

# DYNAMIC MODELS IN SPACE AND TIME

J. Paul Elhorst<sup>1</sup>

SOM-Theme C: Coordination and growth in economies

## Abstract

This paper presents a first-order autoregressive distributed lag model in both space and time. It is shown that this model encompasses a wide series of simpler models frequently used in the analysis of space-time data as well as models that better fit the data and have never been used before. A framework is developed to determine which model is the most likely candidate to study space-time data. As an application, the relationship between the labor force participation rate and the unemployment rate is estimated using regional data of Germany, France and the UK derived from Eurostat, 1983-1993.

---

<sup>1</sup> Faculty of Economics, University of Groningen, P.O. Box 800, 9700 AV Groningen, The Netherlands, tel. 31-50-3633893, fax. 31-50-3637337, e-mail: [J.P.Elhorst@eco.rug.nl](mailto:J.P.Elhorst@eco.rug.nl)

## 1 INTRODUCTION

Econometric modeling of space-time data calls for quite complex stochastic specifications. Two particular problems one must deal with are serial dependence between the observations on each spatial unit over time, and spatial dependence between the observations on the spatial units at each point in time. These problems might occur, for example, when the error term of a model formulated in static form and estimated by OLS is tested for serial and spatial autocorrelation. Frequently, the null hypotheses of no serial and no spatial autocorrelation must be rejected under these circumstances. This is problematic since failing to account for serial and spatial autocorrelation when present causes the OLS estimator to lose its property of efficiency. One remedial reaction could be to re-estimate the model using methods that assume the errors are generated by a first-order serial and spatial autoregressive process. Major textbooks present this approach as an appropriate reaction, either for serial autocorrelation (Greene, 1993, Ch.15; Griffiths et al., 1993, Ch.16) or spatial autocorrelation (Anselin, 1988, Ch.8; Odland, 1988). By contrast, this approach has also been highly criticized. Hendry and Mizon (1978) were among the first to point out that serial autocorrelation correction cannot be considered a serious effort to find the 'correct' equation.<sup>2</sup> Instead of improving an initial econometric model when it appears to be unsatisfactory, one better starts with a more general model containing, nested within it as special cases, a series of simpler models that ideally should represent all the alternative economic hypotheses requiring consideration. The general model Hendry and Mizon have recommended as a generalization of the first-order serial autocorrelation model is the first-order serial autoregressive distributed lag model; a linear dynamic regression model in which  $Y_t$  is regressed on  $Y_{t-1}$ ,  $X_t$  and  $X_{t-1}$  and the error term is Gaussian white noise.

Hendry and Mizon's model is a typical time series model. Its counterpart in the spatial regression literature, the first-order spatial autoregressive distributed lag model covering the first-order spatial autocorrelation model as a special case, has been

---

<sup>2</sup>See also Gilbert (1986) on professor Hendry's econometric methodology, Mizon (1995) for a recent update, and Hendry (1995, Ch.7) for a recent update and related work.

discussed by Burridge (1981), Bivand (1984), Blommestein and Nijkamp (1986), Anselin (1988: 226-230), and Florax and Folmer (1992).

This paper presents a first-order autoregressive distributed lag model in both space and time. It is shown that this model encompasses not only a wide series of simpler models frequently used in the analysis of space-time data, but also models that better fit the data and have not yet been used. In addition, it should be stressed that while most previous studies in the analysis of space-time data are oriented toward spatial cross-section analysis, especially books dedicated to developments in spatial econometrics (Anselin, 1988; Anselin and Florax, 1995; Fischer and Getis, 1997; and Griffith et al., 1998), this paper shifts the emphasis to time series modeling techniques.

