Nonlinear mixed-effects models using Stata

Yulia Marchenko

Executive Director of Statistics
StataCorp LLC

2018 UK Stata Conference
What is NLMEM?
Simple NLMEM
Residual covariance structures
Heteroskedasticity
Linear combinations and random coefficients
Three-level model: CES production function
Pharmacokinetic model
Nonlinear marginal models
Summary
References
Introduction to NLMEM

- Nonlinear mixed-effects models (NLMEMs)
  - mixed effects = fixed effects + random effects
- Nonlinear multilevel models
- Nonlinear hierarchical models
NLMEMs are popular in studies of biological and agricultural growth processes, population pharmacokinetics, bioassays, and more. For example, NLMEMs have been used to model drug absorption in the body, intensity of earthquakes, and growth of plants.
Nonlinear mixed-effects models

What is NLMEM?

Two ways of thinking: Nonlinear regression + REs

- Nonlinear regression:

\[
y = \frac{1}{\beta_1 + \beta_2 x + \beta_3 x^2} + \epsilon
\]

where \( \epsilon \sim N(0, \sigma^2) \).

- Let, e.g., \( \beta_1 \) vary randomly across \( G \) groups:

\[
\beta_1 = \beta_{1j} = b_1 + u_j, \quad j = 1, 2, \ldots, G
\]

where \( u_j \sim N(0, \sigma_u^2) \).

- Variance components: error variance \( \sigma^2 \) and between-group variance \( \sigma_u^2 \).

- Coefficients \( \beta_2 \) and \( \beta_3 \) can also be group-specific.
Alternatively, consider a linear mixed-effects model:

\[ y_{ij} = \beta_1 + \beta_2 x_{ij} + \beta_3 x_{ij}^2 + u_j + \epsilon_{ij} \]

where \( \epsilon_{ij} \sim N(0, \sigma^2) \) and \( u_j \sim N(0, \sigma^2_u) \).

In the nonlinear mixed-effects model

\[ y_{ij} = \frac{1}{\beta_1 + \beta_2 x_{ij} + \beta_3 x_{ij}^2 + u_j} + \epsilon_{ij} \]

all coefficients and random intercept \( u_j \) enter nonlinearly.
Simple NLMEM: Growth of orange trees

```
webuse orange  
(Growth of orange trees (Draper and Smith, 1998))
twoway connected circumf age, connect(L) title(Growth of orange trees)
```

![Growth of orange trees graph]

- Trunk circumference (mm) vs. Time since Dec 31, 1968 (days)
Consider the following nonlinear growth model:

\[
\text{circumf}_{ij} = \frac{\beta_1}{1 + \exp \left\{ - \left( \text{age}_{ij} - \beta_2 \right) / \beta_3 \right\}} + \epsilon_{ij}
\]

where \( \epsilon_{ij} \sim N(0, \sigma^2) \).

- \( \beta_1 \) is the average asymptotic trunk circumference of trees as \( \text{age} \to \infty \).
- \( \beta_2 \) estimates the age at which a tree attains half of \( \beta_1 \).
- \( \beta_3 \) represents the number of days it takes for a tree to grow from 50\% to about 73\% of its average asymptotic trunk circumference \( \beta_1 \).
$\beta_1 \approx 175$ mm, $\beta_2 \approx 700$ days, and $\beta_3 \approx 1,000 - 700 = 300$ days.

Notice that the variability between trees increases with age.
Let’s incorporate the between-tree variability into the model.

