Description

\texttt{xtrc} fits the Swamy (1970) random-coefficients linear regression model, which does not impose the assumption of constant parameters across panels. Average coefficient estimates are reported by default, but panel-specific coefficients may be requested.

Quick start

Random-coefficients regression of \textit{y} on \textit{x1} and \textit{x2} using \texttt{xtset} data
\begin{verbatim}
xtrc y x1 x2
\end{verbatim}

As above, but report panel-specific best linear predictors
\begin{verbatim}
xtrc y x1 x2, betas
\end{verbatim}

Multiple-imputation estimates of random-coefficients regression using \texttt{mi xtset} data
\begin{verbatim}
mi estimate: xtrc y x
\end{verbatim}

Menu

Statistics > Longitudinal/panel data > Random-coefficients regression by GLS
```plaintext
Syntax

    xtrc  depvar indepvars  [if]  [in]  [ ,  options ]

options       Description
--------------------------
Main
   noconstant          suppress constant term
   offset( varname )   include varname in model with coefficient constrained to 1
SE
   vce( vcetype )     vcetype may be conventional, bootstrap, or jackknife
Reporting
   level(#)           set confidence level; default is level(95)
   betas              display group-specific best linear predictors
   display_options    control columns and column formats, row spacing, line width,
                      display of omitted variables and base and empty cells, and
                      factor-variable labeling
   coeflegend        display legend instead of statistics

A panel variable must be specified; use xtset; see [XT] xtset.
indepvars may contain factor variables; see [U] 11.4.3 Factor variables.
by, mi estimate, and statsby are allowed; see [U] 11.1.10 Prefix commands.
vce(bootstrap) and vce(jackknife) are not allowed with the mi estimate prefix; see [MI] mi estimate.
coeflegend does not appear in the dialog box.
See [U] 20 Estimation and postestimation commands for more capabilities of estimation commands.

Options

   noconstant, offset( varname ): see [R] Estimation options

   vce(vcetype) specifies the type of standard error reported, which includes types that are derived from
   asymptotic theory (conventional) and that use bootstrap or jackknife methods (bootstrap, jackknife); see [XT] vce_options.
   vce(conventional), the default, uses the conventionally derived variance estimator for generalized
   least-squares regression.

   level(#) ; see [R] Estimation options.
   betas requests that the group-specific best linear predictors also be displayed.
   display_options : noci, nopvalues, noomitted, vsquish, noemptycells, baselevels,
   allbaselevels, nofvlabel, fvwrap(#), fvwrapon(style), cformat(%,fmt), pformat(%,fmt),
   sformat(%,fmt), and nolstretch; see [R] Estimation options.
```
The following option is available with `xtrc` but is not shown in the dialog box: `coeflegend`; see [R] Estimation options.

Remarks and examples

In random-coefficients models, we wish to treat the parameter vector as a realization (in each panel) of a stochastic process. `xtrc` fits the Swamy (1970) random-coefficients model, which is suitable for linear regression of panel data. See Greene (2012, chap. 11) and Poi (2003) for more information about this and other panel-data models.

Example 1

Greene (2012, 1112) reprints data from a classic study of investment demand by Grunfeld and Griliches (1960). In [XT] `xtgls`, we use this dataset to illustrate many of the possible models that may be fit with the `xtgls` command. Although the models included in the `xtgls` command offer considerable flexibility, they all assume that there is no parameter variation across firms (the cross-sectional units).

To take a first look at the assumption of parameter constancy, we should `reshape` our data so that we may fit a simultaneous-equation model with `sureg`; see [R] `sureg`. Because there are only five panels here, this is not too difficult.

```
  . use https://www.stata-press.com/data/r16/invest2
  . reshape wide invest market stock, i(time) j(company)
     (note: j = 1 2 3 4 5)
     Data  long  ->  wide
     Number of obs.  100  ->  20
     Number of variables  5  ->  16
     j variable (5 values)  company  ->  (dropped)
     xij variables:  invest  ->  invest1 invest2 ... invest5
                     market  ->  market1 market2 ... market5
                     stock  ->  stock1 stock2 ... stock5
```
. sureg (invest1 market1 stock1) (invest2 market2 stock2)>
   (invest3 market3 stock3) (invest4 market4 stock4) (invest5 market5 stock5)
Seemingly unrelated regression

