Title

stata.com

Intro 8 — The Sp estimation commands

Description Remarks and examples Reference Also see

Description

There are three Sp estimation commands for spatial data:

- `spregress`—linear regression for cross-sectional data
- `spivregress`—instrumental-variables linear regression for cross-sectional data
- `spxtregress`—fixed- and random-effects linear regression models for panel data

This entry provides an overview of these estimation commands.

You may also be interested in introductions to other aspects of Sp. Below, we provide links to those other introductions.

Intro 1	A brief introduction to SAR models
Intro 2	The **W** matrix
Intro 3	Preparing data for analysis
Intro 4	Preparing data: Data with shapefiles
Intro 5	Preparing data: Data containing locations (no shapefiles)
Intro 6	Preparing data: Data without shapefiles or locations
Intro 7	Example from start to finish

Remarks and examples

stata.com

Remarks are presented under the following headings:

spregress, gs2sls
spregress, ml
spivregress
spxtregress
spxtregress, re
spxtregress, fe

spregress, gs2sls

`spregress` is the equivalent of `regress` for spatial data. You have two choices of estimator: `gs2sls` or `ml`.

The `gs2sls` estimator is a generalized method of moments estimator. With `gs2sls`, you can fit multiple spatial lags of the dependent variable (that is, multiple spatial weighting matrices), multiple spatial autoregressive error terms, and multiple spatial lags of covariates. To fit a model, you issue a command like

```
spregress y x1 x2, gs2sls dvarlag(W) errorlag(W) ivarlag(M: x1 x2)
```

where `W` and `M` are weighting matrices. See [SP] **spregress**.

To interpret your results after fitting the model, it is essential that you run `estat impact`. `estat impact` works after all the Sp estimation commands. Explanations and examples are given in [SP] **Intro 7**, example 1 of [SP] **spregress**, [SP] **spivregress postestimation**, [SP] **spregress postestimation**, and [SP] **spxtregress postestimation**.

The `gs2sls` estimator assumes that the errors are independent and identically distributed (i.i.d.) but does not require normality. The i.i.d. requirement is relaxed when you use the `heteroskedastic` option; only independence is required.

```
spregress y x1 x2, gs2sls heteroskedastic dvarlag(W) errorlag(W) ///
   ivarlag(M: x1 x2)
```

The `heteroskedastic` option uses different formulas for the spatial autoregressive error correlations and the standard errors. See *Methods and formulas* in [SP] **spregress**.

spregress, ml

The `spregress, ml` estimator is a maximum likelihood (ML) estimator. With `ml`, you can fit only one spatial lag of the dependent variable and only one spatial autoregressive error term, but you can fit multiple spatial lags of covariates. To fit a model, type

```
spregress y x1 x2, ml dvarlag(W) errorlag(W) ivarlag(W: x1 x2) ///
   ivarlag(M: x1 x2)
```

The `ml` estimator assumes that the errors are normal and i.i.d. The command `spregress, ml` is typically slower than `spregress, gs2sls`, but `spregress, ml` may be more efficient (smaller standard errors) when errors are normal.

The requirement of normality is removed if you use the `vce(robust)` option, just as it is for Stata's other ML estimators that allow this option:

```
spregress y x1 x2, ml vce(robust) dvarlag(W) errorlag(W) ///
     ivarlag(M: x1 x2)
```

See *Methods and formulas* in [SP] **spregress**.

spivregress

`spivregress` is the equivalent of `ivregress` for spatial data. `spivregress` uses the same estimator as `spregress, gs2sls`, but it allows endogenous regressors. You can fit multiple spatial lags of the dependent variable, multiple spatial autoregressive error terms, and multiple spatial lags of included exogenous regressors. You cannot specify a spatial lag for the endogenous regressors or for the excluded exogenous regressors. See *Remarks and examples* in [SP] **spivregress**.

To fit a model using `spivregress`, you would issue a command like

```
spivregress y x1 x2 (z = x3), dvarlag(W) errorlag(W) ivarlag(M: x1 x2)
```

`spivregress` also has a `heteroskedastic` option that provides the same properties it does when used with `spregress, gs2sls`.

spxtregress

`spxtregress` is the Sp estimation command for panel data. It fits fixed-effects (`fe`) and random-effects (`re`) models. `spxtregress, fe` and `re` are the spatial data equivalent of `xtreg, fe` and `re`. To use `spxtregress`, you must have strongly balanced data, and your data must be `xtset`. See [SP] **Intro 3**, [SP] **Intro 7**, and [SP] **spbalance**.

With `spxtregress, fe` and `re`, you can fit only one spatial lag of the dependent variable and only one spatial autoregressive error term. You can fit multiple spatial lags of covariates.

spxtregress, re

The random-effects model is fit using a maximum likelihood estimator. It assumes that the panel-level effects are normal i.i.d. across the panels and that the errors are normal i.i.d. across panels and time.

To fit this model, you issue a command like

```
spxtregress y x1 x2, re dvarlag(W) errorlag(W) ivarlag(M: x1 x2)
```

`spxtregress, re` has a `sarpanel` option that uses a different formulation of the random-effects estimator due to Kapoor, Kelejian, and Prucha (2007). The panel-level effects are considered a disturbance in the error equation, and the panel-level effects have the same autoregressive form as the time-level errors. To fit such models, you issue a command like

```
spxtregress y x1 x2, re sarpanel dvarlag(W) errorlag(W) ///
   ivarlag(M: x1 x2)
```

spxtregress, fe

The fixed-effects model also uses a maximum likelihood estimator. In this estimator, panel effects and effects that are constant within time are conditioned out of the likelihood. No distributional assumptions are made about the panel effects. Only covariates that vary across both panels and time can be fit with this estimator.

To fit this model, you issue a command like

```
spxtregress y x1 x2, fe dvarlag(W) errorlag(W) ivarlag(M: x1 x2)
```

See *Methods and formulas* in [SP] **spxtregress**.

Reference

Kapoor, M., H. H. Kelejian, and I. R. Prucha. 2007. Panel data models with spatially correlated error components. *Journal of Econometrics* 140: 97–130. https://doi.org/10.1016/j.jeconom.2006.09.004.

Also see

[SP] **Intro** — Introduction to spatial data and SAR models

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

[SP] **spregress** — Spatial autoregressive models

[SP] **spxtregress** — Spatial autoregressive models for panel data