Estimating spatial heterogeneous panel data: an Information-Theoretic approach

ROSA BERNARDINI PAPALIA
Department of Statistics,
University of Bologna
Via Belle Arti, 41 - 40100 Bologna
ITALY
rossella.bernardini@unibo.it

Abstract: - This paper introduces an entropy-based estimation strategy for spatial heterogeneous panel data models where separate processes for each unit are considered. The starting point is a general model specification which account for both temporal and spatial lagged effects in a panel data context by treating individual relationships as a system of seemingly unrelated regression equations. An empirical application is provided to demonstrate practical implementation of the GME estimator when one has to deal with estimation of ill-posed or ill-conditioned models in analyzing spatial structures.

Keywords: - Heterogeneous panel data, Generalized maximum entropy estimation, Spatial models

1 Introduction

In recent years, there has been a growing interest in the estimation of econometric relationships based on panel data (Hsiao 1989). In this paper we focus on spatial panels, a family of models for which there is an increasing interest in estimation problems (Elhorst 2001 2003).

A spatial panel data model takes the form of a linear equation extended with a variable intercept, a spatially lagged dependent variable (known as spatial lag model) or a spatially autoregressive process incorporated in the error term (known as spatial error model). With respect to a standard space-time model (Space-Time Autoregressive /Integrated Moving Average model, STARMA, STARIMA (Hepple 1978), and the spatial autoregression space-time forecasting model (Griffith 1996)) which assumes that the spatial units are completely homogeneous, a panel data approach would presume that spatial heterogeneity is a feature of data and attempt to model that heterogeneity.

The need to account for spatial heterogeneity is that spatial units are likely to differ in their background variables, which are usually space-specific time-invariant variables that affect the dependent variable, but are difficult to measure or hard to obtain. Omission of these variables leads to bias in the resulting estimates. To overcome these problems, one possibility is to introduce a variable intercept representing the effect of the omitted variables that are peculiar to each spatial unit considered. More specifically, in the fixed effects model, a dummy variable is introduced for each spatial unit as a measure of the variable intercept, while, in the random effects model, the variable intercept is treated as a random variable that is independently and identically distributed with zero mean. T dynamics are heterogeneous across the cross section units.

This work aims at developing an entropy-based estimation strategy of heterogeneous spatial panel data models. The starting point is a general model specification which account for both temporal and spatial lagged effects in a panel data context.

There is now a considerable body of work, which has given an application of the entropy criterion to a wide class of models (Marsh et al. 1998, 2003; Golan et al. 1996; Kullback 1959; Samilov 2006; Fragoso et al. 2008; Wu et al. 2008). As regards traditional estimation techniques, the formulation of the constrained maximization problem in the maximum entropy view does not require the use of restrictive parametric assumptions on the model. Restrictions expressed in terms of inequality can be introduced and it is possible to calibrate the precision in the estimation. Good results are produced in the case of small-sized samples, in the presence of high numbers of explanatory parameters and variables (highly correlated). An empirical application is provided to demonstrate practical implementation of the GME estimator when one has to deal with estimation of ill-posed or ill-conditioned panel data models with spatial structures.

The paper is organized as follows. Section 2 introduces the basic model introducing the typical assumptions of an information theoretic framework. More specifically, section 2.2. develops the maximum entropy based formulation for alternative spatial error specifications. In section 3 we present an application of GME methodology for estimating spatial models. Finally, section 4 concludes and lists some potential advantages and investigations of the proposed approach.

2 The information-theoretic framework

2.1 The basic problem

Following the heterogeneous panel approach proposed by Pesaran and Smith (1995), we take into account the possibility of cross-sectional correlation by treating individual relationships as a system of seemingly unrelated regression equations.

In the context of this work, the general specification of the model for cross section unit i (i=1,..,N) and t=1,..,T time periods which involves specifying a different intercept coefficient for each cross-sectional unit (equation) is:

$$Y_{it} = \rho_i w_1 Y_{it} + X_{it}\beta_i + \varepsilon_{it} \qquad (1)$$

We start by considering the basic model in vector form for a cross-section unit i:

$$Y_i = \rho_i w_1 Y_i + X_i \beta_i + \varepsilon_i$$
$$\varepsilon_i = \left(I - \lambda_i w_2\right)^{-1} u_i \qquad (2)$$

where Y_i is of dimension (T×1), X_i is (T×K_i) and β_i is (K_i×1).

