

[Description](#)
[Options](#)
[References](#)

[Quick start](#)
[Remarks and examples](#)
[Also see](#)

[Menu](#)
[Stored results](#)

[Syntax](#)
[Methods and formulas](#)

Description

`sureg` fits seemingly unrelated regression models (Zellner 1962; Zellner and Huang 1962; Zellner 1963). The acronyms SURE and SUR are often used for the estimator.

Quick start

Model of y_1 and y_2 as a function of x_1 , x_2 , and x_3

```
sureg (y1 x1 x2 x3) (y2 x1 x2 x3)
```

Same as above

```
sureg (y1 y2 = x1 x2 x3)
```

Same as above, but obtain small-sample statistics and use small-sample adjustment

```
sureg (y1 y2 = x1 x2 x3), small dfk
```

Also perform Breusch–Pagan test

```
sureg (y1 y2 = x1 x2 x3), small dfk corr
```

Model of y_1 as a function of x_1 and x_2 and y_2 as a function of x_2 and lag of x_2 using `tsset` data

```
sureg (y1 x1 x2) (y2 x2 L.x2)
```

Menu

Statistics > Linear models and related > Multiple-equation models > Seemingly unrelated regression

Syntax

Basic syntax

```
sureg (depvar1 varlist1) (depvar2 varlist2) ... (depvarN varlistN)
      [if] [in] [weight]
```

Full syntax

```
sureg ([eqname1:]depvar1a [depvar1b ... = ]varlist1 [ , noconstant ])
      ([eqname2:]depvar2a [depvar2b ... = ]varlist2 [ , noconstant ])
      ...
      ([eqnameN:]depvarNa [depvarNb ... = ]varlistN [ , noconstant ])
      [if] [in] [weight] [ , options ]
```

Explicit equation naming (*eqname*:) cannot be combined with multiple dependent variables in an equation specification.

| <i>options</i> | Description |
|---|--|
| Model | |
| <u>i</u> sure | iterate until estimates converge |
| <u>c</u> onstraints(<i>constraints</i>) | apply specified linear constraints |
| SE/Robust | |
| <u>v</u> ce(<i>vcetype</i>) | <i>vcetype</i> may be <u>u</u> nadjusted, <u>r</u> obust, or <u>c</u> luster <i>clustvar</i> |
| df adj. | |
| <u>s</u> mall | report small-sample statistics |
| <u>d</u> fk | use small-sample adjustment |
| <u>d</u> fk2 | use alternate adjustment |
| Reporting | |
| <u>l</u> evel(#) | set confidence level; default is level(95) |
| <u>c</u> orr | perform Breusch–Pagan test |
| <u>n</u> ocnsreport | do not display constraints |
| <i>display_options</i> | control columns and column formats, row spacing, line width, display of omitted variables and base and empty cells, and factor-variable labeling |
| Optimization | |
| <i>optimization_options</i> | control the optimization process; seldom used |
| <u>n</u> oheader | suppress header table from above coefficient table |
| <u>n</u> otable | suppress coefficient table |
| <u>c</u> oeflegend | display legend instead of statistics |

*varlist*₁, ..., *varlist*_{*N*} may contain factor variables; see [U] 11.4.3 Factor variables. You must have the same levels of factor variables in all equations that have factor variables.

deprvars and the *varlists* may contain time-series operators; see [U] 11.4.4 Time-series varlists.

bootstrap, *by*, *collect*, *fp*, *jackknife*, *rolling*, and *statsby* are allowed; see [U] 11.1.10 Prefix commands.

Weights are not allowed with the *bootstrap* prefix; see [R] *bootstrap*.

*aweight*s are not allowed with the *jackknife* prefix; see [R] *jackknife*.

*aweight*s and *fweight*s are allowed; see [U] 11.1.6 weight.

noheader, *notable*, and *coeflegend* do not appear in the dialog box.

See [U] 20 Estimation and postestimation commands for more capabilities of estimation commands.

Options

Model

isure specifies that *sureg* iterate over the estimated disturbance covariance matrix and parameter estimates until the parameter estimates converge. Under seemingly unrelated regression, this iteration converges to the maximum likelihood results. If this option is not specified, *sureg* produces two-step estimates.

constraints (*constraints*); see [R] Estimation options.