Consider the following two-level nonlinear growth model (Pinheiro and Bates 2000):

\[
\text{circumf}_{ij} = \frac{\beta_1 + u_{1j}}{1 + \exp\left\{- \left(\text{age}_{ij} - \beta_2\right) / \beta_3\right\}} + \epsilon_{ij}
\]

where \( u_{1j} \sim N(0, \sigma_{u_1}^2) \) and \( \epsilon_{ij} \sim N(0, \sigma^2) \).
We use `menl` to fit the model.

```
    . menl circumf = ({b1}+{U1[tree]})/(1+exp(-(age-{b2})/{b3}))
```

Mixed-effects ML nonlinear regression

Group variable: tree

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Std. Err.</th>
<th>[95% Conf. Interval]</th>
</tr>
</thead>
<tbody>
<tr>
<td>/b1</td>
<td>191.049</td>
<td>16.15403</td>
<td>159.3877 222.7103</td>
</tr>
<tr>
<td>/b2</td>
<td>722.556</td>
<td>35.15082</td>
<td>653.6616 791.4503</td>
</tr>
<tr>
<td>/b3</td>
<td>344.1624</td>
<td>27.14739</td>
<td>290.9545 397.3703</td>
</tr>
</tbody>
</table>

Obs per group:

- min = 7
- avg = 7.0
- max = 7

Mixed-effects ML nonlinear regression

Number of obs = 35
Number of groups = 5

Obs per group:

- min = 7
- avg = 7.0
- max = 7

Linearization log likelihood = -131.58458

Random-effects Parameters

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Std. Err.</th>
<th>[95% Conf. Interval]</th>
</tr>
</thead>
<tbody>
<tr>
<td>tree: Identity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>var(U1)</td>
<td>991.1514</td>
<td>639.4636</td>
<td>279.8776 3510.038</td>
</tr>
<tr>
<td>var(Residual)</td>
<td>61.56371</td>
<td>15.89568</td>
<td>37.11466 102.1184</td>
</tr>
</tbody>
</table>
Similarly, we can let $\beta_2$ and $\beta_3$ vary across trees.

We use a more convenient multistage formulation:

$$\text{circumf}_{ij} = \frac{\beta_{1j}}{1 + \exp \left\{ - \frac{\text{age}_{ij} - \beta_{2j}}{\beta_{3j}} \right\}} + \epsilon_{ij}$$

where

$$\beta_{1j} = b_1 + u_{1j}$$

$$\beta_{2j} = b_2 + u_{2j}$$

$$\beta_{3j} = b_3 + u_{3j}$$

and where $u_{1j} \sim N(0, \sigma_{u_1}^2)$, $u_{2j} \sim N(0, \sigma_{u_2}^2)$ and $u_{3j} \sim N(0, \sigma_{u_3}^2)$. 
. menl circumf = ({beta1:})/(1+exp(-(age-{beta2:})/{beta3:})),
> define(beta1:{b1}+{U1[tree]})
> define(beta2:{b2}+{U2[tree]})
> define(beta3:{b3}+{U3[tree]})

Mixed-effects ML nonlinear regression
Number of obs = 35
Group variable: tree
Number of groups = 5

Obs per group:
  min = 7
  avg = 7.0
  max = 7

Linearization log likelihood = -131.55076

beta1: {b1}+{U1[tree]}
beta2: {b2}+{U2[tree]}
beta3: {b3}+{U3[tree]}

|    | Coef.  | Std. Err.  |    z   | P>|z|  | [95% Conf. Interval] |
|----|--------|------------|--------|------|----------------------|
| /b1| 191.1332| 15.96228   | 11.97  | 0.000| 159.8477              | 222.4187   |
| /b2| 722.7144| 34.94627   | 20.68  | 0.000| 654.2209              | 791.2078   |
| /b3| 345.2863| 27.70935   | 12.46  | 0.000| 290.977              | 399.5956   |

Random-effects Parameters

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Std. Err.</th>
<th>[95% Conf. Interval]</th>
</tr>
</thead>
<tbody>
<tr>
<td>tree: Independent</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
<pre><code>| var(U1)   | 970.67 | 665.4967 | 253.2113 | 3721.004 |
| var(U2)   | 140.9707 | 2669.433 | 1.07e-14 | 1.85e+18 |
| var(U3)   | 248.5962 | 1397.996 | .0040617 | 1.52e+07 |
</code></pre>

    | var(Residual) | 59.43549 | 18.44102 | 32.35519 | 109.1812 |
- With only five trees, the previous model is already too rich for these data.