The resulting complete model, expressed as a system of seemingly unrelated system of equations, can be written as:

$$Y = \rho W_1 Y + X\beta + \left(I - \lambda W_2\right)^{-1} U$$

$$\text{where } Y \equiv \begin{bmatrix} y_1 \\ \dots \\ \dots \\ y_N \end{bmatrix} \quad X \equiv \begin{bmatrix} X_1 & 0 & .. & 0 \\ 0 & X_2 & .. & 0 \\ .. & & .. & .. \\ 0 & 0 & .. & X_N \end{bmatrix} \qquad (3)$$

$$\beta \equiv \begin{bmatrix} \beta_1 \\ \dots \\ \dots \\ \beta_N \end{bmatrix} \quad \rho \equiv \begin{bmatrix} \rho_1 \\ \dots \\ \dots \\ \rho_N \end{bmatrix} \quad U \equiv \begin{bmatrix} u_1 \\ \dots \\ \dots \\ u_N \end{bmatrix} \quad \lambda \equiv \begin{bmatrix} \lambda_1 \\ \dots \\ \dots \\ \lambda_N \end{bmatrix}$$

where *Y* is a vector of dimension (TN×1), X is a block diagonal matrix of dimension (TN×K) with $K=\Sigma K_i$, and $\beta=(\beta_1, \beta_2,.. \beta_N)'$, $\lambda=(\lambda_1, \lambda_2,.., \lambda_N)'$ $\rho=(\rho_1, \rho_2,.., \rho_N)'$ are unknown vectors of dimension (KN×1) (N×1) and (N×1), respectively. This model specification assumes that spatial effects are not identical across spatial units; by varying spatial residual correlation coefficients across units it takes into account jointly structural instability and differentiated spatial effects within and between spatial units.

W_1 is a (TN×TN) block diagonal matrix which expresses for each observation (row) those units (columns) that belong to its neighbourhood set as non-zero elements, that is: for pairs of units (i,j), $w_{ij} \neq 0$ for 'neighbours' and $w_{ij} = 0$ for others. It is common practice to derive spatial weights from the location and spatial arrangements of observation by means of a geographic information system. In this case, units are defined 'neighbours' when they are within a given distance of each other, ie $w_{ij}=1$ for $d_{ij} \leq \delta$ and $i \neq j$, where d_{ij} is the great circle distance chosen, and δ is the critical cut-off value. More specifically, a spatial weights matrix W* is defined as follow:

$$w_{ij}^* = \begin{cases} 0 & \text{if} \quad i = j \\ 1 & \text{if} \quad d_{ij} \leq \delta, i \neq j \\ 0 & \text{if} \quad d_{ij} > \delta, i \neq j \end{cases} . \qquad (4)$$

and the elements of the row-standardized spatial weights matrix W (with elements of a row sum to one) result:

$$w_{ij} = \frac{w_{ij}^*}{\sum_{j=1}^{N} w_{ij}^*}, \qquad i, j = 1,.., N. \qquad (5)$$

W_2 is the (TN×TN) block diagonal matrix with T copies of the (N×N) spatial weight matrix, and U is a (TN×1) vector of iid errors with variance σ_u^2.

The specification in the form of a nonlinear in parameters form of the simultaneous equations model result:

$$Y=\left(\rho W_1+\lambda W_2-(\lambda W_2)(\rho W_1)\right)Y+\left(I-\lambda W_2\right)X\beta+u \quad (6)$$

Following Theil (1971) and Zellner (1998), from the joint spatial model (6) and from substitution of the reduced form equation:

$$Y = S\pi + \tau \quad (7)$$

for Y, the resulting joint spatial model is derived:

$$Y=\left(\rho w_1+\lambda w_2-(\lambda w_2)(\rho w_1)\right)\left(S\pi\right)+\left(I-\lambda w_2\right)X\beta+u^* \quad (8)$$

where u* is a (TN×1) vector of appropriately transformed residuals, S is a (TN×L) matrix of instrumental variables, π is a (NL×1) vector of unknown parameters, and τ is a (TN×1) vector of reduced form residuals. Including the reduced form model is necessary to identify the reduced form parameters (Marsh, 2003).