<table>
<thead>
<tr>
<th>Equation</th>
<th>Obs</th>
<th>Parms</th>
<th>RMSE</th>
<th>&quot;R-sq&quot;</th>
<th>chi2</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>invest1</td>
<td>20</td>
<td>2</td>
<td>84.94729</td>
<td>0.9207</td>
<td>261.32</td>
<td>0.0000</td>
</tr>
<tr>
<td>invest2</td>
<td>20</td>
<td>2</td>
<td>12.36322</td>
<td>0.9119</td>
<td>207.21</td>
<td>0.0000</td>
</tr>
<tr>
<td>invest3</td>
<td>20</td>
<td>2</td>
<td>26.46612</td>
<td>0.6876</td>
<td>46.88</td>
<td>0.0000</td>
</tr>
<tr>
<td>invest4</td>
<td>20</td>
<td>2</td>
<td>9.742303</td>
<td>0.7264</td>
<td>59.15</td>
<td>0.0000</td>
</tr>
<tr>
<td>invest5</td>
<td>20</td>
<td>2</td>
<td>95.85484</td>
<td>0.4220</td>
<td>14.97</td>
<td>0.0006</td>
</tr>
</tbody>
</table>

| Coef.  | Std. Err. | z   | P>|z| | [95% Conf. Interval] |
|--------|-----------|-----|------|-----------------------|
| invest1 |           |     |      |                       |
| market1 | 0.120493  | 0.021629 | 5.57 | 0.000 | 0.0781007 | 0.1628853 |
| stock1  | 0.3827462 | 0.032768 | 11.68 | 0.000 | 0.318522  | 0.4469703 |
| _cons  | -162.3641 | 89.45922 | -1.81 | 0.070 | -337.7009 | 12.97279 |

| invest2 |           |     |      |                       |
| market2 | 0.0695456 | 0.0168975 | 4.12 | 0.000 | 0.0364271 | 0.1026641 |
| stock2  | 0.3085445 | 0.0258635 | 11.93 | 0.000 | 0.2578529 | 0.3592362 |
| _cons  | 0.5043112 | 11.51283 | 0.04  | 0.965 | -22.06042 | 23.06904 |

| invest3 |           |     |      |                       |
| market3 | 0.0372914 | 0.0122631 | 3.04 | 0.002 | 0.0132561 | 0.0613268 |
| stock3  | 0.130783  | 0.0220497 | 5.93 | 0.000 | 0.0875663 | 0.1739997 |
| _cons  | -22.43892 | 25.51859 | -0.88 | 0.379 | -72.45443 | 27.57659 |

| invest4 |           |     |      |                       |
| market4 | 0.0570091 | 0.0113623 | 5.02 | 0.000 | 0.0347395 | 0.0792788 |
| stock4  | 0.0415065 | 0.0120161 | 1.01  | 0.314 | -0.0392472 | 0.1226202 |
| _cons  | 1.088878  | 6.258805 | 0.17  | 0.862 | -11.17815 | 13.35591 |

| invest5 |           |     |      |                       |
| market5 | 0.1014782 | 0.0547837 | 1.85 | 0.064 | -0.0058958 | 0.2088523 |
| stock5  | 0.3999914 | 0.1277946 | 3.13 | 0.002 | 0.1495186 | 0.6504642 |
| _cons  | 85.42324  | 111.8774 | 0.76  | 0.445 | -133.8525 | 304.6989 |

Test of parameter constancy:  

| Coef.  | Std. Err. | z   | P>|z| | [95% Conf. Interval] |
|--------|-----------|-----|------|-----------------------|
| market | 0.0807646 | 0.0250829 | 3.22 | 0.001 | 0.0316031 | 0.1299261 |
| stock  | 0.2839885 | 0.0677899 | 4.19 | 0.000 | 0.1512229 | 0.4168542 |
| _cons  | -23.58361 | 34.55547 | -0.68 | 0.495 | -91.31108 | 54.14386 |

Here we instead fit a random-coefficients model:

. use https://www.stata-press.com/data/r16/invest2
. xtrc invest market stock

Random-coefficients regression

| Coef.  | Std. Err. | z   | P>|z| | [95% Conf. Interval] |
|--------|-----------|-----|------|-----------------------|
| market | 0.0807646 | 0.0250829 | 3.22 | 0.001 | 0.0316031 | 0.1299261 |
| stock  | 0.2839885 | 0.0677899 | 4.19 | 0.000 | 0.1512229 | 0.4168542 |
| _cons  | -23.58361 | 34.55547 | -0.68 | 0.495 | -91.31108 | 54.14386 |

Test of parameter constancy:  

chi2(12) = 603.99  Prob > chi2 = 0.0000
Just as the results of our simultaneous-equation model do not support the assumption of parameter constancy, the test included with the random-coefficients model also indicates that the assumption is not valid for these data. With large panel datasets, we would not want to take the time to look at a simultaneous-equations model (aside from the fact that our doing so was subjective).

