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

\[
y_i = X_i(\beta + v_i) + \epsilon_i = X_i\beta + u_i
\]

with \( u_i \equiv X_i v_i + \epsilon_i \). Moreover,

\[
\mathbb{E}\{u_i u'_i\} = \mathbb{E}\{(X_i v_i + \epsilon_i)(X_i v_i + \epsilon'_i)\}
\]

\[
= X_i \Sigma X'_i + \sigma_{ii} I
\]

\[
\equiv \Pi_i
\]

Stacking the equations for the \( P \) panels,

\[
y = X\beta + u
\]

(3)

where

\[
\Pi \equiv \mathbb{E}\{uu'\} = \begin{bmatrix}
\Pi_1 & 0 & \cdots & 0 \\
0 & \Pi_2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \Pi_P
\end{bmatrix}
\]

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

\[
\hat{\beta} = (X'\Pi^{-1}X)^{-1} X'\Pi^{-1} y
\]

\[
= \left( \sum_i X'_i \Pi_i^{-1} X_i \right)^{-1} \sum_i X'_i \Pi_i^{-1} y_i
\]

(4)

where

\[
W_i = \left[ \sum_j \left\{ \Sigma + \sigma_{jj} (X'_i X_j)^{-1} \right\}^{-1} \left\{ \Sigma + \sigma_{ii} (X'_i X_i)^{-1} \right\}^{-1} \right]^{-1}
\]

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

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

where \( E = (b' A^{-1} b)^{-1} \). See Rao (1973, 33).

The variance of \( \hat{\beta} \) is

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

In addition to estimating \( \beta \), one often wishes to obtain estimates of the panel-specific \( \beta_i \) 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 \(b_i\) is appropriate. However, if one does not condition on \(\beta_i\), then the best linear unbiased predictor is

\[
\hat{\beta}_i = \hat{\beta} + \Sigma X_i' (X_i \Sigma X_i' + \sigma_{ii} I)^{-1} \left( y_i - X_i \hat{\beta} \right)
\]

\[
= (\Sigma^{-1} + \sigma_{ii}^{-1} X_i' X_i)^{-1} \left( \sigma_{ii}^{-1} X_i' X_i b_i + \Sigma^{-1} \hat{\beta} \right)
\]

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

\[
\hat{\beta}_i = \left[ A_i (\hat{\beta} - A_i) \right] \left[ A_i' \right]
\]

and

\[
\text{Var}(\hat{\beta}_i) = [A_i (I - A_i)] \text{Var}(\hat{\beta} - A_i) [A_i' (I - A_i)]'
\]

Note that

\[
\text{Var}(\hat{\beta} - A_i) = \text{Var}(\hat{\beta}) + (I - A_i) \left\{ \text{Var}(b_i) - \text{Var}(\hat{\beta}) \right\} (I - A_i)'
\]

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

\[
\hat{\sigma}_{ii} = \frac{(y_i - X_i b_i)' (y_i - X_i b_i)}{T_i - k}
\]

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

\[
\hat{\Sigma} = \frac{1}{P - 1} \left( \sum_{i=1}^{P} b_i b_i' - P \bar{b} \bar{b}' \right) - \frac{1}{P} \sum_{i=1}^{P} \hat{\sigma}_{ii} (X_i' X_i)^{-1}
\]

where \(\bar{b} \equiv \frac{1}{P} \sum_i b_i\). 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 \texttt{xtrchh} command and the \texttt{xtrchh2} 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

\[
H_0 : \beta_1 = \beta_2 = \cdots = \beta_p
\]
the test statistic

\[ T \equiv \sum_{i=1}^{P} (b_i - \beta^\dagger)' [\hat{\sigma}^{-1}_{ii}(X_iX_i)] (b_i - \beta^\dagger) \]

where

\[ \beta^\dagger \equiv \left\{ \sum_{i=1}^{P} \hat{\sigma}^{-1}_{ii}(X_iX_i) \right\}^{-1} \sum_{i=1}^{P} \hat{\sigma}^{-1}_{ii}(X_iX_i)b_i \]

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 `iw` 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 \( \hat{\beta} \)'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 $x_{it}b$ instead of $x_{it}b + \text{offset}_{it}$.

### 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 $\beta$. Below the estimate of $\beta$ 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

|     | Coef. | Std. Err. | z     | P>|z| | [95% Conf. Interval] |
|-----|-------|-----------|-------|------|---------------------|
| invest | 0.0807646 | 0.0250829 | 3.22  | 0.001 | (0.0316031, 0.1299261) |
|       | 0.2839885 | 0.0677899 | 4.19  | 0.000 | (0.1511229, 0.4168542) |
| _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

| Group  | Coef. | Std. Err. | z     | P>|z| | [95% Conf. Interval] |
|--------|-------|-----------|-------|------|---------------------|
| Group 1 |       |           |       |      |                     |
| market | 0.1027848 | 0.0108566 | 9.47  | 0.000 | (0.0815062, 0.1240634) |
| stock  | 0.3678493 | 0.0331352 | 11.10 | 0.000 | (0.3029055, 0.4327931) |
| _cons  | -71.62927 | 37.46663 | -1.91 | 0.056 | (-145.0625, 1.803978) |
| Group 2 |       |           |       |      |                     |
| market | 0.084236 | 0.0155761 | 5.41  | 0.000 | (0.0537074, 0.1147647) |
| stock  | 0.3092167 | 0.0301806 | 10.25 | 0.000 | (0.2500638, 0.3683695) |
| _cons  | -9.819343 | 14.07496 | -0.70 | 0.485 | (-37.40575, 17.76707) |
| Group 3 |       |           |       |      |                     |
| market | 0.0279384 | 0.013477 | 2.07  | 0.038 | (0.0015241, 0.0543528) |
| stock  | 0.1508282 | 0.0286904 | 5.26  | 0.000 | (0.0945961, 0.2070603) |
| _cons  | -12.03268 | 29.58063 | -0.41 | 0.684 | (-70.01004, 45.94467) |
| Group 4 |       |           |       |      |                     |
| market | 0.0411089 | 0.0118179 | 3.48  | 0.001 | (0.0179461, 0.0642717) |
| stock  | 0.1407172 | 0.0340279 | 4.14  | 0.000 | (0.0760237, 0.2074108) |
| _cons  | 3.269523 | 9.510794 | 0.34  | 0.731 | (-15.37129, 21.91034) |
| Group 5 |       |           |       |      |                     |
| market | 0.147755 | 0.0181902 | 8.12  | 0.000 | (0.1121028, 0.1834072) |
| stock  | 0.4513312 | 0.0569299 | 7.93  | 0.000 | (0.3397506, 0.5629118) |
| _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(chi2)` \( \chi^2 \) for comparison test
- `e(chi2c)` degrees of freedom for comparison test
- `e(N_g)` number of groups
- `e(df_m)` model degrees of freedom
- `e(g_max)` largest group size
- `e(g_avg)` average group size
- `e(g_min)` smallest group size
- `e(df_chi2c)` degrees of freedom for comparison test

Macros

- `e(cmd)` xtrchh2
- `e(predict)` program used to implement
- `e(i1var)` variable denoting groups
- `e(depvar)` name of dependent variable
- `e(chi2type)` Wald; type of model \( \chi^2 \) test
- `e(title)` title in estimation output

Matrices

- `e(V)` \( \hat{\beta} \) vector
- `e(V_{i,i})` \( \hat{\beta}_i \) vector, \( i=1...P \)
- `e(Sigma)` \( \hat{\Sigma} \)

Functions

- `e(sample)` marks estimation sample

6 References


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.