

# THE STATA JOURNAL

## Editor

H. Joseph Newton  
Department of Statistics  
Texas A & M University  
College Station, Texas 77843  
979-845-3142; FAX 979-845-3144  
jnewton@stata-journal.com

## Executive Editor

Nicholas J. Cox  
Department of Geography  
University of Durham  
South Road  
Durham City DH1 3LE UK  
n.j.cox@stata-journal.com

## Associate Editors

Christopher Baum  
Boston College  
Rino Bellocco  
Karolinska Institutet  
David Clayton  
Cambridge Inst. for Medical Research  
Mario A. Cleves  
Univ. of Arkansas for Medical Sciences  
Charles Franklin  
University of Wisconsin, Madison  
Joanne M. Garrett  
University of North Carolina  
Allan Gregory  
Queen's University  
James Hardin  
Texas A&M University  
Stephen Jenkins  
University of Essex  
Jens Lauritsen  
Odense University Hospital  
Stanley Lemeshow  
Ohio State University

J. Scott Long  
Indiana University  
Thomas Lumley  
University of Washington, Seattle  
Roger Newson  
King's College, London  
Marcello Pagano  
Harvard School of Public Health  
Sophia Rabe-Hesketh  
Inst. of Psychiatry, King's College London  
J. Patrick Royston  
MRC Clinical Trials Unit, London  
Philip Ryan  
University of Adelaide  
Mark E. Schaffer  
Heriot-Watt University, Edinburgh  
Jeroen Weesie  
Utrecht University  
Jeffrey Wooldridge  
Michigan State University

**Copyright Statement:** The Stata Journal and the contents of the supporting files (programs, datasets, and help files) are copyright © by Stata Corporation. The contents of the supporting files (programs, datasets, and help files) may be copied or reproduced by any means whatsoever, in whole or in part, as long as any copy or reproduction includes attribution to both (1) the author and (2) the Stata Journal.

The articles appearing in the Stata Journal may be copied or reproduced as printed copies, in whole or in part, as long as any copy or reproduction includes attribution to both (1) the author and (2) the Stata Journal.

Written permission must be obtained from Stata Corporation if you wish to make electronic copies of the insertions. This precludes placing electronic copies of the Stata Journal, in whole or in part, on publicly accessible web sites, file servers, or other locations where the copy may be accessed by anyone other than the subscriber.

Users of any of the software, ideas, data, or other materials published in the Stata Journal or the supporting files understand that such use is made without warranty of any kind, by either the Stata Journal, the author, or Stata Corporation. In particular, there is no warranty of fitness of purpose or merchantability, nor for special, incidental, or consequential damages such as loss of profits. The purpose of the Stata Journal is to promote free communication among Stata users.

The *Stata Technical Journal* (ISSN 1536-867X) is a publication of Stata Press, and Stata is a registered trademark of Stata Corporation.

**The Stata Journal** publishes reviewed papers together with shorter notes or comments, regular columns, book reviews, and other material of interest to Stata users. Examples of the types of papers include 1) expository papers that link the use of Stata commands or programs to associated principles, such as those that will serve as tutorials for users first encountering a new field of statistics or a major new technique; 2) papers that go “beyond the Stata manual” in explaining key features or uses of Stata that are of interest to intermediate or advanced users of Stata; 3) papers that discuss new commands or Stata programs of interest either to a wide spectrum of users (e.g., in data management or graphics) or to some large segment of Stata users (e.g., in survey statistics, survival analysis, panel analysis, or limited dependent variable modeling); 4) papers analyzing the statistical properties of new or existing estimators and tests in Stata; 5) papers that could be of interest or usefulness to researchers, especially in fields that are of practical importance but are not often included in texts or other journals, such as the use of Stata in managing datasets, especially large datasets, with advice from hard-won experience; and 6) papers of interest to those teaching, including Stata with topics such as extended examples of techniques and interpretation of results, simulations of statistical concepts, and overviews of subject areas.