- Otherwise, we could have considered a more complicated covariance structure for the random effects:

\[
(u_{1j}, u_{2j}, u_{3j}) \sim \text{MVN}(0, \Sigma), \quad \Sigma = \begin{bmatrix}
\sigma_{11} & \sigma_{12} & \sigma_{13} \\
\sigma_{12} & \sigma_{22} & \sigma_{23} \\
\sigma_{13} & \sigma_{23} & \sigma_{33}
\end{bmatrix}
\]

- Or assuming dependence only between some random effects such as \(u_{1j}\) and \(u_{2j}\):

\[
\Sigma = \begin{bmatrix}
\sigma_{11} & \sigma_{12} & 0 \\
\sigma_{12} & \sigma_{22} & 0 \\
0 & 0 & \sigma_{33}
\end{bmatrix}
\]

- And variations of the above.
For example,

```
. menl circumf = ({beta1:})/(1+exp(-(age-{beta2:})/{beta3:})),
    > define(beta1:{b1}+{U1[tree]})
    > define(beta2:{b2}+{U2[tree]})
    > define(beta3:{b3}+{U3[tree]})
    > covariance(U1 U2 U3, unstructured)
```

The above is also equivalent to:

```
. menl ..., ... covariance(U*, unstructured)
```

Or, assuming correlation only between U1 and U2

```
. menl ..., ... covariance(U1 U2, unstructured)
```
menl provides flexible modeling of within-group error structures (or residual covariance structures).

Use option `resvariance()` to model error heteroskedasticity as a linear, power, or exponential function of other covariates or of predicted values.

Use option `rescorrelation()` to model the dependence of the within-group errors as, e.g., AR or MA processes.

Combine `resvariance()` and `rescorrelation()` to build flexible residual covariance structures.
Continuing with growth processes, consider the growth of soybean plants.

Variable weight records an average leaf weight per plant in grams.

Variable time records the number of days after planting at which plants were weighed.

The data are obtained from 48 plots.

```
.webuse soybean
(Growth of soybean plants (Davidian and Giltinan, 1995))
```
Consider the following growth model:

\[
weight_{ij} = \frac{\phi_{1j}}{1 + \exp \left\{ - \left( \text{time}_{ij} - \phi_{2j} \right) / \phi_{3j} \right\}} + \epsilon_{ij}
\]

where

\[
\begin{align*}
\phi_{1j} &= b_1 + u_{1j} \\
\phi_{2j} &= b_2 + u_{2j} \\
\phi_{3j} &= b_3 + u_{3j}
\end{align*}
\]

and where \((u_{1j}, u_{2j}, u_{3j}) \sim \text{MVN}(0, \Sigma)\) with

\[
\Sigma = \begin{bmatrix}
\sigma_{11} & \sigma_{12} & \sigma_{13} \\
\sigma_{12} & \sigma_{22} & \sigma_{23} \\
\sigma_{13} & \sigma_{23} & \sigma_{33}
\end{bmatrix}
\]

and \(\epsilon_{ij} \sim N(0, \sigma^2)\).
We use the following specification of `menl`:

```
. menl weight = {phi1:}/(1+exp(-(time-{phi2:})/{phi3:})),
> define(phi1: U1[plot], xb)
> define(phi2: U2[plot], xb)
> define(phi3: U3[plot], xb)
> covariance(U1 U2 U3, unstructured)
```

Option

```
define(phi1: U1[plot], xb)
```

is essentially a shortcut for

```
define(phi1: {b1}+{U1[plot]})
```

The above shortcut is useful to specify linear combinations.
Estimates of regression coefficients:

mixed-effects ML nonlinear regression  Number of obs = 412
Group variable: plot  Number of groups = 48