We hasten to stress that models with serial and spatial dynamic effects are not totally absent from the literature; short overviews can be found in Cressie (1991: 449-452) and Robinson (1998: 319-328). The point is that previous studies are rather restrictive. Bronars and Jansen (1987) have specified a spatiotemporal model by starting from lattice data, the two-dimensional analogue of a time series, but this type of data is rather unusual in that it oversimplifies the spatial arrangement of the data. In space, there is typically no natural order for arranging sample data. The same holds for a recent study by Pace et al. (1998). In this study it was assumed that the spatial weight matrix is lower triangular, which is rather unusual in that it implies that two spatial units cannot mutually affect each other. Earlier textbooks on spatial econometrics also dealt with autoregressive specifications in which the regressand is lagged in both space and time, but on further consideration not together in estimation unlike OLS, a method that does not take into account that lagged dependent variables can create bias problems (Cliff et al., 1975; Bartels, 1979; Bennett, 1979; Wrigley and Bennett, 1981). Other studies have struggled with the initial values of lagged variables. If the process generating the data in the sample period is stationary, or can be made so by a suitable transformation of the data, the initial values convey a great deal of information about this process since they reflect how it has operated in the past. Conditioning on those initial values or simply setting the initial values at zero

(Pfeifer and Deutsch, 1980: 41) is an undesirable feature, especially when the time dimension of the space-time data set is short.

The plan of this paper is as follows: Along with the first-order autoregressive distributed lag model in both space and time, nine simpler econometric models are presented that are subsumed by this model, some of which are frequently used in applied research. These models are then arranged in a framework to determine which is the most likely candidate to study space-time data. As an application, the relationship between the labor force participation rate and the unemployment rate is estimated using regional data of Germany, France and the UK derived from Eurostat, 1983-1993. We conclude this paper with a number of important implications for econometric modeling of relationships based on space-time data.

## 2 DYNAMIC MODELS IN SPACE AND TIME

### 2.1 A GENERAL FIRST-ORDER SERIAL AND SPATIAL AUTOREGRESSIVE DISTRIBUTED LAG MODEL

This paper focuses on a first-order serial and spatial autoregressive distributed lag model. The model is considered in vector form for a cross-section of observations at time  $t$

$$Y_t = \tau Y_{t-1} + \delta WY_t + \eta WY_{t-1} + \beta_1 X_t + \beta_2 X_{t-1} + \beta_3 WX_t + \beta_4 WX_{t-1} + u_t, \quad (1)$$

where  $Y_t$  denotes a  $n \times 1$  vector consisting of one observation for every spatial unit ( $i=1, \dots, n$ ) of the dependent variable in the  $t^{\text{th}}$  time period ( $t=1, \dots, T$ );  $X_t$  denotes a  $n \times 1$  vector of the independent variable. For reasons of simplicity, only one regressor  $X$  is considered, leaving the generalization to a set of regressors as an exercise;  $\tau$ ,  $\delta$ ,  $\eta$ ,  $\beta_1$ ,  $\beta_2$ ,  $\beta_3$  and  $\beta_4$  are the response parameters;  $u_t$  is a  $n \times 1$  vector containing the error terms and is normally distributed with  $E(u_t)=0$  and  $E(u_t u_t') = \sigma^2 I_n$ ;  $W$  denotes an  $n \times n$  weight matrix describing the spatial arrangement of the spatial units; a variable with

subscript t-1 denotes its serially lagged value, and a variable premultiplied by W denotes its spatially lagged value. In addition, we maintain the following assumption:

**Assumption 1:** (a) W is a matrix of known constants. (b) All diagonal elements of W are zero. (c) The characteristic roots of W  $\{\omega_1, \dots, \omega_n\}$  are known. (d)  $1/\omega_{\min} < \delta < 1/\omega_{\max}$  ( $\omega_{\min} < 0, \omega_{\max} > 0$ ).

