1-2011

Instrumental Variable Estimation of a Spatial Autoregressive Panel Model with Random Effects

Badi H. Baltagi
Syracuse University

Long Liu
University of Texas at San Antonio

ISSN: 1525-3066

Center for Policy Research
Working Paper No. 127

INSTRUMENTAL VARIABLE ESTIMATION OF A SPATIAL AUTOREGRESSIVE PANEL MODEL WITH RANDOM EFFECTS

BADI H. BALTAGI AND LONG LIU

Center for Policy Research
Maxwell School of Citizenship and Public Affairs
Syracuse University
426 Eggers Hall
Syracuse, New York 13244-1020
(315) 443-3114 | Fax (315) 443-1081
e-mail: ctrpol@syr.edu

January 2011

$5.00

Up-to-date information about CPR's research projects and other activities is available from our World Wide Web site at **www-cpr.maxwell.syr.edu.** All recent working papers and Policy Briefs can be read and/or printed from there as well.

CENTER FOR POLICY RESEARCH – Spring 2011

Christine L. Himes, Director
Maxwell Professor of Sociology

Associate Directors

Margaret Austin
Associate Director
Budget and Administration

Douglas Wolf
Gerald B. Cramer Professor of Aging Studies
Associate Director, Aging Studies Program

John Yinger
Professor of Economics and Public Administration
Associate Director, Metropolitan Studies Program

SENIOR RESEARCH ASSOCIATES

Badi Baltagi ..... Economics
Robert Bifulco ..... Public Administration
Leonard Burman ..Public Administration/Economics
Kalena Cortes ..... Education
Thomas Dennison ..... Public Administration
William Duncombe ..... Public Administration
Gary Engelhardt ..... Economics
Madonna Harrington Meyer ..... Sociology
William C. Horrace ..... Economics
Duke Kao ..... Economics
Eric Kingson ..... Social Work
Sharon Kioko ..... Public Administration
Thomas Kniesner ..... Economics
Jeffrey Kubik ..... Economics
Andrew London ..... Sociology
Len Lopoo ..... Public Administration
Amy Lutz ..... Sociology
Jerry Miner ..... Economics
Jan Ondrich ..... Economics
John Palmer ..... Public Administration
David Popp ..... Public Administration
Gretchen Purser ..... Sociology
Christopher Rohlfs ..... Economics
Stuart Rosenthal ..... Economics
Ross Rubenstein ..... Public Administration
Perry Singleton ..... Economics
Margaret Usdansky ..... Sociology
Michael Wasylenko ..... Economics
Jeffrey Weinstein ..... Economics
Janet Wilmoth ..... Sociology

GRADUATE ASSOCIATES

Charles Alamo ..... Public Administration
Kanika Arora ..... Public Administration
Samuel Brown ..... Public Administration
Christian Buerger ..... Public Administration
Il Hwan Chung ..... Public Administration
Alissa Dubnicki ..... Economics
Andrew Friedson ..... Economics
Virgilio Galdo ..... Economics
Jenna Harkabus ..... Public Administration
Clorise Harvey ..... Public Administration
Biff Jones ..... Public Administration
Hee Seung Lee ..... Public Administration
Chong Li ..... Economics
Jing Li ..... Economics
Allison Marier ..... Economics
Qing Miao ..... Public Administration
Wael Moussa ..... Economics
Casey Muhm ..... Public Administration
Kerri Raissian ..... Public Administration
Morgan Romine ..... Public Administration
Amanda Ross ..... Economics
Natalee Simpson ..... Sociology
Liu Tian ..... Public Administration
Ryan Yeung ..... Public Administration

STAFF

Kelly Bogart ..... Administrative Secretary
Martha Bonney ..... Publications/Events Coordinator
Karen Cimilluca ..... Office Coordinator
Kitty Nasto ..... Administrative Secretary
Candi Patterson ..... Computer Consultant
Roseann Presutti ..... Administrative Secretary
Mary Santy ..... Administrative Secretary

Abstract

This paper extends the instrumental variable estimators of Kelejian and Prucha (1998) and Lee (2003) proposed for the cross-sectional spatial autoregressive model to the random effects spatial autoregressive panel data model. It also suggests an extension of the Baltagi (1981) error component 2SLS estimator to this spatial panel model.

