spivregress postestimation — Postestimation tools for spivregress

Postestimation commands predict margins estat impact
Methods and formulas References Also see

Postestimation commands

The following postestimation command is of special interest after `spivregress`:

Command	Description
`estat impact`	direct, indirect, and total impacts

The following standard postestimation commands are also available:

Command	Description
`contrast`	contrasts and ANOVA-style joint tests of parameters
`estat summarize`	summary statistics for the estimation sample
`estat vce`	variance–covariance matrix of the estimators (VCE)
`estimates`	cataloging estimation results
`etable`	table of estimation results
`lincom`	point estimates, standard errors, testing, and inference for linear combinations of parameters
`margins`	marginal means, predictive margins, marginal effects, and average marginal effects
`marginsplot`	graph the results from margins (profile plots, interaction plots, etc.)
`nlcom`	point estimates, standard errors, testing, and inference for nonlinear combinations of parameters
`predict`	predictions, residuals, influence statistics, and other diagnostic measures
`predictnl`	point estimates, standard errors, testing, and inference for generalized predictions
`pwcompare`	pairwise comparisons of parameters
`test`	Wald tests of simple and composite linear hypotheses
`testnl`	Wald tests of nonlinear hypotheses

predict

Description for predict

`predict` creates a new variable containing predictions such as the reduced-form mean, the direct mean, the indirect mean, the limited-information mean, the full-information mean, the naïve-form prediction, the linear prediction, the residuals, or the uncorrelated residuals.

Menu for predict

Statistics > Postestimation

Syntax for predict

`predict` [*type*] *newvar* [*if*] [*in*] [, *statistic*]

statistic	Description
Main	
`rform`	reduced-form mean; the default
`direct`	direct mean
`indirect`	indirect mean
`limited`	limited-information mean
`full`	full-information mean
`naive`	naïve-form prediction
`xb`	linear prediction
`residuals`	residuals
`ucresiduals`	uncorrelated residuals

These statistics are only available in a subset of the estimation sample.

Options for predict

Main

`rform`, the default, calculates the reduced-form mean. It is the predicted mean of the dependent variable conditional on the independent variables and any spatial lags of the independent variables. See *Methods and formulas*.

`direct` calculates the direct mean. It is a unit's predicted contribution to its own reduced-form mean. The direct and indirect means sum to the reduced-form mean.

`indirect` calculates the indirect mean. It is the predicted sum of the other units' contributions to a unit's reduced-form mean.

`limited` calculates the limited-information mean. It is the predicted mean of the dependent variable conditional on the independent variables, any spatial lags of the independent variables, and any spatial lags of the dependent variable. `limited` is not available when the `heteroskedastic` option is used with `spivregress`.

`full` calculates the full-information mean. It is the predicted mean of the dependent variable conditional on the independent variables, any spatial lags of the independent variables, and the other units' values of the dependent variable. `full` is not available when the `heteroskedastic` option is used with `spivregress`.

`naive` calculates the naïve-form prediction. It is the predicted linear combination of the independent variables, any spatial lags of the independent variables, and any spatial lags of the dependent variable. It is not a consistent estimator of an expectation. See *Methods and formulas*.

`xb` calculates the predicted linear combination of the independent variables.

`residuals` calculates the residuals, including any autoregressive error term.

`ucresiduals` calculates the uncorrelated residuals, which are estimates of the uncorrelated error term.

margins

Description for margins

`margins` estimates margins of response for reduced-form mean, direct mean, indirect mean, and linear predictions.

Menu for margins

Statistics > Postestimation

Syntax for margins

`margins` [*marginlist*] [, *options*]

`margins` [*marginlist*] , `predict(`*statistic* `...)` [`predict(`*statistic* `...)` `...`] [*options*]

statistic	Description
`rform`	reduced-form mean; the default
`direct`	direct mean
`indirect`	indirect mean
`xb`	linear prediction
`limited`	not allowed with `margins`
`full`	not allowed with `margins`
`naive`	not allowed with `margins`
`residuals`	not allowed with `margins`
`ucresiduals`	not allowed with `margins`

Statistics not allowed with `margins` are functions of stochastic quantities other than `e(b)`.

For the full syntax, see [R] **margins**.

Remarks for margins

The computations that `margins` must do to calculate standard errors can sometimes be time consuming. Time will depend on the complexity of the spatial model and the number of spatial units in the data. You may want to fit your model with a subsample of your data, run `margins`, and extrapolate to estimate the time required to run `margins` on the full sample. See [P] **timer** and [P] **rmsg**.

estat impact

Description for estat impact

`estat impact` estimates the mean of the direct, indirect, and total impacts of independent variables on the reduced-form mean of the dependent variable.