The basic generalized maximum entropy (GME) formulation we present refer to a model specification where a constant spatial structure for error variance in each equation is assumed so that the disturbance vector is assumed to have a zero mean and a non-diagonal covariance matrix:

$$\Phi = \Sigma \otimes I_T \quad (9)$$

where $\otimes$ is the Kronecker product operator and I_T is an identity matrix of dimension T.

Under the GME framework we recover simultaneously the unknown parameters, the unknown errors by defining an inverse problem, which is based only on indirect, partial or incomplete information. The objective is to estimate the parameters of spatial models with minimal distributional assumptions. Each parameter is treated as a discrete random variable with a compact support Z and M possible outcomes, 2 ≤M ≤∞. The uncertainty about the outcome of the error process is represented by treating each error as a finite and discrete random variable with J possible outcomes, 2≤J≤∞. To this end, we start by choosing a set of discrete points, the support space $V=[v_1,v_2,\dots,v_J]'$ of dimension $J\geq 2$, that are at uniform intervals and symmetric around zero. Each error term has corresponding unknown weights $r_i=[r_{i1},r_{i2}\dots,r_{iJ}]'$ that have the properties of probabilities $0\leq r_{ij}\leq 1$ and $\sum_j r_{ij}=1$.

Re-parameterizing the set of equations (7) and (8), so that $\beta=Zp^\beta$ and $\pi=Zp^\pi$, $\rho=Zp^\rho$, $\lambda=Zp^\lambda$, $u^*=Vr^{u^*}$ and $\tau=Vr^\tau$ yields:

$$Y=\left(\left(Z^\rho p^\rho\right)w_1+\left(Z^\lambda p^\lambda\right)w_2-\left(Z^\lambda p^\lambda\right)w_2\left(Z^\rho p^\rho\right)w_1\right)\left(S\left(Z^\pi p^\pi\right)\right) + \left(I-\left(Z^\lambda p^\lambda\right)w_2\right)X\left(Z^\beta p^\beta\right)+\left(Z^{u^*}p^{u^*}\right) \quad (10)$$

$$Y = S\left(Z^\pi p^\pi\right)+\left(Z^\tau p^\tau\right) \quad (11)$$

where $p=vec(p^\beta p^\pi p^\rho p^\lambda)$ and $r=vec(p^{u^*}p^\tau)$ are vectors of proper probability distributions for parameters and errors, respectively.

2.2 Estimation issues

The estimation of panel data models incorporating both spatial heterogeneity and spatial dependence poses identification, endogeneity and collinearity problems and as a consequence the standard estimation procedures can produce (i) biased parameter estimates, (ii) unbiased but inefficient parameter estimates, or (iii) biased estimates of the standard errors.

Another potential problem is that for large N, the usual spatial econometric procedures are problematic because the eigenvalues of spatial weight matrices of dimensions over 400 cannot be estimated with sufficient reliability (Kelejian and Prucha, 1999). One solution is to use the GMM estimator in the case of the fixed effects spatial error model (Bell and Bockstael, 2000). Another solution, based on maximum likelihood estimation, is not to express the Jacobian term in the individual eigenvalues but in the coefficients of a characteristic polynomial (Smirnov and Anselin, 2001) or to approximate the Jacobian term in its original form using a Monte Carlo approach (Barry and Pace, 1999). More specifically, the spatial LAG-SUR specification may be consistently (but not efficiently) estimated by feasible generalized least squares, F-GLS (Zellner 1997), while the SAR-SUR specification cannot be consistently estimated by F-GLS due to the endogeneity of spatial spillovers (Anselin, 1988). Consistent estimates may be obtained for both of these specifications using: (i) a combination of F-GLS and Maximum Likelihood estimation (Anselin 1988), (ii) two-stage estimation procedures (Zellner 1997), or (iii) moment conditions for the GMM estimation derived by Honoré and Hu (2004).

A quite important problem of fixed coefficients models, also expressed in spatial forms, is the large number of parameters causing the estimators to be infeasible. Furthermore, even if the estimators are made feasible by introducing restrictions on the parameters, the quality of the asymptotic approximation used to justify the approach remains rather suspect, unless the ratio N/T tends to zero. In this respect, as a suitable alternative, in presence of both endogeneity and ill-posed problems, consistently and asymptotically normal estimates may be obtained by using the Generalized Maximum Entropy estimation approach (Golan et al. 1996) which avoids some of the strong parametric assumptions required with traditional procedures and performs well over a range of non-Normal error distributions and in presence of small samples.