For more information on the Stata Journal, including information for authors, see the web page

<http://www.stata-journal.com>

**Subscriptions** are available from Stata Corporation, 4905 Lakeway Drive, College Station, Texas 77845, telephone 979-696-4600 or 800-STATA-PC, fax 979-696-4601, or online at

<http://www.stata.com/bookstore/sj.html>

#### **Subscription rates:**

Subscriptions mailed to US and Canadian addresses:

3-year subscription (includes printed and electronic copy)	\$153
2-year subscription (includes printed and electronic copy)	\$110
1-year subscription (includes printed and electronic copy)	\$ 59

Subscriptions mailed to other countries:

3-year subscription (includes printed and electronic copy)	\$225
2-year subscription (includes printed and electronic copy)	\$158
1-year subscription (includes printed and electronic copy)	\$ 83
3-year subscription (electronic only)	\$153

Back issues of the Stata Journal may be ordered online at

<http://www.stata.com/bookstore/sj.html>

The Stata Journal is published quarterly by the Stata Press, College Station, Texas, USA.

Address changes should be sent to The Stata Journal, Stata Corporation, 4905 Lakeway Drive, College Station TX 77845, USA, or email [sj@stata.com](mailto:sj@stata.com).

# From the help desk: Swamy's random-coefficients model

Brian P. Poi  
Stata Corporation

**Abstract.** This article discusses the Swamy (1970) random-coefficients model and presents a command that extends Stata's `xtrchh` command by also providing estimates of the panel-specific coefficients.

**Keywords:** st0046, panel data, random-coefficients models

## 1 Introduction

Fixed- and random-effects models incorporate panel-specific heterogeneity by including a set of nuisance parameters that essentially provide each panel with its own constant term. However, all panels share common slope parameters. Random-coefficients models are more general in that they allow each panel to have its own vector of slopes randomly drawn from a distribution common to all panels. Stata's `xtrchh` command provides estimates of the parameters characterizing the distribution from which the panel-specific parameters are drawn. The command included with this article extends Stata's implementation by also providing best linear unbiased predictors of the panel-specific draws from that distribution.

Section 2 develops the Swamy (1970) random-coefficients model. Section 3 then presents the syntax and usage of a command called `xtrchh2` that implements the estimator in Stata, and section 4 presents an example. Section 5 lists the results stored by `xtrchh2`.

## 2 Swamy's random-coefficients model

Following Swamy (1970), consider a random-coefficients model of the form

$$\mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta}_i + \boldsymbol{\epsilon}_i \quad (1)$$

where  $i = 1 \dots P$  denotes panels,  $\mathbf{y}_i$  is a  $T_i \times 1$  vector of observations for the  $i$ th panel,  $\mathbf{X}_i$  is a  $T_i \times k$  matrix of nonstochastic covariates, and  $\boldsymbol{\beta}_i$  is a  $k \times 1$  vector of parameters specific to panel  $i$ . The error term vector  $\boldsymbol{\epsilon}_i$  is distributed with mean zero and variance  $\sigma_{ii}\mathbf{I}$ . The panels do not need to be balanced.

Each panel-specific  $\boldsymbol{\beta}_i$  is related to an underlying common parameter vector  $\boldsymbol{\beta}$ :

$$\boldsymbol{\beta}_i = \boldsymbol{\beta} + \mathbf{v}_i \quad (2)$$

where  $E\{\mathbf{v}_i\} = \mathbf{0}$ ,  $E\{\mathbf{v}_i\mathbf{v}_i'\} = \Sigma$ ,  $E\{\mathbf{v}_i\mathbf{v}_j'\} = \mathbf{0}$  for  $j \neq i$ , and  $E\{\mathbf{v}_i\epsilon_j'\} = \mathbf{0}$  for all  $i$  and  $j$ . Combining (1) and (2),

$$\begin{aligned}\mathbf{y}_i &= \mathbf{X}_i(\boldsymbol{\beta} + \mathbf{v}_i) + \epsilon_i \\ &= \mathbf{X}_i\boldsymbol{\beta} + \mathbf{u}_i\end{aligned}$$

with  $\mathbf{u}_i \equiv \mathbf{X}_i\mathbf{v}_i + \epsilon_i$ . Moreover,

$$\begin{aligned}E\{\mathbf{u}_i\mathbf{u}_i'\} &= E\{(\mathbf{X}_i\mathbf{v}_i + \epsilon_i)(\mathbf{X}_i\mathbf{v}_i + \epsilon_i)'\} \\ &= \mathbf{X}_i\Sigma\mathbf{X}_i' + \sigma_{ii}\mathbf{I} \\ &\equiv \boldsymbol{\Pi}_i\end{aligned}$$