SE/Robust

vce (*vcetype*) specifies the type of standard error reported, which includes types that are derived from asymptotic theory (*unadjusted*), that are robust to some kinds of misspecification (*robust*), and that allow for intragroup correlation (*cluster* *clustvar*); see [R] *vce_option*.

vce(*unadjusted*), the default, specifies that an unadjusted (nonrobust) VCE matrix be used; this results in efficient estimates when assuming homoskedasticity.

df adj.

small specifies that small-sample statistics be computed. It shifts the test statistics from χ^2 and *z* statistics to *F* statistics and *t* statistics. Although the standard errors from each equation are computed using the degrees of freedom for the equation, the degrees of freedom for the *t* statistics are all taken to be those for the first equation.

dfk specifies the use of an alternative divisor in computing the covariance matrix for the equation residuals. As an asymptotically justified estimator, *sureg* by default uses the number of sample observations (*n*) as a divisor. When the *dfk* option is set, a small-sample adjustment is made and the divisor is taken to be $\sqrt{(n - k_i)(n - k_j)}$, where *k*_{*i*} and *k*_{*j*} are the number of parameters in equations *i* and *j*, respectively.

dfk2 specifies the use of an alternative divisor in computing the covariance matrix for the equation residuals. When the *dfk2* option is set, the divisor is taken to be the mean of the residual degrees of freedom from the individual equations.

Reporting

level(#); see [R] Estimation options.

corr displays the correlation matrix of the residuals between equations and performs a Breusch–Pagan test for independent equations; that is, the disturbance covariance matrix is diagonal.

nocnsreport; see [R] [Estimation options](#).

display_options: `nocl`, `nopvalues`, `noomitted`, `vsquish`, `noemptycells`, `baselevels`, `allbaselevels`, `nofvlabel`, `fvwrap(#)`, `fvwrapon(style)`, `cformat(%fmt)`, `pformat(%fmt)`, `sformat(%fmt)`, and `no1stretch`; see [R] [Estimation options](#).

Optimization

optimization_options control the iterative process that minimizes the sum of squared errors when `isure` is specified. These options are seldom used.

`iterate(#)` specifies the maximum number of iterations. When the number of iterations equals `#`, the optimizer stops and presents the current results, even if the convergence tolerance has not been reached. The default is the number set using `set maxiter`, which is 300 by default.

`trace` adds to the iteration log a display of the current parameter vector.

`log` and `nolog` specify whether to display the iteration log. The iteration log is displayed by default unless you used `set iterlog off` to suppress it; see `set iterlog` in [R] [set iter](#).

`tolerance(#)` specifies the tolerance for the coefficient vector. When the relative change in the coefficient vector from one iteration to the next is less than or equal to `#`, the optimization process is stopped. `tolerance(1e-6)` is the default.

The following options are available with `sureg` but are not shown in the dialog box:

`noheader` suppresses display of the header reporting F statistics, R^2 , and root mean squared error above the coefficient table.

`notable` suppresses display of the coefficient table.

`coeflegend`; see [R] [Estimation options](#).

Remarks and examples

Seemingly unrelated regression models are so called because they appear to be joint estimates from several regression models, each with its own error term. The regressions are related because the (contemporaneous) errors associated with the dependent variables may be correlated. Chapter 6 of [Cameron and Trivedi \(2022\)](#) contains a discussion of the seemingly unrelated regression model and the feasible generalized least-squares estimator underlying it.

▷ Example 1

When we fit models with the same set of right-hand-side variables, the seemingly unrelated regression results (in terms of coefficients and standard errors) are the same as fitting the models separately (using, say, `regress`). The same is true when the models are nested. Even in such cases, `sureg` is useful when we want to perform joint tests. For instance, let us assume that we think

$$\begin{aligned}\text{price} &= \beta_0 + \beta_1 \text{foreign} + \beta_2 \text{length} + u_1 \\ \text{weight} &= \gamma_0 + \gamma_1 \text{foreign} + \gamma_2 \text{length} + u_2\end{aligned}$$

Because the models have the same set of explanatory variables, we could estimate the two equations separately. Yet, we might still choose to estimate them with `sureg` because we want to perform the joint test $\beta_1 = \gamma_1 = 0$.