The GME estimator can be viewed as a shrinkage estimator that shrinks the data to the priors (uniform distributions) and toward the centre of their supports. It should be pointed out that unlike ML estimators, the GME approach does not require any explicit error-distribution assumptions: in fact, the GME method selects the most uniform distribution consistent with the information provided by the constraints. In this respect, we do not need to specify a parametric family for the likelihood function, and the estimation rule is flexible with respect to: (i) the dynamic, stochastic nature of economic data, (ii) a non-random survey design, as well as to the model selection problems. Within the GME framework, all coefficients and errors are expressed in terms of proper probabilities. The basic idea is that rather than search for the point parameter estimates, each parameter is viewed as the mean value of some well-defined random variable. The unobserved error vector is also viewed as another set of unknowns, and as in the case of the signal vector, each error is constructed as the mean value of a random variable. Under the GME framework, the full distribution of each parameter and of each error (within their support spaces) is simultaneously estimated under minimal distributional assumptions.

Given the data consistency (10) and (11), and the covariance's relationship (9) the GME objective function relative to our formulation problem may be formulated as:

$$\max_{p_i, r_i} H(p,r) = -p'\ln p - r'\ln r \tag{12}$$

subject to:

(i) data consistency conditions:

$$\begin{aligned} Y = & \left(\left(Z^{\rho}p^{\rho}\right)w_1 + \left(Z^{\lambda}p^{\lambda}\right)w_2 - \left(Z^{\lambda}p^{\lambda}\right)w_2\left(Z^{\rho}p^{\rho}\right)w_1\right)\left(S\left(Z^{\pi}p^{\pi}\right)\right) \\ & + \left(I - \left(Z^{\lambda}p^{\lambda}\right)w_2\right)X\left(Z^{\beta}p^{\beta}\right) + \left(Z^{u^*}p^{u^*}\right) \end{aligned} \tag{13}$$

$$Y = S\left(Z^{\pi}p^{\pi}\right) + \left(Z^{\tau}p^{\tau}\right) \tag{14}$$

(ii) adding-up constraints:

$$\begin{aligned} & 1'p_k^{\beta} = 1\,\forall k; \qquad 1'p_k^{\pi} = 1 \quad \forall k; \\ & 1'p^{\rho} = 1; \qquad 1'p^{\lambda} = 1; \\ & 1'p_i^{u^*} = 1\,\forall i; \qquad 1'p_i^{\tau} = 1 \quad \forall i; \end{aligned} \tag{15}$$

where $p = vec(p^{\beta}p^{\pi}p^{\rho}p^{\lambda})$ and $r = vec(p^{u^*}p^{\tau})$ are vectors of proper probability distributions for parameters and errors, respectively.

If the correlation between time periods varies with time, the error covariance that results is more general than the block diagonal structure we have considered. In such a cases it is possible to introduce different covariances for each period by adding a new set of constraints in the optimization problem.

It is important to point out some advantages of our GME specification for spatial panel datamodels. First, by assuming heterogeneity of parameters across units it is possible to analyze the spatial patterns in the estimated coefficients. Second, by using the GME procedure it is possible to obtain consistent parameter estimates when the number of these coefficients increases as the number of observations increases (the curse of dimensionality problem), when the number of time periods involved, T, is less than the number of observations N and the corresponding variance-covariance matrix for errors is singular.