Obs per group:
    min = 8
    avg = 8.6
    max = 10

Linearization log likelihood = -739.83445

<table>
<thead>
<tr>
<th>phi1: U1[plot], xb</th>
<th>phi2: U2[plot], xb</th>
<th>phi3: U3[plot], xb</th>
</tr>
</thead>
</table>

| weight | Coef.  | Std. Err.  | z     | P>|z|  | [95% Conf. Interval] |
|--------|--------|------------|-------|------|---------------------|
| phi1   |        |            |       |      |                     |
| _cons  | 19.25314 | 0.8031811  | 23.97 | 0.000 | 17.67893  20.82734 |
| phi2   |        |            |       |      |                     |
| _cons  | 55.01999 | 0.7272491  | 75.65 | 0.000 | 53.59461  56.44537 |
| phi3   |        |            |       |      |                     |
| _cons  | 8.403468 | 0.3152551  | 26.66 | 0.000 | 7.78558  9.021357  |
Estimates of variance components:

<table>
<thead>
<tr>
<th>Random-effects Parameters</th>
<th>Estimate</th>
<th>Std. Err.</th>
<th>[95% Conf. Interval]</th>
</tr>
</thead>
<tbody>
<tr>
<td>plot: Unstructured</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>var(U1)</td>
<td>27.05081</td>
<td>6.776516</td>
<td>16.55561 44.19929</td>
</tr>
<tr>
<td>var(U2)</td>
<td>17.61605</td>
<td>5.317899</td>
<td>9.748766 31.83227</td>
</tr>
<tr>
<td>var(U3)</td>
<td>1.972036</td>
<td>0.9849825</td>
<td>0.7409021 5.248904</td>
</tr>
<tr>
<td>cov(U1,U2)</td>
<td>15.73304</td>
<td>5.413365</td>
<td>5.123042 26.34304</td>
</tr>
<tr>
<td>cov(U1,U3)</td>
<td>5.193819</td>
<td>2.165586</td>
<td>0.9493488 9.438289</td>
</tr>
<tr>
<td>cov(U2,U3)</td>
<td>5.649306</td>
<td>2.049458</td>
<td>1.632442 9.66617</td>
</tr>
<tr>
<td>var(Residual)</td>
<td>1.262237</td>
<td>0.1111686</td>
<td>1.062119 1.50006</td>
</tr>
</tbody>
</table>

Store estimation results for later comparison

.estimates store nohet
Residuals-versus-fitted plot

. predict fitweight, yhat
. predict res, residuals
. scatter res fitweight
Davidian and Giltinan (1995) proposed to model heteroskedasticity (the error variance) in this example as a power function of the mean:

$$\text{Var}(\epsilon_{ij}) = \sigma^2(\hat{\text{weight}}_{ij})^{2\delta}$$

where $\hat{\text{weight}}_{ij}$ denotes predicted mean weight values.

The corresponding \texttt{menl} specification is

```stata
. menl weight = {phi1:}/(1+exp(-(time-{phi2:})/{phi3:})), > define(phi1: U1[plot], xb) > define(phi2: U2[plot], xb) > define(phi3: U3[plot], xb) > covariance(U1 U2 U3, unstructured) > resvariance(power _yhat, noconstant)
```
Nonlinear mixed-effects models

Heteroskedasticity

menl, resvar(power): Regression coefficients

Estimates of regression coefficients:

Mixed-effects ML nonlinear regression
Group variable: plot

Number of obs = 412
Number of groups = 48
Obs per group:
min = 8
avg = 8.6
max = 10

Linearization log likelihood = -357.49994

phi1: U1[plot], xb
phi2: U2[plot], xb
phi3: U3[plot], xb

| weight | Coef.  | Std. Err. | z      | P>|z|  | [95% Conf. Interval] |
|--------|--------|-----------|--------|------|----------------------|
| phi1  | _cons | 16.9453   | .6097869 | 27.79 | 0.000 | 15.75014 18.14047 |
| phi2  | _cons | 51.77782  | .4625946 | 111.93 | 0.000 | 50.87115 52.68449 |
| phi3  | _cons | 7.542116  | .0967274 | 77.97  | 0.000 | 7.352534 7.731699 |
Estimates of variance components:

<table>
<thead>
<tr>
<th>plot: Unstructured</th>
<th>var(U1)</th>
<th>var(U2)</th>
<th>var(U3)</th>
<th>cov(U1,U2)</th>
<th>cov(U1,U3)</th>
<th>cov(U2,U3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>11.69387</td>
<td>2.782124</td>
<td>7.335783</td>
<td>.7796133</td>
<td>.9554937</td>
<td>.3484718</td>
</tr>
<tr>
<td></td>
<td>3.023475</td>
<td>1.293342</td>
<td>1.307347</td>
<td>.1937439</td>
<td>.2547011</td>
<td>.1110401</td>
</tr>
<tr>
<td></td>
<td>.105363</td>
<td>.0447326</td>
<td>.0458463</td>
<td>.3998823</td>
<td>.4562887</td>
<td>.1308373</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.3484718</td>
<td>.1308373</td>
<td></td>
</tr>
</tbody>
</table>

Residual variance:

<table>
<thead>
<tr>
<th>Power _yhat</th>
<th>sigma2</th>
<th>delta</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>.0496825</td>
<td>.9371342</td>
</tr>
<tr>
<td></td>
<td>.0043247</td>
<td>.0253618</td>
</tr>
<tr>
<td></td>
<td>.0418899</td>
<td>.8874261</td>
</tr>
<tr>
<td></td>
<td>.0589247</td>
<td>.9868424</td>
</tr>
</tbody>
</table>

Store estimation results for comparison

`. estimates store het`
Nonlinear mixed-effects models

Heteroskedasticity

Model comparison

- **Likelihood-ratio test:**

  `. lrtest het nohet`

  Likelihood-ratio test
  (Assumption: nohet nested in het)

  LR chi2(1) = 764.67
  Prob > chi2 = 0.0000

- **Information criteria:**

  `. estimates stats het nohet`

  Akaike’s information criterion and Bayesian information criterion

<table>
<thead>
<tr>
<th>Model</th>
<th>Obs</th>
<th>ll(null)</th>
<th>ll(model)</th>
<th>df</th>
<th>AIC</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>het</td>
<td>412</td>
<td>.</td>
<td>-357.4999</td>
<td>11</td>
<td>736.9999</td>
<td>781.2311</td>
</tr>
<tr>
<td>nohet</td>
<td>412</td>
<td>.</td>
<td>-739.8344</td>
<td>10</td>
<td>1499.669</td>
<td>1539.879</td>
</tr>
</tbody>
</table>

  Note: N=Obs used in calculating BIC; see [R] BIC note.

- A heteroskedastic model fits data better.
The actual objective of the soybean study was to compare the growth patterns of two genotypes of soybean plants in three types of growing seasons.

- Genotypes, variety: commercial variety F and experimental variety P

Growth pattern in wet growing season 1989
We can include the main effects of genotypes and of years, and their interaction in the equation for the asymptotic rate:

\[ \phi_{1j} = b_1 + \beta_G^\top G + \beta_Y^\top Y + \cdots + u_{1j} \]

**menl specification:**

```plaintext
. menl weight = {phi1:}/(1+exp(-(time-{phi2:})/{phi3:})),
>   define(phi1: i.variety##i.year U1[plot])
>   define(phi2: U2[plot], xb)
>   define(phi3: U3[plot], xb)
>   covariance(U1 U2 U3, unstructured)
>   resvariance(power _yhat, noconstant)
```
We can also let the coefficients for, e.g., genotypes vary across plots:

$$\phi_{1j} = b_1 + \beta_G^\top G + \beta_Y^\top Y + \cdots + u_{1j} + F \times f_{1j} + P \times p_{1j}$$

where $F$ and $P$ are genotype indicators and $f_{1j} \sim N(0, \sigma^2_F)$ and $p_{1j} \sim N(0, \sigma^2_P)$.