We use the small and dfk options to obtain small-sample statistics comparable with regress or mvreg.

```
. use https://www.stata-press.com/data/r19/auto
(1978 automobile data)
. sureg (price foreign length) (weight foreign length), small dfk
Seemingly unrelated regression
```

| Equation | Obs | Params | RMSE | "R-squared" | F | P>F |
|----------|-----|--------|----------|-------------|--------|--------|
| price | 74 | 2 | 2474.593 | 0.3154 | 16.35 | 0.0000 |
| weight | 74 | 2 | 250.2515 | 0.8992 | 316.54 | 0.0000 |

| | Coefficient | Std. err. | t | P> t | [95% conf. interval] | |
|---------|-------------|-----------|-------|-------|----------------------|-----------|
| price | | | | | | |
| foreign | 2801.143 | 766.117 | 3.66 | 0.000 | 1286.674 | 4315.611 |
| length | 90.21239 | 15.83368 | 5.70 | 0.000 | 58.91219 | 121.5126 |
| _cons | -11621.35 | 3124.436 | -3.72 | 0.000 | -17797.77 | -5444.93 |
| weight | | | | | | |
| foreign | -133.6775 | 77.47615 | -1.73 | 0.087 | -286.8332 | 19.4782 |
| length | 31.44455 | 1.601234 | 19.64 | 0.000 | 28.27921 | 34.60989 |
| _cons | -2850.25 | 315.9691 | -9.02 | 0.000 | -3474.861 | -2225.639 |

Note: Small-sample degrees-of-freedom adjustment applied when estimating covariance matrix of residuals.

These two equations have a common set of regressors, and we could have used a shorthand syntax to specify the equations:

```
. sureg (price weight = foreign length), small dfk
```

Here the results presented by sureg are the same as if we had estimated the equations separately:

```
. regress price foreign length
(output omitted)
. regress weight foreign length
(output omitted)
```

There is, however, a difference. We have allowed u_1 and u_2 to be correlated and have estimated the full variance–covariance matrix of the coefficients. sureg has estimated the correlations, but it does not report them unless we specify the corr option. We did not remember to specify corr when we fit the model, but we can redisplay the results:

```
. sureg, notable noheader corr
```

Note: Small-sample degrees-of-freedom adjustment applied when estimating covariance matrix of residuals.

Correlation matrix of residuals:

```
  price  weight
price  1.0000
weight 0.5840 1.0000
```

```
Breusch-Pagan test of independence: chi2(1) = 25.237, Pr = 0.0000
```

The notable and noheader options prevented sureg from redisplaying the header and coefficient tables. We find that, for the same cars, the correlation of the residuals in the price and weight equations is 0.5840 and that we can reject the hypothesis that this correlation is zero.

We can test that the coefficients on `foreign` are jointly zero in both equations—as we set out to do—by typing `test foreign`; see [R] `test`. When we type a variable without specifying the equation, that variable is tested for zero in all equations in which it appears:

```
. test foreign
( 1) [price]foreign = 0
( 2) [weight]foreign = 0
      F( 2, 142) = 17.99
      Prob > F = 0.0000
```

◀

▶ Example 2

When the models do not have the same set of explanatory variables and are not nested, `sureg` may lead to more efficient estimates than running the models separately as well as allowing joint tests. This time, let us assume that we believe

$$\begin{aligned} \text{price} &= \beta_0 + \beta_1 \text{foreign} + \beta_2 \text{mpg} + \beta_3 \text{displ} + u_1 \\ \text{weight} &= \gamma_0 + \gamma_1 \text{foreign} + \gamma_2 \text{length} + u_2 \end{aligned}$$

To fit this model, we type

```
. sureg (price foreign mpg displ) (weight foreign length), corr
Seemingly unrelated regression
```

| Equation | Obs | Params | RMSE | "R-squared" | chi2 | P>chi2 |
|----------|-----|--------|----------|-------------|--------|--------|
| price | 74 | 3 | 2165.321 | 0.4537 | 49.64 | 0.0000 |
| weight | 74 | 2 | 245.2916 | 0.8990 | 661.84 | 0.0000 |

| | Coefficient | Std. err. | z | P> z | [95% conf. interval] | |
|--------------|-------------|-----------|-------|-------|----------------------|-----------|
| price | | | | | | |
| foreign | 3058.25 | 685.7357 | 4.46 | 0.000 | 1714.233 | 4402.267 |
| mpg | -104.9591 | 58.47209 | -1.80 | 0.073 | -219.5623 | 9.644042 |
| displacement | 18.18098 | 4.286372 | 4.24 | 0.000 | 9.779842 | 26.58211 |
| _cons | 3904.336 | 1966.521 | 1.99 | 0.047 | 50.0263 | 7758.645 |
| weight | | | | | | |
| foreign | -147.3481 | 75.44314 | -1.95 | 0.051 | -295.2139 | .517755 |
| length | 30.94905 | 1.539895 | 20.10 | 0.000 | 27.93091 | 33.96718 |
| _cons | -2753.064 | 303.9336 | -9.06 | 0.000 | -3348.763 | -2157.365 |