JEL No. C13, C21

Key Words: Panel Data; Spatial Model; Two Stage Least Squares; Error Components

Address correspondence to: Badi H. Baltagi, Center for Policy Research, 426 Eggers Hall, Syracuse University, Syracuse, NY 13244-1020; e-mail: bbaltagi@maxwell.syr.edu.

Long Liu: Department of Economics, College of Business, University of Texas at San Antonio, One UTSA Circle, TX; 78249-0633; e-mail: long.liu@utsa.edu.

Instrument Variable Estimation of a Spatial Autoregressive Panel Model with Random Effects

Badi H. Baltagi*, Long Liu†

December 22, 2010

Abstract

This paper extends the instrumental variable estimators of Kelejian and Prucha (1998) and Lee (2003) proposed for the cross-sectional spatial autoregressive model to the random effects spatial autoregressive panel data model. It also suggests an extension of the Baltagi (1981) error component 2SLS estimator to this spatial panel model.

Key Words: *Panel Data; Spatial Model; Two Stage Least Squares; Error Components*

JEL: C13,C21

1 Model and Results

Consider the following autoregressive spatial panel data model:

$$\begin{aligned} y &= \lambda W y + X\beta + u, \\ u &= Z_\mu \mu + \nu, \end{aligned} \tag{1}$$

where y is of dimension $NT \times 1$, X is $NT \times k$, β is $k \times 1$ and u is $NT \times 1$, see Anselin (1988), Anselin, Le Gallo and Jayet (2008), and Elhorst (2003) to mention a few. The observations are ordered with t being the slow running index and i the fast running index, i.e., $y' = (y_{11}, \ldots, y_{1N}, \ldots, y_{T1}, \ldots, y_{TN})$. X is assumed to be of full column rank and its elements are assumed to be asymptotically bounded in absolute value. The error component structure is given by the second equation with $Z_\mu = \iota_T \otimes I_N$ denoting the selector matrix for the $(N \times 1)$ random vector of individual effects μ which is assumed to be i.i.d. $(0, \sigma_\mu^2 I_N)$. Here ι_T is a vector of ones of dimension T and I_N is an identity matrix of dimension N. ν is a vector of $NT \times 1$ remainder disturbances which is assumed to be i.i.d. $(0, \sigma_\nu^2 I_{NT})$. Also, μ and ν are independent of each other and the regressor matrix X.

*Address correspondence to: Badi H. Baltagi, Center for Policy Research, 426 Eggers Hall, Syracuse University, Syracuse, NY 13244-1020; e-mail: bbaltagi@maxwell.syr.edu.

†Long Liu: Department of Economics, College of Business, University of Texas at San Antonio, One UTSA Circle, TX 78249-0633; e-mail: long.liu@utsa.edu.

Ordering the data first by time (with index $t = 1, ..., T$) and then by individual units (with index $i = 1, ..., N$), we get $W = I_T \otimes W_N$, where the $(N \times N)$ spatial weight matrix W_N has zero diagonal elements and is row-normalized with its entries usually declining with distance. This in turn results in row and column sums of W_N that are uniformly bounded in absolute value. In addition the column sums of W_N, as well as the row and column sums of $(I_N - \lambda W_N)^{-1}$ are bounded uniformly in absolute value by some finite constant. We also assume that λ is bounded in absolute value, i.e., $|\lambda| < 1$. This is a panel data version of the cross-section spatial autoregressive model considered by Kelejian, and Prucha (1998) and Lee (2003). Let $A = I_T \otimes A_N$ where $A_N = I_N - \lambda W_N$, then one can write

$$y = A^{-1} (X\beta + u).$$

Note that

$$E[Wyu'] = E\left[WA^{-1}(X\beta + u) u'\right] = WA^{-1}\Omega \neq 0,$$

where $\Omega = E(uu')$. The spatially lagged dependent variable Wy is correlated with the disturbance u. Therefore, the Ordinary Least Squares estimator will be inconsistent. Let $Z = (X, Wy)$ and $\delta = (\beta, \lambda)'$, the model in Equation (1) can be written as