Stacking the equations for the  $P$  panels,

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u} \quad (3)$$

where

$$\boldsymbol{\Pi} \equiv E\{\mathbf{u}\mathbf{u}'\} = \begin{bmatrix} \boldsymbol{\Pi}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Pi}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{\Pi}_P \end{bmatrix}$$

Estimating the parameters of (3) is a standard problem in generalized least squares (GLS), so

$$\begin{aligned}\hat{\boldsymbol{\beta}} &= (\mathbf{X}'\boldsymbol{\Pi}^{-1}\mathbf{X})^{-1} \mathbf{X}'\boldsymbol{\Pi}^{-1}\mathbf{y} \\ &= \left( \sum_i \mathbf{X}_i'\boldsymbol{\Pi}_i^{-1}\mathbf{X}_i \right)^{-1} \sum_i \mathbf{X}_i'\boldsymbol{\Pi}_i^{-1}\mathbf{y}_i \\ &= \sum_i \mathbf{W}_i\mathbf{b}_i\end{aligned} \quad (4)$$

where

$$\mathbf{W}_i = \left[ \sum_j \left\{ \Sigma + \sigma_{jj} (\mathbf{X}_j'\mathbf{X}_j)^{-1} \right\}^{-1} \right]^{-1} \left\{ \Sigma + \sigma_{ii} (\mathbf{X}_i'\mathbf{X}_i)^{-1} \right\}^{-1}$$

and  $\mathbf{b}_i \equiv (\mathbf{X}_i'\mathbf{X}_i)^{-1} \mathbf{X}_i'\mathbf{y}_i$ , showing that  $\hat{\boldsymbol{\beta}}$  is a weighted average of the panel-specific OLS estimates. The final equality in (4) makes use of the fact that

$$(\mathbf{A} + \mathbf{BDB}')^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{BEB}'\mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{BE}(\mathbf{E} + \mathbf{D})^{-1}\mathbf{EBA}'^{-1}$$

where  $\mathbf{E} \equiv (\mathbf{B}'\mathbf{A}^{-1}\mathbf{B})^{-1}$ . See Rao (1973, 33).

The variance of  $\hat{\boldsymbol{\beta}}$  is

$$\text{Var}(\hat{\boldsymbol{\beta}}) = (\mathbf{X}'\boldsymbol{\Pi}^{-1}\mathbf{X})^{-1} = \sum_i \left\{ \Sigma + \sigma_{ii}(\mathbf{X}_i'\mathbf{X}_i)^{-1} \right\}^{-1}$$

In addition to estimating  $\boldsymbol{\beta}$ , one often wishes to obtain estimates of the panel-specific  $\boldsymbol{\beta}_i$  vectors as well. As discussed by Judge et al. (1985, 541), if attention is restricted

to the class of estimators  $\{\beta_i^*\}$  for which  $E\{\beta_i^* | \beta_i\} = \beta_i$ , then the panel-specific OLS estimator  $\mathbf{b}_i$  is appropriate. However, if one does not condition on  $\beta_i$ , then the best linear unbiased predictor is

$$\begin{aligned}\widehat{\beta}_i &= \widehat{\beta} + \Sigma \mathbf{X}_i' (\mathbf{X}_i \Sigma \mathbf{X}_i' + \sigma_{ii} \mathbf{I})^{-1} (\mathbf{y}_i - \mathbf{X}_i \widehat{\beta}) \\ &= (\Sigma^{-1} + \sigma_{ii}^{-1} \mathbf{X}_i' \mathbf{X}_i)^{-1} (\sigma_{ii}^{-1} \mathbf{X}_i' \mathbf{X}_i \mathbf{b}_i + \Sigma^{-1} \widehat{\beta})\end{aligned}$$