Syntax for estat impact

`estat impact` [*varlist*] [*if*] [*in*] [, `nolog` `vce(`*vcetype*`)`]

varlist is a list of independent variables, including factor variables, taken from the fitted model. By default, all independent variables from the fitted model are used.

`collect` is allowed; see **[U] 11.1.10 Prefix commands**.

Options for estat impact

Main

`nolog` suppresses the calculation progress log that shows the percentage completed. By default, the log is displayed.

VCE

`vce(`*vcetype*`)` specifies how the standard errors of the impacts are calculated.

`vce(delta)`, the default, is the delta method and treats the independent variables as fixed.

`vce(unconditional)` specifies that standard errors account for sampling variance in the independent variables. This option is not available when `if` or `in` is specified with `estat impact`.

Remarks for estat impact

`estat impact` is essential for interpreting the output of `spivregress`. See [SP] **Intro 7** and example 1 of [SP] **spregress** for explanations and examples.

Stored results for estat impact

`estat impact` stores the following in `r()`:

Scalars

`r(N)`	number of observations

Macros

`r(vce)`	*vcetype* specified in `vce()`
`r(xvars)`	names of independent variables

Matrices

`r(b_direct)`	vector of estimated direct impacts
`r(Jacobian_direct)`	Jacobian matrix for direct impacts
`r(V_direct)`	estimated variance–covariance matrix of direct impacts
`r(b_indirect)`	vector of estimated indirect impacts
`r(Jacobian_indirect)`	Jacobian matrix for indirect impacts
`r(V_indirect)`	estimated variance–covariance matrix of indirect impacts
`r(b_total)`	vector of estimated total impacts
`r(Jacobian_total)`	Jacobian matrix for total impacts
`r(V_total)`	estimated variance–covariance matrix of total impacts

Methods and formulas

Methods and formulas are presented under the following headings:

Predictions
- *Reduced-form mean*
- *Direct and indirect means*
- *Limited-information mean*
- *Full-information mean*
- *Naïve-form predictor*
- *Linear predictor*
- *Residuals*
- *Uncorrelated residuals*

Impacts

Predictions

To motivate the predictions, consider the vector form of a spatial autoregressive model

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

where

$\mathbf{y}$ is the vector containing each unit's dependent-variable observation,

$\mathbf{W}\mathbf{y}$ is a spatial lag of $\mathbf{y}$,

$\mathbf{X}$ is the matrix of independent-variable observations,

$\boldsymbol{\epsilon}$ is a vector of errors, and

λ and $\boldsymbol{\beta}$ are the coefficients.

Any spatial lags of the independent variables are assumed to be in $\mathbf{X}$. Spatial lags of the error do not affect the reduced-form, direct, or indirect means, so they are not included in (1) for simplicity.

Reduced-form mean

Equation (1) represents the spatial autoregressive model as a system of equations. The solution

$$\mathbf{y} = (\mathbf{I} - \lambda\mathbf{W})^{-1}(\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}) \tag{2}$$

implies that the mean of $\mathbf{y}$ given the independent variables and the spatial weighting matrix is

$$E(\mathbf{y} \,|\, \mathbf{X}, \mathbf{W}) = (\mathbf{I} - \lambda\mathbf{W})^{-1}(\mathbf{X}\boldsymbol{\beta}) \tag{3}$$

This is known as the reduced-form mean because the solution in (2) is known as the reduced form of the model. The predicted reduced-form mean substitutes estimates of λ and $\boldsymbol{\beta}$ into (3).

Direct and indirect means

To define the direct mean and the indirect mean, let

$$\mathbf{S} = (\mathbf{I} - \lambda\mathbf{W})^{-1}$$

and let $\mathbf{S}_d$ be a matrix with diagonal elements of $\mathbf{S}$ on its diagonal and with off-diagonal elements set to 0.

The direct means are

$$\mathbf{S}_d\mathbf{X}\boldsymbol{\beta}$$

which capture the contributions of each unit's independent variables on its own reduced-form mean. Substituting estimates of λ and $\boldsymbol{\beta}$ produces the predictions.

