma_0^j z_t) + (\beta_1^j + \gamma_1^j z_t) t + (\beta_2^j + \gamma_2^j z_t) t^2, \quad (10) \]
Attention to multicolinearity issues!

We analyze the influence of the consumer price index (CPI) on the salary. CPI and time have a correlation of 0.995. Hence a model like

\[ S_{it} = (\beta_0^i + \gamma_0^i z_t) + (\beta_1^i + \gamma_1^i z_t)t + (\beta_2^i + \gamma_2^i z_t)t^2 + (\beta_3^i + \gamma_3^i z_t)t^3, \quad (9) \]

where \( S \) denotes the salary and \( z_t \) is Luxembourg’s CPI in year \( t \) of the study, makes no sense.

Because of obvious multicolinearity problems, almost none of the parameters would be significant.
Attention to multicolinearity issues!

We analyze the influence of the consumer price index (CPI) on the salary. CPI and time have a correlation of 0.995. Hence a model like

\[ S_{it} = (\beta_0^j + \gamma_0^j z_t) + (\beta_1^j + \gamma_1^j z_t) t + (\beta_2^j + \gamma_2^j z_t) t^2 + (\beta_3^j + \gamma_3^j z_t) t^3, \]  

(9)

where \( S \) denotes the salary and \( z_t \) is Luxembourg’s CPI in year \( t \) of the study, makes no sense.

Because of obvious multicolinearity problems, almost none of the parameters would be significant.

Therefore, we simplify the model and calibrate

\[ S_{it} = (\beta_0^j + \gamma_0^j z_t) + \gamma_1^j z_t t + \gamma_2^j z_t t^2 + \gamma_3^j z_t t^3. \]  

(10)
### Results for group 1

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Standard error</th>
<th>95% confidence intervals</th>
<th>intervals</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\beta_0)</td>
<td>321.381</td>
<td>1189.430</td>
<td>-2213.502</td>
<td>2856.093</td>
</tr>
<tr>
<td>(\gamma_0)</td>
<td>1689.492</td>
<td>277.834</td>
<td>-4.232</td>
<td>7.611</td>
</tr>
<tr>
<td>(\gamma_1)</td>
<td>0.400</td>
<td>0.120</td>
<td>0.143</td>
<td>0.656</td>
</tr>
<tr>
<td>(\gamma_2)</td>
<td>-0.034</td>
<td>0.007</td>
<td>-0.049</td>
<td>-0.019</td>
</tr>
<tr>
<td>(\gamma_3)</td>
<td>0.0008</td>
<td>0.0002</td>
<td>0.0005</td>
<td>0.0019</td>
</tr>
</tbody>
</table>

### Results for group 2

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Standard error</th>
<th>95% confidence intervals</th>
<th>intervals</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\beta_0)</td>
<td>7688.158</td>
<td>951.103</td>
<td>5660.197</td>
<td>9714.832</td>
</tr>
<tr>
<td>(\gamma_0)</td>
<td>-13.095</td>
<td>2.222</td>
<td>-17.822</td>
<td>-8.350</td>
</tr>
<tr>
<td>(\gamma_1)</td>
<td>1.260</td>
<td>0.096</td>
<td>1.055</td>
<td>1.465</td>
</tr>
<tr>
<td>(\gamma_2)</td>
<td>-0.097</td>
<td>0.006</td>
<td>-0.109</td>
<td>-0.085</td>
</tr>
<tr>
<td>(\gamma_3)</td>
<td>0.0025</td>
<td>0.0002</td>
<td>0.0022</td>
<td>0.0028</td>
</tr>
</tbody>
</table>

### Results for group 3

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Standard error</th>
<th>95% confidence intervals</th>
<th>intervals</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\beta_0)</td>
<td>682.638</td>
<td>196.327</td>
<td>141.924</td>
<td>1101.045</td>
</tr>
<tr>
<td>(\gamma_0)</td>
<td>-11.367</td>
<td>4.586</td>
<td>-21.135</td>
<td>-1.586</td>
</tr>
<tr>
<td>(\gamma_1)</td>
<td>0.983</td>
<td>0.199</td>
<td>0.559</td>
<td>1.406</td>
</tr>
<tr>
<td>(\gamma_2)</td>
<td>-0.048</td>
<td>0.012</td>
<td>-0.073</td>
<td>-0.023</td>
</tr>
<tr>
<td>(\gamma_3)</td>
<td>0.0010</td>
<td>0.0003</td>
<td>0.0003</td>
<td>0.0017</td>
</tr>
</tbody>
</table>
### Results for group 4

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Standard error</th>
<th>95% confidence intervals</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_0$</td>
<td>8473.081</td>
<td>1859.349</td>
<td>4511.016 - 12434.892</td>
</tr>
<tr>
<td>$\gamma_0$</td>
<td>-13.083</td>
<td>4.342</td>
<td>-22.335 - 3.825</td>
</tr>
<tr>
<td>$\gamma_1$</td>
<td>0.927</td>
<td>0.188</td>
<td>0.527 - 1.328</td>
</tr>
<tr>
<td>$\gamma_2$</td>
<td>-0.013</td>
<td>0.011</td>
<td>-0.036 - 0.010</td>
</tr>
<tr>
<td>$\gamma_3$</td>
<td>-0.0003</td>
<td>0.0003</td>
<td>-0.0009 - 0.0004</td>
</tr>
</tbody>
</table>

### Results for group 5

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Standard error</th>
<th>95% confidence intervals</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_0$</td>
<td>4798.276</td>
<td>3205.141</td>
<td>-2034.302 - 11630.238</td>
</tr>
<tr>
<td>$\gamma_0$</td>
<td>-2.846</td>
<td>7.488</td>
<td>-18.806 - 13.115</td>
</tr>
<tr>
<td>$\gamma_1$</td>
<td>1.315</td>
<td>0.324</td>
<td>0.0624 - 2.006</td>
</tr>
<tr>
<td>$\gamma_2$</td>
<td>-0.081</td>
<td>0.019</td>
<td>-0.122 - 0.040</td>
</tr>
<tr>
<td>$\gamma_3$</td>
<td>0.0016</td>
<td>0.0005</td>
<td>0.0005 - 0.0027</td>
</tr>
</tbody>
</table>

### Results for group 6

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Standard error</th>
<th>95% confidence intervals</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_0$</td>
<td>8332.439</td>
<td>1139.127</td>
<td>5903.348 - 10759.713</td>
</tr>
<tr>
<td>$\gamma_0$</td>
<td>-12.472</td>
<td>2.661</td>
<td>-18.145 - 6.800</td>
</tr>
<tr>
<td>$\gamma_1$</td>
<td>1.378</td>
<td>0.015</td>
<td>1.132 - 1.623</td>
</tr>
<tr>
<td>$\gamma_2$</td>
<td>-0.094</td>
<td>0.007</td>
<td>-0.108 - 0.079</td>
</tr>
<tr>
<td>$\gamma_3$</td>
<td>0.0022</td>
<td>0.0002</td>
<td>0.0018 - 0.0026</td>
</tr>
</tbody>
</table>
Disturbance terms

The disturbance terms for the six groups are $\sigma_1 = 41.49$, $\sigma_2 = 33.18$, $\sigma_3 = 68.48$, $\sigma_4 = 64.84$, $\sigma_5 = 111.83$ and $\sigma_6 = 39.74$
Bibliography