Correlation matrix of residuals:

```
  price  weight
price  1.0000
weight 0.3285 1.0000
```

Breusch-Pagan test of independence: `chi2(1) = 7.984, Pr = 0.0047`

In comparison, if we had fit the price model separately,

```
. regress price foreign mpg displ
```

| Source | SS | df | MS | Number of obs | = | 74 |
|----------|-----------|----|------------|---------------|---|--------|
| Model | 294104790 | 3 | 98034929.9 | F(3, 70) | = | 20.13 |
| Residual | 340960606 | 70 | 4870865.81 | Prob > F | = | 0.0000 |
| | | | | R-squared | = | 0.4631 |
| | | | | Adj R-squared | = | 0.4401 |
| Total | 635065396 | 73 | 8699525.97 | Root MSE | = | 2207 |

| price | Coefficient | Std. err. | t | P> t | [95% conf. interval] | |
|--------------|-------------|-----------|-------|-------|----------------------|----------|
| foreign | 3545.484 | 712.7763 | 4.97 | 0.000 | 2123.897 | 4967.072 |
| mpg | -98.88559 | 63.17063 | -1.57 | 0.122 | -224.8754 | 27.10426 |
| displacement | 22.40416 | 4.634239 | 4.83 | 0.000 | 13.16146 | 31.64686 |
| _cons | 2796.91 | 2137.873 | 1.31 | 0.195 | -1466.943 | 7060.763 |

The coefficients are slightly different, but the standard errors are uniformly larger. This would still be true if we specified the `dfk` option to make a small-sample adjustment to the estimated covariance of the disturbances.

◀

□ Technical note

Constraints can be applied to SURE models using Stata's standard syntax for constraints. For a general discussion of constraints, see [R] [constraint](#); for examples similar to seemingly unrelated regression models, see [R] [reg3](#).

□

Stored results

sureg stores the following in `e()`:

Scalars

| | |
|--------------------------|--|
| <code>e(N)</code> | number of observations |
| <code>e(k)</code> | number of parameters |
| <code>e(k_eq)</code> | number of equations in <code>e(b)</code> |
| <code>e(mss_#)</code> | model sum of squares for equation # |
| <code>e(df_m#)</code> | model degrees of freedom for equation # |
| <code>e(rss_#)</code> | residual sum of squares for equation # |
| <code>e(df_r)</code> | residual degrees of freedom |
| <code>e(r2_#)</code> | R^2 for equation # |
| <code>e(F_#)</code> | F statistic for equation # (small only) |
| <code>e(rmse_#)</code> | root mean squared error for equation # |
| <code>e(dfk2_adj)</code> | divisor used with VCE when <code>dfk2</code> specified |
| <code>e(ll)</code> | log likelihood |
| <code>e(N_clust)</code> | number of clusters |
| <code>e(chi2_#)</code> | χ^2 for equation # |
| <code>e(p_#)</code> | p -value for equation # |
| <code>e(cons_#)</code> | 1 if equation # has a constant, 0 otherwise |
| <code>e(chi2_bp)</code> | Breusch–Pagan χ^2 |
| <code>e(df_bp)</code> | degrees of freedom for Breusch–Pagan χ^2 test |
| <code>e(rank)</code> | rank of <code>e(V)</code> |
| <code>e(ic)</code> | number of iterations |

Macros

| | |
|-------------------|--|
| e(cmd) | sureg |
| e(cmdline) | command as typed |
| e(method) | sure or isure |
| e(depvar) | names of dependent variables |
| e(exog) | names of exogenous variables |
| e(eqnames) | names of equations |
| e(wtype) | weight type |
| e(wexp) | weight expression |
| e(corr) | correlation structure |
| e(small) | small, if specified |
| e(dfk) | dfk or dfk2, if specified |
| e(clustvar) | name of cluster variable |
| e(vce) | <i>vcetype</i> specified in <i>vce()</i> |
| e(vcetype) | title used to label Std. err. |
| e(properties) | b V |
| e(predict) | program used to implement predict |
| e(marginsok) | predictions allowed by margins |
| e(marginsnotok) | predictions disallowed by margins |
| e(marginsdefault) | default <i>predict()</i> specification for margins |
| e(asbalanced) | factor variables <i>fvset</i> as asbalanced |
| e(asobserved) | factor variables <i>fvset</i> as asobserved |