Stored results

`xtrc` stores the following in `e()`:

Scalars

- `e(N)`: number of observations
- `e(N_g)`: number of groups
- `e(df_m)`: model degrees of freedom
- `e(chi2)`: $\chi^2$
- `e(chi2_c)`: $\chi^2$ for comparison test
- `e(df_chic2c)`: degrees of freedom for comparison $\chi^2$ test
- `e(g_min)`: smallest group size
- `e(g_avg)`: average group size
- `e(g_max)`: largest group size
- `e(rank)`: rank of `e(V)`

Macros

- `e(cmd)`: `xtrc`
- `e(cmdline)`: command as typed
- `e(depvar)`: name of dependent variable
- `e(ivar)`: variable denoting groups
- `e(tvar)`: variable denoting time within groups
- `e(title)`: title in estimation output
- `e(offset)`: linear offset variable
- `e(chi2type)`: Wald; type of model $\chi^2$ test
- `e(vce)`: `vcetype` specified in `vce()`
- `e(properties)`: `b V`
- `e(predict)`: program used to implement `predict`
- `e(marginsnotok)`: predictions disallowed by `margins`
- `e(asbalanced)`: factor variables `fvset` as `asbalanced`
- `e(asobserved)`: factor variables `fvset` as `asobserved`

Matrices

- `e(b)`: coefficient vector
- `e(Sigma)`: $\Sigma$ matrix
- `e(beta_ps)`: matrix of best linear predictors
- `e(V)`: variance–covariance matrix of the estimators
- `e(V_ps)`: matrix of variances for the best linear predictors; row $i$ contains vec of variance matrix for group $i$ predictor

Functions

- `e(sample)`: marks estimation sample

In addition to the above, the following is stored in `r()`:

Matrices

- `r(table)`: matrix containing the coefficients with their standard errors, test statistics, $p$-values, and confidence intervals

Note that results stored in `r()` are updated when the command is replayed and will be replaced when any r-class command is run after the estimation command.
Methods and formulas

In a random-coefficients model, the parameter heterogeneity is treated as stochastic variation. Assume that we write

\[ y_i = X_i \beta_i + \epsilon_i \]

where \( i = 1, \ldots, m \), and \( \beta_i \) is the coefficient vector \((k \times 1)\) for the \( i \)th cross-sectional unit, such that

\[ \beta_i = \beta + \nu_i \quad E(\nu_i) = 0 \quad E(\nu_i \nu_i') = \Sigma \]

Our goal is to find \( \hat{\beta} \) and \( \hat{\Sigma} \).

The derivation of the estimator assumes that the cross-sectional specific coefficient vector \( \beta_i \) is the outcome of a random process with mean vector \( \beta \) and covariance matrix \( \Sigma \),

\[ y_i = X_i \beta_i + \epsilon_i = X_i(\beta + \nu_i) + \epsilon_i = X_i\beta + (X_i\nu_i + \epsilon_i) = X_i\beta + \omega_i \]

where \( E(\omega_i) = 0 \) and

\[ E(\omega_i \omega_i') = E\left\{ (X_i\nu_i + \epsilon_i)(X_i\nu_i + \epsilon_i)' \right\} = E(\epsilon_i \epsilon_i') + X_i E(\nu_i \nu_i')X_i' = \sigma_i^2 I + X_i \Sigma X_i' = \Pi_i \]

Stacking the \( m \) equations, we have

\[ y = X\beta + \omega \]

where \( \Pi \equiv E(\omega \omega') \) is a block diagonal matrix with \( \Pi_i, \ i = 1 \ldots m \), along the main diagonal and zeros elsewhere. The GLS estimator of \( \hat{\beta} \) is then

\[ \hat{\beta} = \left( \sum_i X_i' \Pi_i^{-1} X_i \right)^{-1} \sum_i X_i' \Pi_i^{-1} y_i = \sum_{i=1}^m W_i b_i \]

where

\[ W_i = \left\{ \sum_{i=1}^m (\Sigma + V_i)^{-1} \right\}^{-1} (\Sigma + V_i)^{-1} \]