3 An empirical application

In the empirical application, firms' competitive effects from FDI inflows have been studied with data on foreign direct investment inflows in Italy, collected by the Italian Foreign Exchange Office for all economic sectors (1999-2004). The application examines the importance of local agglomeration externalities in determining the Foreign Direct Investment (FDI) intensity by analyzing the link between the degree of FDI inflows penetration and its determinants at the Italian regional level. In contrast with most previous studies (Konings, 2001), we focus on the spillover effects related to inter/intra-industry linkages on FDI intensity. More specifically, we are interested in testing: (i) the hypothesis that FDI intensity is related to the characteristics of the regional economic system; (ii) the hypothesis that sector specificities are relevant in explaining FDI patterns in Italy. In this respect, we argue the importance of a panel data approach and the need of testing for spatial heterogeneity and for any remaining spatial autocorrelation, since ignoring it could result in biased coefficients. Our analysis is concentrated in the 1999-2004 period and Balance of payments data on Italy's inward FDI flows are used. Data on FDI are collected by region, and aggregated into 4 macro-sectors, agriculture, industry in the strict sense, construction, and services. Regional data used for the construction of the explanatory variables come from Istat. The dependent variable, FDI-intensity, is calculated by using FDI inflow divided by value added, for each region and sector. Our specification includes as regressors time and sectoral fixed effects and the model has been estimated in log-linear form. Spatial heterogeneity is modelled by introducing fixed effects at sectoral level and heterogeneous coefficients across regions. To account for residual spatial dependence, spatial lag and spatial error specifications have also been considered.

Spillovers connected to specialization (Marshall-type economies) and to diversity are introduced. With reference to *Marshall-type agglomeration spillovers*, capturing the positive effects of the agglomeration of firms belonging to the same sector, we consider a sector specialization index computed on industry employment:

$$Spec_{ij} = \frac{IS_{ij} - 1}{IS_{ij} + 1} \text{ with } IS_{ij} = \frac{L_{ij} \Big/ \sum_j L_{ij}}{L_{ITA\,j} \Big/ \sum_j L_{ITA\,j}}$$

and $i = 1, \ldots, 20$ j = *agriculture, industry in the strict sense, construction, services*

where L_{ij} is employment in region i and industry j, and $L_{ITA\,j}$ is employment at national level in industry j. The index is standardized and constrained within the interval (–1, 1).

To measure the *Jacobs-type externalities* we employ the relative Hirschman-Herfindal index:

$$Div_{ij} = \frac{H_{ij}}{H_{ITA\,j}}$$

with $i = 1, \ldots, 20$ j = *agriculture, industry in the strict sense, construction, services,* where $H_{ij} = \sum_{j^* \neq j} s^2_{ij^*}$, and $s^2_{ij^*} = L_{ij} \Big/ \sum_{j^* \neq j} L_{ij^*}$.

For region i and sector j, the index is measured over all the industrial sectors except j and is decreasing with the relative diversity of the area compared with the national average, that is higher indexes indicate less diversified areas. A positive effect of vertical externalities is detected by a negative sign for the corresponding coefficient.

In addition, to account for spillovers among firms localized in the same geographical area (Porter-type spillovers), a district variable has been introduced. Number of workers employed in districts, identified by ISTAT, in a region divided by the total number of workers in the same region, as a measure of the local development of firms' networks across Italian regions. Finally, as a measure of trade openness we calculated a trade variable by using the sum of imports and exports divided by GDP, for each region and sector.

Due to the panel data size, standard techniques cannot be applied when heterogeneity in coefficients of explanatory variables across both regions and sectors is introduced. Estimates are then computed using the GME-based estimator previously introduced. We use M=5 and J=3 since with higher values produced the same estimates. Results of GME estimates of the spatial lag model with varying (region specific) spatial dependence (Table 2-Appendix) confirm the high heterogeneity of all types of spillovers across regions, mostly concentrated in Centre and Northern regions. Endogenous spillovers connected to geographical proximity produce diversified effects on regions: FDI intensity in a region is positively affected by FDI flows in neighboring regions for Piemonte, Lombardia, Veneto, and Toscana. The effect is negative for Liguria, Umbria, Marche, Lazio, and Sardegna. Specialization externalities have mixed effects on FDI intensity. With reference to vertical linkages across industries, no clear evidence comes out, except for agriculture and construction where inter-industry externalities positively influence MNE's investment decisions. Finally, there is evidence that the importance of spatial dependence varies across regions on data disaggregated in four macro-sectors. Differences in the estimated coefficients on the region-specific spatial dependence terms emerge for regions that are contiguous to the North and Center of Italy while peripheral regions experience much smaller and negligible effects.