Assumption 1(a) excludes the possibility that the spatial weight matrix is parametric. Assumption 1(b) implies that no spatial unit can be viewed as its own neighbor. Assumption 1(c) is needed to ensure that the log-likelihood function of the model can be computed and presupposes that the characteristic roots of W can be computed accurately based on computing technology typically available to empirical researchers. Kelejian and Prucha (1999) have pointed out that this might be problematic even for moderate sample sizes ( $n=400$ ). The explanation behind assumption 1(d) is as follows: Basically, the spatial weight matrix is symmetric. Consequently, all its characteristics roots are real, while the largest characteristic root is greater than zero and the smallest characteristic root is smaller than zero. If a spatial weight matrix is row-normalized, these properties retain (Ord, 1975). Kelejian and Robinson (1995) have pointed out that the restriction  $1/\omega_{\min} < \delta < 1/\omega_{\max}$  may nonetheless be unnecessarily restrictive since any first-order spatial model is defined for every  $\delta$  so long as the matrix  $(I-\delta W)$  is nonsingular.<sup>3</sup> Note that this matrix is singular if  $\delta$  would be equal to the reciprocal of just one of the characteristic roots of the spatial weight matrix W. The main reason to restrict  $\delta$  to  $1/\omega_{\min} < \delta < 1/\omega_{\max}$  is to facilitate the maximum likelihood estimation of  $\delta$  and to ensure invertibility of the matrix  $(I-\delta W)$  (Ord, 1981).

Figure 1 summarizes nine econometric models subsumed by (1), some of which are frequently used in applied research. These models are briefly discussed below in

---

<sup>3</sup>One simple example may illustrate this. Let  $\alpha_t = \delta W \alpha_t + a_t$  and W a spatial weight matrix of two spatial units whose off-diagonal elements are unity, as a result of which  $\omega_{\min} = -1$  and  $\omega_{\max} = 1$ . If  $a_t \sim N(0, \sigma^2 I_n)$ , then  $\alpha_t \sim N(0, (1+\delta)/(1-\delta^2) \sigma^2 I_n)$ , which shows that the variance of  $\alpha_t$  is finite when the variance of  $a_t$  is finite for every  $\delta$  unless  $\delta = 1/\omega_{\min}$  or  $\delta = 1/\omega_{\max}$ .

conjunction with the properties of the proposed general model. In section 2.2 we give an explanation for introducing lags in both space and time. In section 2.3 we discuss the conditions under which autoregressive lag models reduce to autocorrelated-error models. In section 2.4 we show that each model can be rewritten as an equilibrium correction model, a notation which is useful for interpretation. In section 2.5 we introduce a new type of model that we call a spatial equilibrium correction model. In section 2.6 we derive the maximum likelihood functions of the models in figure 1. In section 2.7 we present the conditions under which these models are stationary. Finally, we explain in section 2.8 how the likelihood function of the proposed general model might be maximized computationally.

## 2.2 THE EXPLANATION OF LAGS

There are a number of reasons why serial lags appear in econometric equations. First, a household may not change its consumption level and labor supply immediately in response to a change in prices or its income. Similarly, a firm may react with some delay to changes in costs and to changes in the demand for its product. Second, lags can arise because of imperfect information. Economic agents require time to gather relevant information, and this delays the making of decisions. There are also occasions when institutional factors can result in lags. Households may be contractually obliged to supply a certain level of labor hours, even though other conditions would indicate a reduction or increase in labor supply.

Similarly, there are a number of reasons why spatial lags appear in econometric equations. The main reason that one observation associated with a location depends on observations at other locations is that distance affects household and firm behavior. Each household may change its location, consumption and labor supply decisions, and each firm may change its location, input demand and output supply decisions, depending on the market conditions in the home region compared to other regions and on the distance to these regions. These notions have been formulated in regional science theory that relies on notions of spatial interaction and diffusion effects, hierarchies of place and spatial spillovers.

Another reason to expect both serial and spatial lags is that data collection of observations associated with spatial units over time might reflect measurement error. This would occur if the administrative boundaries for collecting information — the arbitrary delineation of space into different units (countries, states, provinces, countries, tracts or zip codes), and of time into different units (years, quarters or months) — do not accurately reflect the nature of the underlying process generating the sample data. As an example, consider the relationship between the labor force participation rate and the unemployment rate. Since laborers may travel up and down daily from one spatial unit to another, labor force participation and unemployment rates measured on the basis of where people live could exhibit spatial dependence. Similarly, as unemployed people may find a job that starts the next time period, labor force participation and unemployment rates measured on the basis of people's labor market status at a particular point in time could exhibit serial dependence.