Greene (1997, 672) suggests using the following method to obtain the variance of  $\widehat{\beta}_i$ . Define  $\mathbf{A}_i \equiv (\Sigma^{-1} + \sigma_{ii}^{-1} \mathbf{X}_i' \mathbf{X}_i)^{-1} \Sigma^{-1}$ . Then

$$\widehat{\beta}_i = [\mathbf{A}_i \quad (\mathbf{I} - \mathbf{A}_i)] \begin{bmatrix} \widehat{\beta} \\ \mathbf{b}_i \end{bmatrix}$$

and

$$\text{Var}(\widehat{\beta}_i) = [\mathbf{A}_i \quad (\mathbf{I} - \mathbf{A}_i)] \text{Var} \begin{pmatrix} \widehat{\beta} \\ \mathbf{b}_i \end{pmatrix} \begin{bmatrix} \mathbf{A}_i' \\ (\mathbf{I} - \mathbf{A}_i) \end{bmatrix}$$

Note that

$$\text{Var} \begin{pmatrix} \widehat{\beta} \\ \mathbf{b}_i \end{pmatrix} = \begin{bmatrix} \text{Var}(\widehat{\beta}) & \text{Cov}(\widehat{\beta}, \mathbf{b}_i) \\ \text{Cov}(\widehat{\beta}, \mathbf{b}_i) & \text{Var}(\mathbf{b}_i) \end{bmatrix}$$

The GLS estimator  $\widehat{\beta}$  is both consistent and efficient; and, although inefficient,  $\mathbf{b}_i$  is nevertheless also a consistent estimator of  $\beta$ . Thus, making use of Lemma 2.1 of Hausman (1978),  $\text{Asy.Cov}(\widehat{\beta}, \mathbf{b}_i) = \text{Asy.Var}(\widehat{\beta}) - \text{Asy.Cov}(\widehat{\beta}, \widehat{\beta} - \mathbf{b}_i) = \text{Asy.Var}(\widehat{\beta})$ . After some algebraic manipulation,

$$\text{Asy.Var}(\widehat{\beta}_i) = \text{Var}(\widehat{\beta}) + (\mathbf{I} - \mathbf{A}_i) \left\{ \text{Var}(\mathbf{b}_i) - \text{Var}(\widehat{\beta}) \right\} (\mathbf{I} - \mathbf{A}_i)'$$

To make the above formulas feasible, each  $\sigma_{ii}$  may be replaced with the consistent OLS estimate

$$\widehat{\sigma}_{ii} = \frac{(\mathbf{y}_i - \mathbf{X}_i \mathbf{b}_i)' (\mathbf{y}_i - \mathbf{X}_i \mathbf{b}_i)}{T_i - k}$$

Swamy (1970) showed that a consistent estimator of  $\Sigma$  is

$$\widehat{\Sigma} = \frac{1}{P-1} \left( \sum_{i=1}^P \mathbf{b}_i \mathbf{b}_i' - P \bar{\mathbf{b}} \bar{\mathbf{b}}' \right) - \frac{1}{P} \sum_{i=1}^P \widehat{\sigma}_{ii} (\mathbf{X}_i' \mathbf{X}_i)^{-1}$$

where  $\bar{\mathbf{b}} \equiv \frac{1}{P} \sum_i \mathbf{b}_i$ . However, that estimator may not always be positive definite in finite samples. A practical solution is to ignore the final term, and both Stata's `xtcrhh` command and the `xtcrhh2` command accompanying this article do that.

A natural question to ask is whether the panel-specific  $\beta_i$ s differ significantly from one another. Under the null hypothesis

$$H_0 : \beta_1 = \beta_2 = \cdots = \beta_P \quad (5)$$

the test statistic

$$T \equiv \sum_{i=1}^P (\mathbf{b}_i - \boldsymbol{\beta}^\dagger)' \{ \hat{\sigma}_{ii}^{-1}(\mathbf{X}_i \mathbf{X}_i) \} (\mathbf{b}_i - \boldsymbol{\beta}^\dagger)$$

where

$$\boldsymbol{\beta}^\dagger \equiv \left\{ \sum_{i=1}^P \hat{\sigma}_{ii}^{-1}(\mathbf{X}_i \mathbf{X}_i) \right\}^{-1} \sum_{i=1}^P \hat{\sigma}_{ii}^{-1}(\mathbf{X}_i \mathbf{X}_i) \mathbf{b}_i$$

is distributed  $\chi^2$  with  $k(P - 1)$  degrees of freedom.