Matrices

| | |
|----------|---|
| e(b) | coefficient vector |
| e(Cns) | constraints matrix |
| e(Sigma) | $\hat{\Sigma}$, covariance matrix of residuals |
| e(V) | variance–covariance matrix of the estimators |

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, <i>p</i> -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

sureg uses the asymptotically efficient, feasible, generalized least-squares algorithm described in Greene (2018, 328–339). The computing formulas are given on page 328–335.

The R^2 reported is the percent of variance explained by the predictors. It may be used for descriptive purposes, but R^2 is not a well-defined concept when GLS is used.

sureg will refuse to compute the estimators if the same equation is named more than once or the covariance matrix of the residuals is singular.

The Breusch and Pagan (1980) χ^2 statistic—a Lagrange multiplier statistic—is given by

$$\lambda = T \sum_{m=2}^M \sum_{n=1}^{m-1} r_{mn}^2$$

where r_{mn} is the estimated correlation between the residuals of the M equations and T is the number of observations. It is distributed as χ^2 with $M(M-1)/2$ degrees of freedom.

This command supports the Huber/White/sandwich estimator of the variance and its clustered version using `vce(robust)` and `vce(cluster clustvar)`, respectively. See [P] [_robust](#), particularly *Introduction* and *Methods and formulas*.

Arnold Zellner (1927–2010) was born in New York. He studied physics at Harvard and economics at Berkeley, and then he taught economics at the Universities of Washington and Wisconsin before settling in Chicago in 1966. Among his many major contributions to econometrics and statistics are his work on seemingly unrelated regression, three-stage least squares, and Bayesian econometrics.

References

- Breusch, T. S., and A. R. Pagan. 1980. The Lagrange multiplier test and its applications to model specification in econometrics. *Review of Economic Studies* 47: 239–253. <https://doi.org/10.2307/2297111>.
- Cameron, A. C., and P. K. Trivedi. 2022. *Microeconometrics Using Stata*. 2nd ed. College Station, TX: Stata Press.
- Greene, W. H. 2018. *Econometric Analysis*. 8th ed. New York: Pearson.
- McDowell, A. W. 2004. From the help desk: Seemingly unrelated regression with unbalanced equations. *Stata Journal* 4: 442–448.
- Rossi, P. E. 1989. The ET interview: Professor Arnold Zellner. *Econometric Theory* 5: 287–317. <https://doi.org/10.1017/S0266466600012445>.
- Zellner, A. 1962. An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *Journal of the American Statistical Association* 57: 348–368. <https://doi.org/10.2307/2281644>.
- . 1963. Estimators for seemingly unrelated regression equations: Some exact finite sample results. *Journal of the American Statistical Association* 58: 977–992. <https://doi.org/10.2307/2283326>.
- Zellner, A., and D. S. Huang. 1962. Further properties of efficient estimators for seemingly unrelated regression equations. *International Economic Review* 3: 300–313. <https://doi.org/10.2307/2525396>.

Also see

- [R] [sureg postestimation](#) — Postestimation tools for `sureg`
- [R] [nlsur](#) — Estimation of nonlinear system of equations
- [R] [reg3](#) — Three-stage estimation for systems of simultaneous equations
- [R] [regress](#) — Linear regression
- [MV] [mvreg](#) — Multivariate regression
- [SEM] [Example 12](#) — Seemingly unrelated regression
- [SEM] [Intro 5](#) — Tour of models
- [TS] [dfactor](#) — Dynamic-factor models
- [U] [20 Estimation and postestimation commands](#)

Stata, Stata Press, Mata, NetCourse, and NetCourseNow are registered trademarks of StataCorp LLC. Stata and Stata Press are registered trademarks with the World Intellectual Property Organization of the United Nations. StataNow is a trademark of StataCorp LLC. Other brand and product names are registered trademarks or trademarks of their respective companies. Copyright © 1985–2025 StataCorp LLC, College Station, TX, USA. All rights reserved.



For suggested citations, see the FAQ on [citing Stata documentation](#).