$$y = Z\delta + u. \tag{2}$$

In the cross-section spatial autoregressive model, Kelejian, and Prucha (1998) suggested a two-stage least square estimator (2SLS) based on feasible instruments like $H = \left(X, WX, W^2X\right)$ which yield:

$$\hat{\delta}_{2SLS} = [Z'P_H Z]^{-1} Z'P_H y, \tag{3}$$

where $P_H = H(H'H)^{-1} H'$ denotes the projection matrix using H. Actually, H could include additional higher order terms like W^3X, W^4X, etc., see also Kelejian, Prucha and Yuzefovich (2004). Kelejian, and Prucha (1998) show that this 2SLS estimator is consistent under some general regularity conditions for this model. Note that the results derived in this paper carry through when H includes these higher order terms. Let $\bar{J}_T = J_T/T$, where J_T is a matrix of ones of dimension T. Also, let $E_T = I_T - \bar{J}_T$, and define P to be the projection matrix on Z_μ, i.e., $P = \bar{J}_T \otimes I_N$, and $Q = I_{NT} - P = E_T \otimes I_N$. Premultiply Equation (2) by Q to obtain

$$\widetilde{y} = \lambda W\widetilde{y} + \widetilde{X}\beta + \widetilde{u}. \tag{4}$$

This follows from the fact that $QW = (E_T \otimes I_N)(I_T \otimes W_N) = (E_T \otimes W_N) = (I_T \otimes W_N)(E_T \otimes I_N) = WQ$. Applying the Kelejian, and Prucha (1998) 2SLS to this Q transformed panel autoregressive spatial model, we get the fixed effects spatial 2SLS (FE-S2SLS) estimator of δ based upon $\widetilde{H} = \left(\widetilde{X}, W\widetilde{X}, W^2\widetilde{X}\right) =$

$\left(QX, QWX, QW^2X\right) = QH$. We denote this by $\hat{\delta}_{FE-S2SLS}$. Note that the s^2 of this FE-S2SLS provides a consistent estimate $\hat{\sigma}_\nu^2$ of σ_ν^2. Similarly, one could premultiply Equation (2) by P to obtain

$$\overline{y} = \lambda W \overline{y} + \overline{X}\beta + \overline{u}. \tag{5}$$

This follows from the fact that $PW = (\bar{J}_T \otimes I_N)(I_T \otimes W_N) = (\bar{J}_T \otimes W_N) = (I_T \otimes W_N)(\bar{J}_T \otimes I_N) = WP$. Applying the Kelejian, and Prucha (1998) 2SLS to this P transformed panel autoregressive spatial model, we get the between effects spatial 2SLS (BE-S2SLS) estimator of δ based upon $\overline{H} = \left(\overline{X}, W\overline{X}, W^2\overline{X}\right) = \left(PX, PWX, PW^2X\right) = PH$. We denote this by $\hat{\delta}_{BE-S2SLS}$. Note that the s^2 of this BE-S2SLS provides a consistent estimate $\hat{\sigma}_1^2$ of $\sigma_1^2 = T\sigma_\mu^2 + \sigma_\nu^2$. As shown in Baltagi (2008), the variance-covariance matrix of u is $\Omega = E\left(uu'\right) = \sigma_1^2 P + \sigma_\nu^2 Q$, and $\Omega^{-1/2} = \left(\sigma_1^{-1}P + \sigma_\nu^{-1}Q\right)$. Left multiply Equation (2) by $\Omega^{-1/2}$, we get

$$y^* = Z^*\delta + u^*, \tag{6}$$

where $y^* = \Omega^{-1/2}y$, $u^* = \Omega^{-1/2}u$, and $Z^* = \Omega^{-1/2}Z = \Omega^{-1/2}(X, Wy) = \left(X^*, \Omega^{-1/2}Wy\right) = \left(X^*, W\Omega^{-1/2}y\right) = (X^*, Wy^*)$. The last expression follows from the fact that $\Omega^{-1/2}W = \left(\sigma_1^{-1}PW + \sigma_\nu^{-1}QW\right) = \left(\sigma_1^{-1}WP + \sigma_\nu^{-1}WQ\right) = W\Omega^{-1/2}$. This means that Equation (6) can also be written as:

$$y^* = \lambda W y^* + X^*\beta + u^*. \tag{7}$$

Applying the Kelejian, and Prucha (1998) 2SLS to this $\Omega^{-1/2}$ transformed panel autoregressive spatial model, we get the random effects spatial two-stage least square estimator (RE-S2SLS) of δ given by

$$\hat{\delta}_{RE-S2SLS} = \left[Z^{*\prime}P_{H^*}Z^*\right]^{-1} Z^{*\prime}P_{H^*}y^* \tag{8}$$

where $H^* = \left(X^*, WX^*, W^2X^*\right) = \left(\Omega^{-1/2}X, W\Omega^{-1/2}X, W^2\Omega^{-1/2}X\right) = \left(\Omega^{-1/2}X, \Omega^{-1/2}WX, \Omega^{-1/2}W^2X\right) = \Omega^{-1/2}H$.

Using the results in Baltagi (1981), one can similarly derive a spatial error component two-stage least square estimators (SEC-2SLS) as follows: Left multiply Equation (4) by $\tilde{H}'$, and (5) by $\bar{H}'$, and stack the two equation system recognizing that they estimate the same δ, we get:

$$\begin{pmatrix} \tilde{H}'\widetilde{y} \\ \bar{H}'\overline{y} \end{pmatrix} = \begin{pmatrix} \tilde{H}'\widetilde{Z} \\ \bar{H}'\overline{Z} \end{pmatrix} \delta + \begin{pmatrix} \tilde{H}'\widetilde{u} \\ \bar{H}'\overline{u} \end{pmatrix},$$

where $E\begin{pmatrix} \tilde{H}'\widetilde{u} \\ \bar{H}'\overline{u} \end{pmatrix} = 0$ and $Var\begin{pmatrix} \tilde{H}'\widetilde{u} \\ \bar{H}'\overline{u} \end{pmatrix} = \begin{bmatrix} \sigma_\nu^2\tilde{H}'\tilde{H} & 0 \\ 0 & \sigma_1^2\bar{H}'\bar{H} \end{bmatrix}$. This follows from the fact that $Q\Omega = \sigma_\nu^2 Q$, and $P\Omega = \sigma_1^2 P$, with $QP = 0$. Performing GLS on this two equation system we get the spatial error component two-stage least squares (SEC-2SLS) estimator:

$$\hat{\delta}_{SEC-2SLS} = \left(\frac{\widetilde{Z}'P_{\tilde{H}}\widetilde{Z}}{\sigma_\nu^2} + \frac{\overline{Z}'P_{\bar{H}}\overline{Z}}{\sigma_1^2}\right)^{-1}\left(\frac{\widetilde{Z}'P_{\tilde{H}}\widetilde{y}}{\sigma_\nu^2} + \frac{\overline{Z}'P_{\bar{H}}\overline{y}}{\sigma_1^2}\right) \tag{9}$$

$$= \left(\sigma_\nu^{-2}Z'P_{\tilde{H}}Z + \sigma_1^{-2}Z'P_{\bar{H}}Z\right)^{-1}\left(\sigma_\nu^{-2}Z'P_{\tilde{H}}y + \sigma_1^{-2}Z'P_{\bar{H}}y\right), \tag{10}$$