\( b_i = (X_i'X_i)^{-1} X_i' y_i \) and \( V_i = \sigma_i^2 (X_i'X_i)^{-1} \), showing that the resulting GLS estimator is a matrix-weighted average of the panel-specific OLS estimators. The variance of \( \hat{\beta} \) is

\[ \text{Var}(\beta) = \sum_{i=1}^m (\Sigma + V_i)^{-1} \]
To calculate the above estimator $\hat{\beta}$ for the unknown $\Sigma$ and $V_i$ parameters, we use the two-step approach suggested by Swamy (1970):

\[ b_i = \text{OLS panel-specific estimator} \]
\[ \hat{\sigma}^2_i = \frac{\varepsilon_i'\varepsilon_i}{n_i - k} \]
\[ \hat{V}_i = \hat{\sigma}^2_i (X_i'X_i)^{-1} \]
\[ \bar{b} = \frac{1}{m} \sum_{i=1}^{m} b_i \]
\[ \hat{\Sigma} = \frac{1}{m-1} \left( \sum_{i=1}^{m} b_i b_i' - m\bar{b} \bar{b}' \right) - \frac{1}{m} \sum_{i=1}^{m} \hat{V}_i \]

The two-step procedure begins with the usual OLS estimates of $\beta_i$. With those estimates, we may proceed by obtaining estimates of $\hat{V}_i$ and $\hat{\Sigma}$ (and thus $\hat{W}_i$) and then obtain an estimate of $\beta$.

Swamy (1970) further points out that the matrix $\hat{\Sigma}$ may not be positive definite and that because the second term is of order $1/(mT)$, it is negligible in large samples. A simple and asymptotically expedient solution is simply to drop this second term and instead use

\[ \hat{\Sigma} = \frac{1}{m-1} \left( \sum_{i=1}^{m} b_i b_i' - m\bar{b} \bar{b}' \right) \]

As discussed by Judge et al. (1985, 541), the feasible best linear predictor of $\beta_i$ is given by

\[ \hat{\beta}_i = \hat{\beta} + \hat{\Sigma}X_i' \left( X_i\hat{\Sigma}X_i' + \hat{\sigma}^2_i I \right)^{-1} \left( y_i - X_i\hat{\beta} \right) \]
\[ = \left( \hat{\Sigma}^{-1} + \hat{V}_i^{-1} \right)^{-1} \left( \hat{\Sigma}^{-1}\hat{\beta} + \hat{V}_i^{-1}b_i \right) \]

The conventional variance of $\hat{\beta}_i$ is given by

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

where

\[ A_i = \left( \hat{\Sigma}^{-1} + \hat{V}_i^{-1} \right)^{-1} \hat{\Sigma}^{-1} \]

To test the model, we may look at the difference between the OLS estimate of $\beta$, ignoring the panel structure of the data and the matrix-weighted average of the panel-specific OLS estimators. The test statistic suggested by Swamy (1970) is given by

\[ \chi^2_{k(m-1)} = \sum_{i=1}^{m} (b_i - \bar{\beta}^*)' \hat{V}_i^{-1} (b_i - \bar{\beta}^*) \quad \text{where} \quad \bar{\beta}^* = \left( \sum_{i=1}^{m} \hat{V}_i^{-1} \right)^{-1} \sum_{i=1}^{m} \hat{V}_i^{-1} b_i \]
Johnston and DiNardo (1997) have shown that the test is algebraically equivalent to testing

$$H_0 : \beta_1 = \beta_2 = \cdots = \beta_m$$

in the generalized (groupwise heteroskedastic) \textit{xtgls} model, where $V$ is block diagonal with $i$th diagonal element $\Pi_i$.

References


Also see

[XT] \textit{xtrc postestimation} — Postestimation tools for \textit{xtrc}
[XT] \textit{xtreg} — Fixed-, between-, and random-effects and population-averaged linear models
[XT] \textit{xtset} — Declare data to be panel data
[ME] \textit{mixed} — Multilevel mixed-effects linear regression
[M] \textit{Estimation} — Estimation commands for use with \textit{mi estimate}
[U] 20 \textit{Estimation and postestimation commands}