One crucial practical question is the order of the lag structure that should be assigned to each variable. One may look to economic theory for guidance, but while economic theory is often quite helpful to entail long-run equilibrium relationships it often has little to say about the short-run dynamics of how this equilibrium is approached. In practice, dynamic factors can often only be uncovered by allowing actual data to determine the appropriate structure. This paper is restricted to lag structures of the first-order. As the emphasis in this paper is on models with both serial and spatial effects, higher order lag structures are not discussed. Nevertheless, the specified model already covers a broad range of econometric models commonly used in empirical time series analysis and spatial econometrics (see Hendry, 1995, Ch. 7; Anselin, 1988, Ch. 4).

It should be noted that some analysts have been troubled with the idea that the spatial autoregressive interaction between  $Y$  and  $WY$  is instantaneous (see Upton and Fingleton, 1985: 369). Instead, they suggest a model in which the autoregressive response is allotted a period in which to take effect,  $Y_t = \eta WY_{t-1}$ . The advantage of this specification is that the Jacobian term,  $|I - \delta W|$ , which is the result of transforming the estimation model from the error term into the dependent variable, disappears. This considerably simplifies estimation of the model by maximum likelihood. By contrast,

many other analysts do not seem to have problems with the idea that  $Y_t$  in one spatial unit is regressed on  $Y_t$  in other spatial units depending on a spatial weight matrix  $W$ ,  $Y_t = \delta W Y_t$ . For that reason we do not wish to preclude this specification in advance. Instead, we start with  $Y_t = \delta W Y_t + \eta W Y_{t-1}$  and try to determine whether the data can help to determine the most appropriate model.

### 2.3 AUTOCORRELATED-ERROR MODELS

Each first-order autoregressive distributed lag model with  $E(u_t) = 0$  and  $E(u_t u_t') = \sigma^2 I_n$  may, when a common factor is present, be reduced to a simpler regression model with fewer parameters and autocorrelated errors. This is shown below for the four models in figure 1 that satisfy these conditions:

(i) The first-order serial autoregressive distributed lag model is  $Y_t = \tau Y_{t-1} + \beta_1 X_t + \beta_2 X_{t-1} + u_t$ , which is a typical time series model. This model reduces to  $[I_n - \tau L] Y_t = [I_n - \tau L] \beta_1 X_t + u_t$ , when  $\beta_2 = -\tau \beta_1$  ( $L$  is defined such that  $LY_t = Y_{t-1}$ ), so that the terms involving  $Y_t$  and  $X_t$  have a common factor of  $(I_n - \tau L)$ . Dividing both sides of this equation by this common factor yields  $Y_t = \beta_1 X_t + \varepsilon_t$  with  $\varepsilon_t = \tau \varepsilon_{t-1} + u_t$ , from which it is seen that  $\varepsilon_t$  is generated by a first-order serial autoregressive process.

(ii) The first-order spatial autoregressive distributed lag model is  $Y_t = \delta W Y_t + \beta_1 X_t + \beta_3 W X_t + u_t$ , which is a typical spatial cross-section model. This model reduces to  $[I_n - \delta W] Y_t = [I_n - \delta W] \beta_1 X_t + u_t$ , when  $\beta_3 = -\delta \beta_1$ , so that the terms involving  $Y_t$  and  $X_t$  have a common factor of  $(I_n - \delta W)$ . Multiplying both sides of this equation by the inverse of this common factor yields  $Y_t = \beta_1 X_t + \varepsilon_t$  with  $\varepsilon_t = \delta W \varepsilon_t + u_t$ , from which it is seen that  $\varepsilon_t$  is generated by a first-order spatial autoregressive process.

(iii) The combined first-order serial and spatial autoregressive distributed lag model is  $Y_t = \tau Y_{t-1} + \delta W Y_t + \beta_1 X_t + \beta_2 X_{t-1} + \beta_3 W X_t + u_t$ . This model reduces to  $[I_n - \tau L - \delta W] Y_t = [I_n - \tau L - \delta W] \beta_1 X_t + u_t$ , when  $\beta_2 = -\tau \beta_1$  and  $\beta_3 = -\delta \beta_1$ , so that the terms involving  $Y_t$  and  $X_t$  have a common factor of  $(I_n - \tau L - \delta W)$ . Multiplying both sides of this equation by the



inverse of this common factor yields  $Y_t = \beta_1 X_t + \varepsilon_t$  with  $\varepsilon_t = \tau \varepsilon_{t-1} + \delta W \varepsilon_t + u_t$ , from which it is seen that  $\varepsilon_t$  is generated by a first-order serial and spatial autoregressive process.