4 Remarks and conclusions

In this paper a GME estimation procedure for heterogeneous spatial panel data model has been derived. Spatial autoregressive processes in either the dependent variable and residuals are here considered by taking into account jointly structural instability and differentiated spatial effects between and within spatial cross sectional units. This model specification allows for much more flexibility of the coefficients across units than do traditional models and also contributes: (i) to avoid the potential bias that arises from constraining the coefficients on the lagged dependent variable to be constant across units, and (ii) to provide a diagnostic tool to investigate the presence of some type of heterogeneity in panel data sets (unobserved heterogeneity, slope heterogeneity, spatial dependence).

An empirical application has been presented with the aim of studying firms' competitive effects from FDI inflows using data on foreign direct investment inflows in Italy. A fixed effects panel data model, also extended to include spatial autocorrelation, has been examined by testing the hypothesis of panel heterogeneity in slope coefficients with varying (region specific) spatial dependence. The empirical findings give support to the hypothesis that different types of agglomeration externalities, such as Marshall-type economies related to the sector specialization of a specific geographical area, and Jacobs-type externalities linked to sector diversity, contribute to affect FDI inflows. We find that controlling for fixed-effects allows us to disentangle the effect of spatial dependence from that of spatial heterogeneity and of omitted variables. The estimated relationship of determinants of FDI intensity is robust to inclusion of terms introduced to capture spatial interdependence, even though such interdependence is estimated to be significant. Though heterogeneity in coefficients across both regions and sectors have been identified, the presence of foreign multinational firms seems to play an important role in terms of linkages with the local context. Inter-industry and intra-industry externalities have shown mixed effects across sectors. Finally, endogenous spillovers connected to geographical proximity have produced diversified effects on regions.

With reference to the use of the GME-based estimation procedures to estimate spatial structures in panel data models, several advantages can be pointed out. The maximum entropy-based estimator is able to produce consistent estimates in "ill-posed" and/or "ill-conditioned" models where the number of parameters exceeds the number of data points and in models characterized by a non-scalar identity covariance matrix. Prior information can be introduced by adding suitable constraints in the formulation without imposing strong distributional assumptions.

Further investigation of GME estimators for spatial panel data models could yield useful results (i) to applied analyses of socio and economic phenomena, and (ii) to space-time problems under complex and nonrandom sample designs.

References:

[1] Anselin L. (1988). *Spatial Econometrics: Methods and models*. Dordrecht, the Netherlands, Kluwer.

[2] Anselin L. (2001). Spatial Econometrics. In *A companion to theoretical econometrics*, ed. B.H. Baltagi, 310-330. Malden, Blackwell.

[3] Barry R.P. and Pace R.K. (1999) Monte Carlo estimates of the log determinant of large sparse matrices, *Linear Algebra and Its Applications* , 289, pp. 41-54.

[4] Bell K.P. and Bockstael N.E. (2000) Applying the generalized-moments estimation approach to spatial problems involving microlevel data, *Review of Economics and Statistics*, 82, pp. 72-82.

[5] Elhorst JP (2001) Dynamic Models in Space and Time. Geographical Analysis 33: 119-140

[6] Fragoso R., De Belem Martins M., Ventura Lucas MR. (2008) Generate disaggregated soil allocation data using a Minimum Cross Entropy Model, WSEAS *Transactions on Envieronment and Development*, 9, 756-566.

[7] Golan A, Judge G, Miller D (1996). Maximum entropy econometrics: robust estimation with limited data. Wiley, New York.

[8] Elhorst JP (2003) Specification and Estimation of Spatial Panel Data Mod*els, International Regional Science Review*, 26:3, 244-268.

[9] Griffith DA (1996) Computational simplifications for Space-Time Forecasting within GIS: The Neighbourhood Spatial Forecasting Models. In Spatial Analysis: Modelling in a GIS Environment, eds, P. Longley and M. Batty. Cambridge, UK: Geo Information International

[10] Hepple LW (1978) The Econometric Specification and Estimation of Spatio-Temporal Models. In Time and Regional Dynamics, eds., T. Carlstein, D. Parkes, and N. Thrift. London, Edward Arnold

[11] Honoré B.E. and Hu L. (2004) Estimation of cross sectional and panel data censored regression models with endogeneity, *Journal of Econometrics*, 122, pp. 293-316.

