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

$$\mathbf{y}_i = \mathbf{X}_i \boldsymbol{\beta}_i + \boldsymbol{\epsilon}_i \tag{1}$$

where $i = 1...P$ denotes panels, $\mathbf{y}_i$ is a $T_i \times 1$ vector of observations for the ith 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 $\boldsymbol{\beta}$:

$$\boldsymbol{\beta}_i = \boldsymbol{\beta} + \mathbf{v}_i \tag{2}$$

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

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

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

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

Stacking the equations for the P panels,

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u} \tag{3}$$

where

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

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

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

where

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

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

$$\left(\mathbf{A} + \mathbf{B}\mathbf{D}\mathbf{B}'\right)^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}\mathbf{E}\mathbf{B}'\mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{B}\mathbf{E}\left(\mathbf{E} + \mathbf{D}\right)^{-1}\mathbf{E}\mathbf{B}\mathbf{A}'^{-1}$$

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

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

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

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 $\{\boldsymbol{\beta}_i^*\}$ for which $\mathrm{E}\{\boldsymbol{\beta}_i^* \mid \boldsymbol{\beta}_i\} = \boldsymbol{\beta}_i$, then the panel-specific OLS estimator $\mathbf{b}_i$ is appropriate. However, if one does not condition on $\boldsymbol{\beta}_i$, then the best linear unbiased predictor is

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

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

$$
\widehat{\boldsymbol{\beta}_i} = \begin{bmatrix} \mathbf{A}_i & (\mathbf{I} - \mathbf{A}_i) \end{bmatrix} \begin{bmatrix} \widehat{\boldsymbol{\beta}} \\ \mathbf{b}_i \end{bmatrix}
$$

and

$$
\mathrm{Var}(\widehat{\boldsymbol{\beta}_i}) = \begin{bmatrix} \mathbf{A}_i & (\mathbf{I} - \mathbf{A}_i) \end{bmatrix} \mathrm{Var}\begin{pmatrix} \widehat{\boldsymbol{\beta}} \\ \mathbf{b}_i \end{pmatrix} \begin{bmatrix} \mathbf{A}_i' \\ (\mathbf{I} - \mathbf{A}_i)' \end{bmatrix}
$$

Note that

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

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

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

To make the above formulas feasible, each σ_{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 $\boldsymbol{\Sigma}$ is

$$
\widehat{\boldsymbol{\Sigma}} = \frac{1}{P-1}\left(\sum_{i=1}^{P} \mathbf{b}_i\mathbf{b}_i' - P\overline{\mathbf{b}}\,\overline{\mathbf{b}}'\right) - \frac{1}{P}\sum_{i=1}^{P} \widehat{\sigma_{ii}}(\mathbf{X}_i'\mathbf{X}_i)^{-1}
$$

where $\overline{\mathbf{b}} \equiv \frac{1}{P}\sum_i \mathbf{b}_i$. 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 `xtrchh` command and the `xtrchh2` command accompanying this article do that.

A natural question to ask is whether the panel-specific $\boldsymbol{\beta}_i$s differ significantly from one another. Under the null hypothesis

$$
H_0 : \boldsymbol{\beta}_1 = \boldsymbol{\beta}_2 = \cdots = \boldsymbol{\beta}_P \tag{5}
$$

the test statistic

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

where

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

is distributed χ^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*)
  <u>level</u>(#) <u>off</u>set(*varname*) <u>nocon</u>stant <u>nobetas</u>]

Syntax for predict

predict [*type*] *newvarname* [if *exp*] [in *range*] [, [xb|stdp|xbi] <u>group</u>(#)
  <u>nooff</u>set]

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 $\widehat{\boldsymbol{\beta}_i}$s not be displayed.

Options for predict

xb, the default, calculates the linear prediction based on $\widehat{\boldsymbol{\beta}}$.

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

xbi calculates the linear prediction based on the group-specific $\widehat{\beta}_i$, 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 $\widehat{\beta}_i$ 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 $\widehat{\beta}$ 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 $\widehat{\boldsymbol{\beta}}$. Below the estimate of $\widehat{\boldsymbol{\beta}}$ is the test
statistic for the null hypothesis shown in (5). The remainder of the output consists of
the estimated panel-specific $\widehat{\boldsymbol{\beta}_i}$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

e(N)	number of observations	e(g_avg)	average group size
e(chi2)	χ^2	e(N_g)	number of groups
e(df_m)	model degrees of freedom	e(chi2_c)	χ^2 for comparison test
e(g_max)	largest group size	e(df_chi2c)	degrees of freedom for
e(g_min)	smallest group size		comparison test

Macros

e(cmd)	xtrchh2	e(depvar)	name of dependent variable
e(predict)	program used to implement	e(chi2type)	Wald; type of model χ^2 test
	predict	e(title)	title in estimation output
e(ivar)	variable denoting groups		

Matrices

e(b)	$\widehat{\beta}$ vector	e(V_i)	estimated $\mathrm{Var}(\widehat{\beta_i})$, $i=1...P$
e(V)	estimated $\mathrm{Var}(\widehat{\beta})$	e(Sigma)	$\widehat{\Sigma}$
e(beta_i)	$\widehat{\beta_i}$ vector, $i=1...P$		

Functions

e(sample)	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.