(iv) The general first-order serial and spatial autoregressive distributed lag model is  $Y_t = \tau Y_{t-1} + \delta W Y_t + \eta W Y_{t-1} + \beta_1 X_t + \beta_2 X_{t-1} + \beta_3 W X_t + \beta_4 W X_{t-1} + u_t$ . This model reduces to  $[I_n - \tau L] Y_t = [I_n - \tau L] \delta W Y_t + [I_n - \tau L] \beta_1 X_t + [I_n - \tau L] \beta_3 W X_t + u_t$ , when  $\eta = -\tau \delta$ ,  $\beta_2 = -\tau \beta_1$  and  $\beta_4 = -\tau \beta_3$ , so that the terms involving  $Y_t$ ,  $W Y_t$ ,  $X_t$  and  $W X_t$  have a common factor of  $(I_n - \tau L)$ . Multiplying both sides of this equation by the inverse of this common factor yields a first-order spatial autoregressive distributed lag model  $Y_t = \delta W Y_t + \beta_1 X_t + \beta_w W X_t + \varepsilon_t$  with first-order serial autocorrelation  $\varepsilon_t = \tau \varepsilon_{t-1} + u_t$ . Similarly, this model reduces to  $[I_n - \delta W] Y_t = [I_n - \delta W] \tau Y_{t-1} + [I_n - \delta W] \beta_1 X_t + [I_n - \delta W] \beta_2 X_{t-1} + u_t$ , when  $\eta = -\delta \tau$ ,  $\beta_3 = -\delta \beta_1$  and  $\beta_4 = -\delta \beta_2$ , so that the terms involving  $Y_t$ ,  $Y_{t-1}$ ,  $X_t$  and  $X_{t-1}$  have a common factor of  $(I_n - \delta W)$ . Multiplying both sides of this equation by the inverse of this common factor yields a first-order serial autoregressive distributed lag model  $Y_t = \delta W Y_t + \beta_1 X_t + \beta_w W X_t + \varepsilon_t$  with first-order spatial autocorrelation  $\varepsilon_t = \delta W \varepsilon_t + u_t$ .

The reason to consider the combined, as well as the general, first-order serial and spatial autoregressive distributed lag model becomes clear by listing these cases. Whereas the combined model may be reduced to a first-order serial *and* spatial autocorrelation model when a common factor is present, the general model may not. Conversely, whereas the general model may be reduced to an autoregressive distributed lag model with first-order serial *or* spatial autocorrelation model when a common factor is present, the combined model may not.

Obviously, autoregressive distributed lag models are more general than autocorrelated-error models. One reason to test for autocorrelated-error models is that a gain in statistical efficiency results when the number of parameters to be estimated reduces. At the same time, it should be stressed that the rejection of a restricted model in favor of a more general model, when it is true, is a less serious error than using a restricted model when the restrictions imposed are incorrect.

## 2.4 REPARAMETERIZING THE MODELS TO EQUILIBRIUM CORRECTION MODELS

While equilibrium economic theories often entail static equations, imposing a static equation on data restricts short-run and long-run response of  $Y$  to  $X$  to be identical across space and instantaneous in time. Nevertheless, it seems preferable to require that a dynamic model reproduces the static equation under equilibrium conditions; this restricts the type of model but not the range of dynamic responses. For this purpose, we consider a reparameterizing of the non-stochastic part of equation (1), known as an equilibrium correction model.