## 3 Stata implementation

### 3.1 Syntax

```
xtrchh2 depvar varlist [if exp] [in range] [, i(varname) t(varname)
    level(#) offset(varname) noconstant nobetas]
```

#### Syntax for predict

```
predict [type] newvarname [if exp] [in range] [, [xb|stdp|xbi] group(#)
    nooffset]
```

### 3.2 Options

`i(varname)` specifies the variable that contains the unit to which the observation belongs. You can specify the `i()` option the first time you estimate, or you can use the `iis` command to set `i()` beforehand. Note that it is not necessary to specify `i()` if the data have been previously `tsset`, or if `iis` has been previously specified—in these cases, the group variable is taken from the previous setting. See [XT] `xt`.

`t(varname)` specifies the variable that contains the time at which the observation was made. You can specify the `t()` option the first time you estimate, or you can use the `tis` command to set `t()` beforehand. Note that it is not necessary to specify `t()` if the data have been previously `tsset`, or if `tis` has been previously specified—in these cases, the time variable is taken from the previous setting. See [XT] `xt`.

`level(#)` specifies the confidence level, in percent, for confidence intervals. The default is `level(95)` or as set by `set level`; see [U] **23.6 Specifying the width of confidence intervals**.

`offset(varname)` specifies that `varname` is to be included in the model with its coefficient constrained to be 1.

`noconstant` suppresses the constant term (intercept) in the regression.

`nobetas` requests that the panel-specific  $\widehat{\beta}_i$ s not be displayed.

### Options for predict

**xb**, the default, calculates the linear prediction based on  $\hat{\beta}$ .

**stdp** calculates the standard error of the linear prediction based on  $\hat{\beta}$ .

**xbi** calculates the linear prediction based on the group-specific  $\hat{\beta}_i$ , where  $i$  is specified with the **group(#)** option. The predictions are calculated for all available observations in the dataset, not just those in group  $i$ ; you can use **if** or **in** to restrict that behavior.

**group(#)** specifies which group-specific  $\hat{\beta}_i$  to use with the **xbi** option. The default is **group(1)**. **group(#)** has no effect if **xbi** is not specified.

**nooffset** is relevant only if you specified **offset(varname)** for **xtrchh2**. It modifies the calculations made by **predict** so that they ignore the offset variable; the linear prediction is treated as  $\mathbf{x}_{it}\mathbf{b}$  instead of  $\mathbf{x}_{it}\mathbf{b} + \text{offset}_{it}$ .

## 3.3 Remarks

The **xtrchh2** command fits Swamy's random-coefficients model as described in the previous section. The estimates of  $\hat{\beta}$  are identical to those produced by **xtrchh**. Additionally, **xtrchh2** displays the best linear unbiased estimates of the panel-specific coefficients; an option allows that output to be suppressed. Note that one can simply use the **statsby** command to obtain the panel-specific OLS estimates if they are desired. Saved results are stored in **e()** macros; see *Saved results* below.

## 4 Example

To illustrate the usage of **xtrchh2**, the following example uses the same dataset as [XT] **xtrchh**.

```
. webuse invest2, clear
. xtrchh2 invest market stock, i(company) t(time)
```

The output is shown on the next page. The header displays the number of observations and summarizes the structure of the panel data. It also contains a Wald test of the joint significance of the slope parameters in  $\hat{\beta}$ . Below the estimate of  $\hat{\beta}$  is the test statistic for the null hypothesis shown in (5). The remainder of the output consists of the estimated panel-specific  $\hat{\beta}_i$ s.