[12]Hsiao C (1989) Modeling Ontario Regional Electricity System Demand Using a Mixed Fixed and Random Coefficient Approach, Regional Science and Urban Economics, 19

[13] Kelejian H.H. and Prucha I.R. (1999) A generalized moments estimator for the autoregressive parameter in a spatial model, *International Economic Review*, 40, pp. 509-533.

[14] Konings J. (2001) The Effects of Foreign Direct Investment on Domestic Firms: Evidence from Firm Level Panel Data in Emerging Economies, *Economics of Transition*, Vol. 9, pp. 619-633.

[15] Kullback (1959) Information theory and statistics. Wiley, New York.

[16] Lesage J.P., Pace R.K. (2004) *Spatial and Spatiotemporal Econometrics* - Advances in Econometrics, Vol. 18. Elsevier Ltd.

[17]Marsh L Mittlehammer RC (2003) Generalized maximum entropy estimation of spatial autoregressive models. Working paper, Kansan State University

[18] Marsh L, Mittlehammer RC, Cardell S (1998) A Structural-Equation GME Estimator. Selected Paper for the AAEA Annual Meeting, 1998, Salt Lake City

[19] Mundlak Y (1978) On the Pooling of Time Series and Cross-Section Data, Econometrica, 36, 69-85

[20] Pesaran MH, Smith R (1995) Estimating long-run relationships from dynamic heterogeneous panels, Journal of Econometrics, 68, p. 79-113

[21] Samilov A. (2006) A development of Entropy Optimization Methods, WSEAS *Transactions on Mathematics*, 5, 568-575.

[22] Smirnov O. and Anselin L. (2001) Fast maximum likelihood estimation of very large spatial autoregressive models: A characteristic polynomial approach, *Computational Statistics & Data Analysis,* 35, pp. 301-319.

[23] Theil H (1971) Principles of Econometrics. John Wiley &Sons: New York.

[24] Wu Y, Fyfe C. (2008) The On-line Cross Entropy Method for Unsupervised Data Exploration, WSEAS *Transactions on Mathematic*.

[25] Zellner A (1996) Models, prior information, and Bayesian Analysis, Journal of Econometrics, 75, 51-68

[26] Zellner A (1997) A Bayesian Method of Moments (BMOM): Theory and Applications. Advances in Econometrics, Vol. 12. Applying Maximum Entropy to Econometric Problems, T.B. Formby and R.C. Hill, (Eds.) 85-105. JAI Press, Grenwich

[27] Zellner A (1998) The finite sample properties of simultaneous equations'estimates and estimators Bayesian and Non-Bayesian Approaches, Journal of Econometrics, 83, 185-212.

Appendix: Description of data

Data related to Italy's inward FDI flows by (destination) region and economic sector are collected by Ufficio Italiano Cambi (UIC) and come from the balance of payments statistics. FDI is defined as "the category of international investment that reflects the objective of a residence entity in one economy obtaining a lasting interest in an enterprise resident in another economy" (IMF, 1993). Following the IMF guidelines, the UIC provides the following definition: the investment in a foreign company is classified as FDI when it involves 10 per cent or more of the company's share; it is classified as a portfolio investment when it is less than 10 per cent. Using balance of payments data on Italy's inward FDI flows (from 1999 to 2004), a wide array of activities related to the internationalization of production is covered, including greenfield investments abroad as well as cross-border M&As. "Non-equity" forms of internationalization are instead not covered. Other data on value added, employment, and trade flows come from Istat.

Tab. 1: Description of sectors

Macro-sector	Sector	Description
A	A1	**AGRICULTURE**
		Agriculture, hunting, forestry and fishing
B		**INDUSTRY**
	B1	**INDUSTRY IN THE STRICT SENSE**
		Paper, paper products and printing
		Agricultural and industrial machinery and equipment
		Office, accounting and computing machinery
		Electrical materials
		Transport equipment
		Ferrous and non-ferrous minerals and metals
		Non-metallic mineral products
		Food products, beverages, and tobacco products
		Chemicals products
		Energy products
		Rubber and plastics products
		Metal products except transport equipment
		Textiles, leather, footwear and clothing
		Other manufacturing products
	B2	**CONSTRUCTION**
		Construction in private and public sectors
C		**SERVICES**
	C1	**MARKET SERVICES**
		Supporting and auxiliary transport activities
		Hotels and restaurants
		Land transport
		Water and air transport
		Maintenance and repair of motor vehicles, personal and household
		Telecommunications
		Other trade services
	C2	**NON-MARKET SERVICES**
		Public administration
		Private households with employed persons
	C3	**FINANCIAL INTERMEDIATION**
		Financial intermediation, except insurance and pension funding
		Insurance and pension funding