As the principle of the equilibrium correction model is a product of the econometric time series literature, it is helpful to first consider the equilibrium correction model of a first-order serial autoregressive distributed lag model. The first-order serial autoregressive distributed lag model is  $Y_t = \tau Y_{t-1} + \beta_1 X_t + \beta_2 X_{t-1} + u_t$ . This equation may be equivalently reformulated as

$$Y_t = -\frac{\tau}{1-\tau} \Delta Y_t + \frac{\beta_1 + \beta_2}{1-\tau} X_t - \frac{\beta_2}{1-\tau} \Delta X_t = \tau^* \Delta Y_t + \beta_1^* X_t + \beta_2^* \Delta X_t + \left(\frac{1}{1-\tau} u_t\right). \quad (2)$$

This equation demonstrates that ‘short-run dynamics’ have been added to the static equation. There still exists a static long-run equilibrium relationship between  $Y$  and  $X$ , but short-run dynamics of how equilibrium is approached are explicitly taken into account.  $\beta_1^*$  reflects the long-run effect of  $Y$  with respect to  $X$ , while  $\beta_2^*$  reflects the short-run or immediate response of  $Y$  to a change in  $X$ . Wickens and Breusch (1988) have shown that a point estimate of the long-run coefficients ( $\beta_1^*$ ) and its covariance matrix can be obtained directly by using an instrumental variables estimator. However, as this property does not hold in a spatial econometric model due to the Jacobian term  $|\mathbf{I} - \delta \mathbf{W}|$ , we leave this aside.

The first-order spatial autoregressive distributed lag model,  $Y_t = \delta \mathbf{W} Y_t + \beta_1 X_t + \beta_3 \mathbf{W} X_t + u_t$ , may be equivalently reformulated as

$$Y_t = [\beta_1 (\mathbf{I}_N - \delta \mathbf{W})^{-1} + \beta_3 (\mathbf{I}_N - \delta \mathbf{W})^{-1} \mathbf{W}] X_t + (\mathbf{I} - \delta \mathbf{W})^{-1} u_t. \quad (3)$$

From this equation it can be seen that a spatial unit in a spatial autoregressive distributed lag model is not only influenced by its local conditions, but also by those

of its neighbors dependent on the structure of the spatial weight matrix  $W$ . Furthermore, the impact of these conditions is not necessarily uniform across spatial units. Assuming that there are  $n$  spatial units and  $k$  regressors, we get  $n \times k$  different 'long-run' parameter estimates and  $n \times k$  different standard errors or t-values. It is clear that the amount of output might be a problem of this model. Even for small values of  $n$  and  $k$ , it may already be rather difficult to report the estimation results compactly. Furthermore, due to this variety in the 'long-run' coefficients, spatial autoregressive distributed lag models are difficult to compare with static or serial autoregressive distributed lag models, as in these models the effect of  $X$  on  $Y$  can be represented by just one long-run coefficient to be valid for all spatial units. In sum, whereas the preference for a model with a spatially lagged dependent variable may improve the plausibility of the estimation results, the preference for a model without a spatially lagged dependent variable may improve the surveyability of the estimation results.

The equilibrium correction model of the general first-order serial and spatial autoregressive distributed lag model is

$$(I - \tau L - \delta W - \eta W) Y_t = -(\tau + \eta W) \Delta Y_t + (\beta_1 + \beta_2) X_t + (\beta_3 + \beta_4) W X_t - \beta_2 \Delta X_t - \beta_4 W \Delta X_t. \quad (4a)$$

This model implies the following static long-run equilibrium relationship between  $Y$  and  $X$ ,

$$Y_t = [(\beta_1 + \beta_2)(I_n - \tau I_n - \delta W - \eta W)^{-1} + (\beta_3 + \beta_4)(I_n - \tau I_n - \delta W - \eta W)^{-1} W] X_t, \quad (4b)$$

whose properties can best be compared with those of the spatial autoregressive distributed lag model.