The indirect means capture the contributions of the other units' independent variables on a unit's reduced-form prediction, and they are

$$\left\{(\mathbf{I} - \lambda\mathbf{W})^{-1} - \mathbf{S}_d\right\}\mathbf{X}\boldsymbol{\beta}$$

Limited-information mean

Instead of solving for the reduced form, the limited-information mean conditions on the spatial lag of y for observation i, which we denote by $(\mathbf{W}\,\mathbf{y})_i$, which yields

$$E\{y_i \,|\, \mathbf{X}, \mathbf{W}, (\mathbf{W}\,\mathbf{y})_i\} = \mathbf{x}_i\boldsymbol{\beta} + \lambda(\mathbf{W}\,\mathbf{y})_i + u_i \tag{4}$$

where u_i is the predictable part of the error term given $(\mathbf{W}\,\mathbf{y})_i$. See Kelejian and Prucha (2007) and Drukker, Prucha, and Raciborski (2013).

Full-information mean

The full-information mean conditions on the dependent-variable values of all the other units instead of conditioning on the spatial lag of the dependent variable, as does the limited-information mean. The additional information produces a better prediction of the error term when a spatial lag of the errors is in the model. See Kelejian and Prucha (2007).

Naïve-form predictor

The naïve-form predictor sets u_i to 0 in (4). It is not consistent for $E\{y_i \,|\, \mathbf{X}, \mathbf{W}, (\mathbf{W}\,\mathbf{y})_i\}$ because it ignores u_i.

Linear predictor

The linear predictor is $\mathbf{X}\beta$.

Residuals

The residuals are u_i from (4).

Uncorrelated residuals

The uncorrelated residuals are

$$\hat{\boldsymbol{\epsilon}} = (\mathbf{I} - \hat{\rho}\,\mathbf{M})^{-1}\,\mathbf{u}$$

where $\mathbf{u}$ is the vector of u_i's, $\mathbf{M}$ is the spatial weighting matrix for the autoregressive error term, and $\hat{\rho}$ is the estimated correlation of $\mathbf{u}$.

Impacts

The total impact of an independent variable $\mathbf{x}$ is the average of the marginal effects it has on the reduced-form mean,

$$\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial E(\mathbf{y}_i \,|\, \mathbf{X}, \mathbf{W})}{\partial x_j}$$

where $E(\mathbf{y}_i \,|\, \mathbf{X}, \mathbf{W})$ is the ith element of the vector $E(\mathbf{y} \,|\, \mathbf{X}, \mathbf{W})$, whose formula is given in (2), and x_j is the jth unit's value for $\mathbf{x}$.

The direct impact of an independent variable $\mathbf{x}$ is the average of the direct, or own, marginal effects

$$\frac{1}{n}\sum_{i=1}^{n}\frac{\partial E(\mathbf{y}_i \,|\, \mathbf{X}, \mathbf{W})}{\partial x_i}$$

The indirect impact of an independent variable $\mathbf{x}$ is the average of the indirect, or spillover, marginal effects.

$$\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n}\frac{\partial E(\mathbf{y}_i \,|\, \mathbf{X}, \mathbf{W})}{\partial x_j}$$

LeSage and Pace (2009, 36–37) call the average direct impact the "average total direct impact", and they call the average indirect impact the "average total indirect impact".

`estat impact` with the default `vce(delta)` uses the delta method to calculate the estimated variance of the impacts. This variance is conditional on the values of the independent variables in the model.

`estat impact` with `vce(unconditional)` uses the generalized method of moments estimation strategy to estimate the unconditional variance of the impacts. It accounts for sampling variance of the independent variables in the model.

References

Drukker, D. M., I. R. Prucha, and R. Raciborski. 2013. Maximum likelihood and generalized spatial two-stage least-squares estimators for a spatial-autoregressive model with spatial-autoregressive disturbances. *Stata Journal* 13: 221–241.

Kelejian, H. H., and I. R. Prucha. 2007. The relative efficiencies of various predictors in spatial econometric models containing spatial lags. *Regional Science and Urban Economics* 37: 363–374. https://doi.org/10.1016/j.regsciurbeco.2006.11.005.

LeSage, J., and R. K. Pace. 2009. *Introduction to Spatial Econometrics.* Boca Raton, FL: Chapman and Hall/CRC. https://doi.org/10.1201/9781420064254.

Liu, D. 2017. How to create animated graphics to illustrate spatial spillover effects. *The Stata Blog: Not Elsewhere Classified.* https://blog.stata.com/2018/03/06/how-to-create-animated-graphics-to-illustrate-spatial-spillover-effects/.

Also see

[SP] **spivregress** — Spatial autoregressive models with endogenous covariates

[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.