**menl** specification:

```stata
. menl weight = {phi1:}/(1+exp(-(time-{phi2:})/{phi3:})),
> define(phi1: i.variety##i.year U1[plot] 1.variety#F1[plot] 2.variety#P1[plot])
> define(phi2: U2[plot], xb)
> define(phi3: U3[plot], xb)
> covariance(U1 U2 U3, unstructured)
> resvariance(power _yhat, noconstant)
```

The *i.* operator is not allowed with factor variables when specifying random coefficients because a distinct name is required for each random coefficient.
Constant elasticity of substitution (CES) production function is used in macroeconomic modeling to model the production process as a function of inputs such as capital and labor. It introduces and estimates the CES parameter, which makes it a flexible modeling tool.

Elasticity of substitution (ES) measures how easy it is to substitute one input such as capital for another such as labor. And constant ES does not depend on input values.

Other common production functions such as Cobb-Douglas and Leontief can be viewed as special cases of the CES production function. For example, Cobb-Douglas function assumes that ES is 1.
Consider fictional data on log(production) from the 50 U.S. states plus D.C. divided into 9 regions over the period of 1990 to 2017.

We wish to fit the CES production function

$$\ln P_{ijt} = \beta_0 - \frac{1}{\rho} \ln \{\delta K_{ijt} - \rho + (1 - \delta)L_{ijt}\} + \epsilon_{ijt}$$

where $\epsilon_{ijt} \sim N(0, \sigma^2)$.

In $P_{ijt}$, $K_{ijt}$, and $L_{ijt}$ are log(production), capital, and labor of state $j$ within region $i$ in year $t$.

Parameters: log-factor productivity $\beta_0$, share $\delta$, and $\rho$ is related to the elasticity of substitution $\sigma = 1/(1 + \rho)$. 
We suspect that $\delta$ may be affected by regions and states-within-regions:

$$\delta = \delta_{ij} = \delta_0 + u_{1i} + u_{2ij}$$

where $u_{1i} \sim N(0, \sigma^2_{u_1})$ and $u_{2ij} \sim N(0, \sigma^2_{u_2})$. $u_2$’s are nested within $u_1$’s.
Nonlinear mixed-effects models

Three-level model: CES production function

```
. menl lnprod = {b0}-1/{rho}*ln({delta:}*capital^(-{rho})+(1-{delta:})*labor^(-{rho})),
> define(delta: {delta0} + {U1[region]} + {U2[region>state]})
```

Mixed-effects ML nonlinear regression

```
Number of obs = 1,377
```

<table>
<thead>
<tr>
<th>Path</th>
<th>No. of Groups</th>
<th>Observations per Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Minimum</td>
<td>Average</td>
</tr>
<tr>
<td>region</td>
<td>9</td>
<td>108</td>
</tr>
<tr>
<td>region&gt;state</td>
<td>51</td>
<td>27</td>
</tr>
</tbody>
</table>

Linearization log likelihood = 1094.2223

delta: {delta0}+{U1[region]}+{U2[region>state]}

| lnprod  | Coef.  | Std. Err. | z       | P>|z|   | [95% Conf. Interval] |
|---------|--------|-----------|---------|-------|----------------------|
| /b0     | 3.49166| 0.0040189 | 868.82  | 0.000 | 3.483783 3.499537    |
| /delta0 | .3439896| 0.0490629 | 7.01    | 0.000 | .2478281 .4401511   |
| /rho    | 1.109318| 0.0272072 | 40.77   | 0.000 | 1.055993 1.162644   |
There is some variability between regions and states-within-regions in the estimates of the share parameter.
Nonlinear mixed-effects models

Three-level model: CES production function

Predict region-specific share parameters

We can predict the share parameter for each region:

```
predict (delta = {delta:}), relevel(region)
list region delta if region[_n]!=region[_n-1], sep(0)
```