Tab. 2: Model with heterogeneity across agriculture, industry in the strict sense, construction and services, Dependent variable: y = ln (1+FDI/VA) – Generalized Maximum Entropy estimates

Regions	**constant**	***y* spatial lag**	**open**	**district**	**ln(spec) in sector 1**	**ln(spec) in sector 2**	**ln(spec) in sector 3**	**ln(spec) in sector 4**	**ln(div) in sector 1**	**ln(div) in sector 2**	**ln(div) in sector 3**	**ln(div) in sector 4**
Piemonte	0.00	2.90*	-0.25*	0.30*	-5.68*	-8.97	-2.62***	-45.46*	-39.99*	24.03	-50.38*	22.76*
Valle d'Aosta	0.62	-0.01	-0.001	0.00	0.28*	-6.53*	3.92*	1.84	-3.47*	5.85*	2.49*	-2.09
Lombardia	-13.10*	3.57*	-0.69*	-126E-17*	-20.32*	-67.34*	47.05*	-78.30*	-127.72*	-90.70*	-77.56*	77.35*
Trentino-A. A.	0.48*	-0.01	-0.05*	7.61E-16**	1.85**	1.41	1.80*	-1.84	2.03**	-1.59	2.44	1.68
Veneto	9.04*	0.20**	0.59***	-0.001*	4.39*	10.60*	-0.56**	-10.72*	3.83*	-5.33*	-4.35***	6.12*
Friuli-. G.	-0.09**	-0.10	0.00	-0.01*	0.01	-0.14	0.03	-0.05	-1.00	-0.07	-1.14	0.13
Liguria	-3.10*	-0.29*	0.01	1.455E-8*	-0.18	-0.45	-0.03	-1.74	-5.67	10.83*	-7.13*	6.82*
Emilia-Romagna	-1.64*	0.09	0.10*	-0.01*	-1.52*	5.75*	-1.62	4.91	15.86*	-5.97*	-14.42*	-9.70**
Toscana	3.78**	1.05*	-0.30*	0.00	1.48*	24.06*	2.58*	0.87*	-1.01**	14.92*	-1.58*	-0.24*
Umbria	-0.54*	-7.55*	0.02*	0.00	2.89*	-23.30*	-1.81***	6.01*	4.77	4.80***	4.62	-2.28**
Marche	0.11	-0.12*	0.01	0.00	-0.005*	-3.81*	0.00	0.82**	-1.01*	0.20	0.16	-0.22**
Lazio	0.00	-0.72*	0.20	-5.40*	5.35*	5.50**	-4.66*	-21.24*	-17.47**	0.17	12.30*	18.23*
Abruzzo	0.00	0.02	-0.02*	0.01*	0.03	0.08	0.07	-0.56	-1.22*	-0.89	-0.29	0.04
Molise	1.32E-16*	0.04	-0.004**	0.00	0.08	-0.12	0.27**	-0.07	-0.19	-0.01	-0.35***	0.02
Campania	0.00	0.17	0.03	-0.07	-0.27	-1.33	0.00	-0.18	-1.61*	1.41	0.94	0.22
Puglia	0.00	9.93	0.00	0.00	0.13	-0.22	0.00	0.05	-0.05	-0.02	0.02	-0.01
Basilicata	0.00	0.05	0.00	0.01	-0.01	0.00	0.00	0.00	0.03	0.00	-0.01	0.00
Calabria	-0.01	0.09	0.00	0.00	-0.13	-0.04	0.01	0.06	0.17	0.24	0.08	0.08
Sicilia	0.00	3.23	0.00	-0.02	-0.01	0.00	0.00	-0.01	-0.01	0.00	-0.01	0.00
Sardegna	0.00	-0.04**	0.03*	-2.91*	-3.10*	-1.27	-1.54	-2.26	-8.05*	4.64	-6.65**	5.74*

All estimates include sector and time dummies. * 1%; ** 5%; *** 10% significance levels