using the fact that Q and P are idempotent. Using a similar argument as in Baltagi (1981) one can show that $\hat{\delta}_{SEC-2SLS}$ is a matrix weighted combination of $\hat{\delta}_{FE-S2SLS}$ and $\hat{\delta}_{BE-S2SLS}$ weighting each by the inverse of their variance-covariance matrix and such that the weights add up to the identity matrix. A feasible SEC-2SLS estimator can be obtained by replacing $\hat{\sigma}_\nu^2$ and $\hat{\sigma}_1^2$ into Equation (9). Following a similar argument as in Cornwell, Schmidt and Wyhowski (1992), one can show that $\hat{\delta}_{SEC-2SLS}$ can also be obtained as 2SLS from the $\Omega^{-1/2}$ transformed equation in (6) using $B = \left(\tilde{H}, \bar{H}\right)$ as instruments. To show this, note that $P_B = P_{\tilde{H}} + P_{\bar{H}}$ using the fact that $\tilde{H}$ and $\bar{H}$ are orthogonal to each other. This means that $Z^{*\prime}P_BZ^* = Z^{*\prime}P_{\tilde{H}}Z^* + Z^{*\prime}P_{\bar{H}}Z^* = \sigma_\nu^{-2}Z'P_{\tilde{H}}Z + \sigma_1^{-2}Z'P_{\bar{H}}Z$, since $Q\Omega^{-1/2} = \sigma_\nu^{-1}Q$, and $P\Omega^{-1/2} = \sigma_1^{-1}P$. Similarly, $Z^{*\prime}P_By^* = Z^{*\prime}P_{\tilde{H}}y^* + Z^{*\prime}P_{\bar{H}}y^* = \sigma_\nu^{-2}Z'P_{\tilde{H}}y + \sigma_1^{-2}Z'P_{\bar{H}}y$. Therefore, $\hat{\delta}_{SEC-2SLS}$ given by Equation (9) can be obtained as 2SLS on (6) using $B = \left(\tilde{H}, \bar{H}\right)$ as instruments.

Lee (2003) argued that in the cross-section spatial model, the optimal instruments for estimating δ in Equation (2) is

$$E(Z) = E[X, Wy] = \left(X, WA^{-1}X\beta\right).$$

Therefore, a Lee (2003) type optimal instruments for estimating δ in the $\Omega^{-1/2}$ transformed panel autoregressive spatial model in Equation (6) is

$$E(Z^*) = E\left(\Omega^{-1/2}Z\right) = E\left[\Omega^{-1/2}(X, Wy)\right] = \left(\Omega^{-1/2}X, \Omega^{-1/2}WA^{-1}X\beta\right),$$

and the resulting best spatial 2SLS estimator is given by

$$\hat{\delta}_{BS-2SLS} = \left(H_b^{*\prime}Z^*\right)^{-1}H_b^{*\prime}y^*, \tag{11}$$

where $H_b^* = \left(\Omega^{-1/2}X, \Omega^{-1/2}WA^{-1}X\beta\right)$. A feasible version of this estimator is based on consistent estimators of σ_1^2, σ_ν^2, λ and β respectively.

Similarly, one can extend Baltagi's (1981) error component two-stage least squares estimator to this spatial autoregressive model using H_b^* as instruments. Left multiply Equation (6) by $\left(\tilde{H}_b^*, \bar{H}_b^*\right)'$, where $\tilde{H}_b^* = QH_b^* = \sigma_\nu^{-1}\left(QX, QWA^{-1}X\beta\right)$ and $\bar{H}^* = PH_b^* = \sigma_1^{-1}\left(PX, PWA^{-1}X\beta\right)$, we get the two equation system recognizing that they estimate the same δ:

$$\begin{pmatrix}\tilde{H}_b^{*\prime}y^* \\ \bar{H}_b^{*\prime}y^*\end{pmatrix} = \begin{pmatrix}\tilde{H}_b^{*\prime}Z^* \\ \bar{H}_b^{*\prime}Z^*\end{pmatrix}\delta + \begin{pmatrix}\tilde{H}_b^{*\prime}u^* \\ \bar{H}_b^{*\prime}u^*\end{pmatrix},$$

where $E\begin{pmatrix}\tilde{H}_b^{*\prime}u^* \\ \bar{H}_b^{*\prime}u^*\end{pmatrix} = 0$ and $Var\begin{pmatrix}\tilde{H}_b^{*\prime}u^* \\ \bar{H}_b^{*\prime}u^*\end{pmatrix} = \begin{bmatrix}\tilde{H}_b^{*\prime}\tilde{H}_b^* & 0 \\ 0 & \bar{H}_b^{*\prime}\bar{H}_b^*\end{bmatrix}$. Performing GLS on this system we obtain the spatial error component best two-stage least squares estimator (SEC-B2SLS):

$$\hat{\delta}_{SEC-B2SLS} = \left(Z^{*\prime}P_{\tilde{H}_b^*}Z^* + Z^{*\prime}P_{\bar{H}_b^*}Z^*\right)^{-1}\left(Z^{*\prime}P_{\tilde{H}_b^*}y^* + Z^{*\prime}P_{\bar{H}_b^*}y^*\right) \tag{12}$$

$$= \left(\sigma_\nu^{-2}Z'P_{\tilde{H}_b^*}Z + \sigma_1^{-2}Z'P_{\bar{H}_b^*}Z\right)^{-1}\left(\sigma_\nu^{-2}Z'P_{\tilde{H}_b^*}y + \sigma_1^{-2}Z'P_{\bar{H}_b^*}y\right). \tag{13}$$