<table>
<thead>
<tr>
<th>region</th>
<th>delta</th>
</tr>
</thead>
<tbody>
<tr>
<td>New England</td>
<td>0.2699136</td>
</tr>
<tr>
<td>Mid Atlantic</td>
<td>0.1453616</td>
</tr>
<tr>
<td>E North Central</td>
<td>0.6366224</td>
</tr>
<tr>
<td>W North Central</td>
<td>0.3761043</td>
</tr>
<tr>
<td>South Atlantic</td>
<td>0.3879336</td>
</tr>
<tr>
<td>E South Central</td>
<td>0.344411</td>
</tr>
<tr>
<td>W South Central</td>
<td>0.17091</td>
</tr>
<tr>
<td>Mountain</td>
<td>0.4102365</td>
</tr>
<tr>
<td>Pacific</td>
<td>0.3544133</td>
</tr>
</tbody>
</table>
We can use `nlcom` to estimate the ES:

```
. nlcom (sigma: 1/(1+_b[/rho]))

sigma: 1/(1+_b[/rho])
```

| lnprod  | Coef.    | Std. Err. | z      | P>|z|  | [95% Conf. Interval] |
|---------|----------|-----------|--------|------|----------------------|
| sigma   | .4740868 | .006115   | 77.53  | 0.000| .4621015   .4860721 |

The estimated ES is 0.47, which is less than one meaning that the capital and labor are not very good substitutes in this example. If the labor price increases, substituting capital for labor will not offset the increase in the total expenditure on labor.
Pharmacokinetic model

- Pharmacokinetics (PKs) studies the distribution of drugs within the body and is often referred to as the study of “what the body does with a drug”.

- It models drug output based on drug input by summarizing concentration-time measurements, while accounting for patient-specific characteristics.
Consider data on the antiasthmatic agent theophylline (Boeckmann, Sheiner, and Beal [1994] 2011).

The drug was administered orally to 12 subjects with the dosage (mg/kg) given on a per weight basis.

Serum concentrations (in mg/L) were obtained at 11 time points per subject over 25 hours following administration.
Concentration-time profiles of 12 subjects:

- `webuse theoph`
  (Theophylline kinetics (Boeckmann et al., [1994] 2011))
- `twoway connected conc time, connect(L)`
The concentration rises rapidly initially and then decays exponentially.

In PKs, such pattern is often described by a so-called one-compartment open model with first-order absorption and elimination. (Body is viewed as one “blood compartment”.)

This model is used for drugs that distribute relatively rapidly throughout the body, which is a reasonable assumption for the kinetics of theophylline after oral administration.
One-compartment open model for theophylline kinetics:

$$\text{conc}_{ij} = \frac{\text{dose}_j k_e k_{aj}}{\text{Cl}_j (k_{aj} - k_e)} \left\{ \exp (-k_e \text{time}_{ij}) - \exp (-k_{aj} \text{time}_{ij}) \right\} + \epsilon_{ij}$$

for $i = 1, \ldots, 11$ and $j = 1, \ldots, 12$.

Parameters: elimination rate constant $k_e$, and, for each subject $j$, absorption rate constant $k_{aj}$ and clearance $\text{Cl}_j$. 
Elimination rate constant describes the rate at which a drug is removed from the body. It is defined as the fraction of drug in the body eliminated per unit time.

Absorption rate constant describes the rate at which a drug is absorbed by the body.

Clearance measures the rate at which a drug is cleared from the plasma. It is defined as the volume of plasma cleared of drug per unit time.
All parameters must be positive, and clearance and absorption rate constant are allowed to vary among subjects:

\[
\begin{align*}
Cl_j &= \exp(\beta_0 + u_{0j}) \\
k_a_j &= \exp(\beta_1 + u_{1j}) \\
k_e &= \exp(\beta_2)
\end{align*}
\]

where \( u_{0j} \sim N(0, \sigma_{u_0}^2) \) and \( u_{1j} \sim N(0, \sigma_{u_1}^2) \).