## 2.5 A SPATIAL EQUILIBRIUM CORRECTION MODEL

Another interesting model occurs when both  $\delta + \eta = 0$  and  $\beta_3 + \beta_4 = 0$ . When these restrictions hold, the general first-order serial and spatial autoregressive distributed lag model reduces to  $Y_t = \tau Y_{t-1} + \delta W \Delta Y_t + \beta_1 X_t + \beta_2 X_{t-1} + \beta_3 W \Delta X_t + u_t$ . Consequently, the static long-run equilibrium relationship between  $Y$  and  $X$  reduces to

$$Y_t = \frac{\beta_1 + \beta_2}{1 - \tau} X_t = \beta_1^* X_t. \quad (5)$$

Although this static long-run equilibrium relationship is equivalent to that of the first-order serial autoregressive distributed lag model, the equilibrium correction model is different. Whereas one spatial unit in a time series model is said to be in equilibrium when  $\Delta Y$  and  $\Delta X$  both converge to zero, a spatial unit in this model is said to be in equilibrium only when  $\Delta Y$  and  $\Delta X$  in all spatial units within the range of influence of that spatial unit converge to zero. Despite this generalization, the long-run effect of  $X$  on  $Y$  can still be represented by just one coefficient to be valid for all spatial units. This has the advantage that this model is better comparable with a static or serial autoregressive distributed lag model. To distinguish this model from an equilibrium correction model, which in the time series literature is usually associated with serial autoregressive distributed lags, it is called a spatial equilibrium correction model.

## 2.6 LIKELIHOOD FUNCTIONS

The way to estimate the static model, the first-order serial and the first-order spatial autocorrelation models, as well as the first-order serial and the first-order spatial autoregressive distributed lag models has been widely discussed in the literature. The way to estimate the other models in figure 1 has not been (widely) discussed before.

All models given in figure 1 can best be estimated by maximum likelihood. We have two reasons for that. The first is that estimation by OLS should not be recommended. The time series literature has shown that the inclusion of a serial lagged dependent variable among the regressors causes the OLS estimator to lose its property of unbiasedness. By contrast, the OLS estimator remains consistent.<sup>4</sup> The spatial econometric literature has shown that the inclusion of a spatial lagged dependent variable among the regressors not only causes the OLS estimator to lose its property of unbiasedness, but also its consistency. By far the most commonly suggested method to overcome this problem is to estimate the model by maximum likelihood (see Anselin, 1988, pp. 181-182; Anselin and Hudak, 1992).

To be able to compare the estimation results of the different models in figure 1, they should be estimated, preferably, on the same set of observations and by the same

---

<sup>4</sup>Provided the sample size is large enough (usually over 30 d.f., see Ramanathan, 1995: 545).

estimation method. Regression equations that include variables lagged one period in time are often estimated by OLS conditional upon the first observation(s). By contrast, when estimating these models by ML it is also possible to obtain unconditional results. Additionally, it is also possible to test restricted against unrestricted models with the help of the Lagrange Multiplier (LM) test, the Wald test and the Likelihood Ratio (LR) test. The models 0,1,2,3 and 5 in figure 1 include explanatory variables lagged one period in time. If these models are estimated conditional upon the first observations, while the models 4,6,7,8 and 10 are estimated upon all observations, we create a dividing line in figure 1 which makes it more difficult to use these tests. This, taken together, is the second reason to choose ML.

Below, we derive the unconditional log-likelihood function of the general first-order serial and spatial autoregressive lag model. After that, the log-likelihood functions of the other models are derived by setting the right restrictions (see figure 1).

The log-likelihood function conditional upon the vector  $Y_1$  of first observations is

$$\text{Log } f_{Y_T, Y_{T-1}, \dots, Y_2 | Y_1} = -\frac{1}{2} n(T-1) \log(2\pi\sigma^2) + (T-1) \log |I - \delta W| - \frac{1}{2\sigma^2} \sum_{t=2}^T e_t' e_t, \quad (6)$$

where the second log at the right-hand side corresponds to the Jacobian term from the error term to the dependent variable and  $e_t = Y_t - \tau Y_{t-1} - \delta W Y_{t-1} - \eta W Y_{t-1} - \beta_1 X_t - \beta_2 X_{t-1} - \beta_3 W X_t - \beta_4 W X_{t-1}$ . To obtain the unconditional log-likelihood function, we rewrite the general first-order serial and spatial autoregressive distributed lag model as

$$B Y_t = \beta_X + A Y_{t-1} + u_t, \quad \text{with } B = I - \delta W, \quad A = \tau I + \eta W, \quad \text{and } \beta_X = \beta_1 X_t + \beta_2 X_{t-1} + \beta_3 W X_t + \beta_4 W X_{t-1}. \quad (7)$$

