Cross-sectional and spatial dependence in panels

Giovanni Millo

March 31, 2008

Abstract

Econometricians have recently turned towards the problems posed by cross-sectional dependence across individuals, which may range from inefficiency of the standard estimators and invalid inference to inconsistency. Panel data are especially useful in this respect, as their double dimensionality allows robust approaches to general cross-sectional dependence.

A general object oriented approach to robust inference is available in the R system (Zeileis, 2004), for which all that's needed are coefficients $\hat{\beta}$ and robust estimators for $vcov(\hat{\beta})$. An useful implementation is, e.g., in linear hypotheses testing (see Fox, package `car`). The `plm` package for pael data econometrics already has features for heteroskedasticity– and serial correlation–robust inference (Croissant and Millo, forthcoming).

If cross-sectional dependence is detected, using a robust covariance estimator allows valid inference. I describe the implementation in the `plm` package for panel data econometrics of:

- tests for detecting cross-sectional dependence in the errors of a panel model (Friedman 1928, Frees 1995, Pesaran 2004)
- robust estimators of covariance matrices for doing valid inference in the presence of cross-sectional dependence (White 1980, Beck and Katz 1995, Driscoll and Kraay 1998)

If a particular spatial structure is assumed, this allows a parsimonious characterization of spatial dependence but, on the converse, the resulting models are computationally expensive to estimate, all the more so in the panel case. Efficient ML estimators for spatial models on a cross-section (Anselin 1988) are implemented in the `spdep` package (Bivand et al.). I describe implementation in a forthcoming package of

- marginal and conditional LM tests for spatial correlation, serial correlation and random effects (Baltagi, Song, Jung and Koh 2007)
- ML estimators for panel models including spatial lags, spatial errors and possibly serial correlation (Anselin 1988, Elhorst 2003, Baltagi, Song, Jung and Koh 2007)

I illustrate the functionalities by application to Munnell's (1990) data on 48 USA states observed over 17 years. On an ordinary desktop machine, the estimators and tests all take under one minute (few seconds for the basic ones). The ML approach is nevertheless structurally limited to a few hundred cross-sectional observations, so further work is warranted to implement Kapoor, Kelejian and Prucha (2007)'s GM approach, which promises to handle problems with n in the thousands.