The second equality uses the fact that $Q\Omega^{-1/2} = \sigma_\nu^{-1}Q$ and $P\Omega^{-1/2} = \sigma_1^{-1}P$. This SEC-B2SLS estimator can also be obtained from the $\Omega^{-1/2}$ transformed equation in (6) using $B = \left(\tilde{H}_b^*, \bar{H}_b^*\right)$ as instruments. In fact, $\tilde{H}_b^*$ and $\bar{H}_b^*$ are orthogonal to each other, since $QP = 0$. Hence, $P_B = P_{\tilde{H}_b^*} + P_{\bar{H}_b^*}$. This also implies that $Z^{*\prime}P_BZ^* = Z^{*\prime}P_{\tilde{H}_b^*}Z^* + Z^{*\prime}P_{\bar{H}_b^*}Z^*$ and $Z^{*\prime}P_By^* = Z^{*\prime}P_{\tilde{H}_b^*}y^* + Z^{*\prime}P_{\bar{H}_b^*}y^*$. Therefore, $\hat{\delta}_{SEC-B2SLS}$ given by Equation (12) is the same as 2SLS on (6) using $B = \left(\tilde{H}_b^*, \bar{H}_b^*\right)$ as instruments.

Note that these 2SLS estimators are easy to apply using standard software. In fact, one can easily extend the *ec2sls* procedure in Stata for panels to the spatial panel case. Note also that SEC-2SLS and SEC-B2SLS use twice the instruments that their counterparts RE-S2SLS and BS-2SLS use. Although the set of instruments in the former is completely spanned by those for the latter estimators, these may yield smaller empirical standard errors in small samples, see Baltagi and Liu (2009) for a similar argument in the panel data case with no spatial correlation.

REFERENCES

Anselin, L. (1988), Spatial Econometrics: Methods and Models. Kluwer Academic Publishers, Dordrecht.

Anselin, L., J. Le Gallo and H. Jayet (2008), Spatial Panel Econometrics. Ch. 19 in L. Mátyás and P. Sevestre, eds., The Econometrics of Panel Data: Fundamentals and Recent Developments in Theory and Practice, Springer-Verlag, Berlin, 625-660.

Baltagi, B.H. (1981), Simultaneous Equations with Error Components, Journal of Econometrics, 17, 189–200.

Baltagi, B.H. (2008), Econometric Analysis of Panel Data, New York, Wiley.

Baltagi, B.H. and L. Liu (2009), A Note on the Application of EC2SLS and EC3SLS Estimators in Panel Data Models, Statistics and Probability Letters, 79, 2189-2192.

Cornwell, C., P. Schmidt and D. Wyhowski (1992), Simultaneous Equations and Panel Data, Journal of Econometrics, 51, 151–181.

Elhorst, P.J. (2003), Specification and Estimation of Spatial Panel Data Models, International Regional Science Review, 26(3), 244–268.

Kelejian, H.H. and I.R.Prucha (1998), A Generalized Spatial Two-Stage Least Squares Procedure for Estimating a Spatial Autoregressive Model with Autoregressive Disturbances, Journal of Real Estate Finance and Economics, 17(1), 99-121.

Kelejian, H.H., I.R.Prucha and Y. Yuzefovich (2004), Instrumental Variable Estimation of a Spatial Autoregressive Model with Autoregressive Disturbances: Large and Small Sample Results, Advances in Econometrics, 18, 163-198.

Lee, L. (2003), Best Spatial Two-Stage Least Square Estimators for a Spatial Autoregressive Model with Autoregressive Disturbances, Econometric Reviews, 22(4), 307-335.