Heteroskedasticity, often present in PK data, is modeled using the power function plus a constant.

\[
\text{Var}(\varepsilon_{ij}) = \sigma^2 \{(\hat{\text{conc}}_{ij})^\delta + c\}^2
\]

Adding a constant avoids the variance of zero at time \( t = 0 \), because the concentration is zero at that time.
Nonlinear mixed-effects models

Pharmacokinetic model

\begin{verbatim}
menl conc = (dose*{ke:}*{ka:}/({cl:}*({ka:}-{ke:}))*\exp(-{ke:}*time)-\exp(-{ka:}*time)),
> define(cl: \exp({b0}+{U0[subject]}))
> define(ka: \exp({b1}+{U1[subject]}))
> define(ke: \exp({b2}))
> resvariance(power _yhat) reml

Mixed-effects REML nonlinear regression
Number of obs = 132
Group variable: subject
Number of groups = 12
Obs per group:
  min = 11
  avg = 11.0
  max = 11

Linear. log restricted-likelihood = -172.44384

  cl: \exp({b0}+{U0[subject]})
  ka: \exp({b1}+{U1[subject]})
  ke: \exp({b2})
\end{verbatim}

Yulia Marchenko (StataCorp) 44 / 49

---

\begin{tabular}{l|cccccc}
  \hline
  conc & Coef. & Std. Err. & z & P>|z| & [95\% Conf. Interval] \\
  \hline
  /b0  & -3.227295 & .0619113 & -52.13 & 0.000 & -3.348639 & -3.105951 \\
  /b1  & .4354519 & .2072387 & 2.10 & 0.036 & .0292716 & .8416322 \\
  /b2  & -2.453743 & .0517991 & -47.37 & 0.000 & -2.555267 & -2.352218 \\
  \hline
\end{tabular}
### Nonlinear mixed-effects models

#### Pharmacokinetic model

#### `menl`: Variance components

<table>
<thead>
<tr>
<th>Random-effects Parameters</th>
<th>Estimate</th>
<th>Std. Err.</th>
<th>[95% Conf. Interval]</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>subject: Independent</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><code>var(U0)</code></td>
<td>0.0316416</td>
<td>0.014531</td>
<td>0.0128634 0.0778326</td>
</tr>
<tr>
<td><code>var(U1)</code></td>
<td>0.4500585</td>
<td>0.2228206</td>
<td>0.1705476 1.187661</td>
</tr>
<tr>
<td><strong>Residual variance: Power _yhat</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><code>sigma2</code></td>
<td>0.1015759</td>
<td>0.086535</td>
<td>0.0191263 0.5394491</td>
</tr>
<tr>
<td><code>delta</code></td>
<td>0.3106636</td>
<td>0.2466547</td>
<td>-0.1727707 0.7940979</td>
</tr>
<tr>
<td><code>_cons</code></td>
<td>0.7150935</td>
<td>0.3745256</td>
<td>0.2561837 1.996063</td>
</tr>
</tbody>
</table>
In the previous `menl` model, we used restricted maximum likelihood estimation (REML) via option `reml` instead of the default maximum likelihood (ML) estimation to account for a moderate number of subjects.

We specified **nonlinear** functions of model parameters in the `define()` options.
As of update 06nov2017 to Stata 15, you can also use `menl` to fit nonlinear marginal models.

These models do not introduce random effects but instead model the within-group error covariance structure directly.

See suboption `group()` available within options `resvariance()` and `rescorrelation()` and example 17 in `[ME] menl`.

You can think of these models as nonlinear models which relax the assumption of i.i.d. errors.
menl fits NLMEMs; see \[\text{[ME] menl}\].

menl implements the Lindstrom–Bates method, which is based on the linearization of the nonlinear mean function with respect to fixed and random effects.

menl supports ML and REML estimation and provides flexible random-effects and residual covariance structures.

menl supports single-stage and multistage specifications.

You can predict random effects and their standard errors, group-specific nonlinear parameters, and more after estimation; see \[\text{[ME] menl postestimation}\].

NLMEMs are known to be sensitive to initial values. menl provides default, but for some models you may need to specify your own. Use option initial().

NLMEMs are known to have difficulty converging or converging to a local maximum. Trying different initial values may help.
References