(Continued on next page)

```

Swamy random-coefficients regression      Number of obs      =      100
Group variable (i): company              Number of groups   =       5

Obs per group: min =      20
                  avg =     20.0
                  max =      20

Wald chi2(2)      =     17.55
Prob > chi2       =     0.0002

```

invest	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
market	.0807646	.0250829	3.22	0.001	.0316031	.1299261
stock	.2839885	.0677899	4.19	0.000	.1511229	.4168542
_cons	-23.58361	34.55547	-0.68	0.495	-91.31108	44.14386

Test of parameter constancy:    chi2(12) =    603.99            Prob > chi2 = 0.0000

Group-specific coefficients

	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
Group 1						
market	.1027848	.0108566	9.47	0.000	.0815062	.1240634
stock	.3678493	.0331352	11.10	0.000	.3029055	.4327931
_cons	-71.62927	37.46663	-1.91	0.056	-145.0625	1.803978
Group 2						
market	.084236	.0155761	5.41	0.000	.0537074	.1147647
stock	.3092167	.0301806	10.25	0.000	.2500638	.3683695
_cons	-9.819343	14.07496	-0.70	0.485	-37.40575	17.76707
Group 3						
market	.0279384	.013477	2.07	0.038	.0015241	.0543528
stock	.1508282	.0286904	5.26	0.000	.0945961	.2070603
_cons	-12.03268	29.58083	-0.41	0.684	-70.01004	45.94467
Group 4						
market	.0411089	.0118179	3.48	0.001	.0179461	.0642717
stock	.1407172	.0340279	4.14	0.000	.0740237	.2074108
_cons	3.269523	9.510794	0.34	0.731	-15.37129	21.91034
Group 5						
market	.147755	.0181902	8.12	0.000	.1121028	.1834072
stock	.4513312	.0569299	7.93	0.000	.3397506	.5629118
_cons	-27.70628	42.12524	-0.66	0.511	-110.2702	54.85766



## 5 Saved results

`xtrchh2` saves in `e()`:

### Scalars

<code>e(N)</code>	number of observations	<code>e(g_avg)</code>	average group size
<code>e(chi2)</code>	$\chi^2$	<code>e(N_g)</code>	number of groups
<code>e(df_m)</code>	model degrees of freedom	<code>e(chi2_c)</code>	$\chi^2$ for comparison test
<code>e(g_max)</code>	largest group size	<code>e(df_chi2c)</code>	degrees of freedom for comparison test
<code>e(g_min)</code>	smallest group size		

### Macros

<code>e(cmd)</code>	<code>xtrchh2</code>	<code>e(depvar)</code>	name of dependent variable
<code>e(predict)</code>	program used to implement <code>predict</code>	<code>e(chi2type)</code>	Wald; type of model $\chi^2$ test
<code>e(ivar)</code>	variable denoting groups	<code>e(title)</code>	title in estimation output

### Matrices

<code>e(b)</code>	$\hat{\beta}$ vector	<code>e(V_i)</code>	estimated $\text{Var}(\hat{\beta}_i)$ , $i=1\dots P$
<code>e(V)</code>	estimated $\text{Var}(\hat{\beta})$	<code>e(Sigma)</code>	$\hat{\Sigma}$
<code>e(beta_i)</code>	$\hat{\beta}_i$ vector, $i=1\dots P$		

### Functions

<code>e(sample)</code>	marks estimation sample
------------------------	-------------------------

## 6 References

- Greene, W. H. 1997. *Econometric Analysis*. 3d ed. Upper Saddle River, NJ: Prentice-Hall.
- Hausman, J. A. 1978. Specification tests in econometrics. *Econometrica* 46(6): 1251–1271.
- Judge, G. G., R. C. Hill, W. E. Griffiths, H. Lütkepohl, and T. C. Lee. 1985. *The Theory and Practice of Econometrics*. 2d ed. New York: John Wiley & Sons.
- Rao, C. R. 1973. *Linear Statistical Inference and its Applications*. 2d ed. New York: John Wiley & Sons.
- Swamy, P. A. V. 1970. Efficient inference in a random coefficient regression model. *Econometrica* 38: 311–323.

### About the Author

Brian P. Poi received his Ph.D. in economics from the University of Michigan and joined Stata Corporation as a staff statistician.