To proceed, we repeatedly lag this equation by one period. For  $m \geq 1$  we get

$$B Y_{t-m} = \beta_X + A Y_{t-(m+1)} + u_{t-m} \quad \text{or} \quad Y_{t-m} = B^{-1} \beta_X + B^{-1} A Y_{t-(m+1)} + B^{-1} u_{t-m}. \quad (8)$$

Then, by substitution of  $Y_{t-1}$  into (7), next  $Y_{t-2}$  into (7) up to  $Y_{t-m}$  into (7), we get

$$BY_t = (\beta_X + u_t) + AB^{-1}(\beta_X + u_{t-1}) + (AB^{-1})^2(\beta_X + u_{t-2}) + \dots + (AB^{-1})^m(\beta_X + u_{t-m}) + (AB^{-1})^m AY_{t-(m+1)}. \quad (9)$$

If  $|AB^{-1}| < 1$ , the term  $(AB^{-1})^m A$  of the last right-hand side variable approaches zero as  $m$  approaches infinity. Thus, if we assume that  $|AB^{-1}| < 1$  and that the process started a very long time ago,

$$BY_t = (I + AB^{-1} + (AB^{-1})^2 + \dots)\beta_X + (u_t + AB^{-1}u_{t-1} + (AB^{-1})^2u_{t-2} + \dots), \quad (10)$$

which implies that  $E(BY_t) = (I - AB^{-1})^{-1}\beta_X$  and that  $E(Y_t) = (B - A)^{-1}\beta_X$ . Since the successive values of  $u_t$  are uncorrelated, the variance of  $BY_t$  under these assumptions is

$$\begin{aligned} \text{Var}(BY_t) &= E(BY_t - E(BY_t))(BY_t - E(BY_t))' = \\ &= E(u_t u_t') + AB^{-1}E(u_{t-1} u_{t-1}') (AB^{-1})' + (AB^{-1})^2 E(u_{t-2} u_{t-2}') ((AB^{-1})')^2 + \dots \\ &= \sigma^2 (I + AB^{-1} (AB^{-1})' + (AB^{-1})^2 ((AB^{-1})')^2 + \dots) = \sigma^2 (I - AB^{-1} (AB^{-1})')^{-1}, \end{aligned} \quad (11)$$

which implies that

$$\text{Var}(Y_t) = B^{-1} \sigma^2 (I - AB^{-1} (AB^{-1})')^{-1} (B^{-1})' = \sigma^2 (B' B - B' A B^{-1} (B' A B^{-1})')^{-1}. \quad (12)$$

It further follows that the  $j^{\text{th}}$  autocovariance is

$$E(Y_t - E(Y_t))(Y_{t-j} - E(Y_{t-j}))' = (AB^{-1})^j \sigma^2 (B' B - B' A B^{-1} (B' A B^{-1})')^{-1}. \quad (13)$$

In sum, the vector process  $Y_t$  has a constant unconditional mean and a constant unconditional variance, independent of time, while the autocovariances depend only on the length of time separating the observations, and not on the date of observation. A process satisfying these conditions is said to be covariance stationary.

Consider the probability function of  $Y_1$ , the vector of first observations in the sample. Under the model assumptions  $Y_1$  is a random variable with mean  $E(Y_1) =$

$(B-A)^{-1}\beta_X$  and variance  $\text{Var}(Y_1)=\sigma^2(B'B-B'AB^{-1}(B'AB^{-1})')^{-1}$ . Hence, the log-likelihood function of the vector of first observations takes the form

$$\begin{aligned} \text{Log } f_{Y_1} = & -\frac{1}{2}n\log(2\pi\sigma^2) + \frac{1}{2}\log|B'B-B'AB^{-1}(B'AB^{-1})'| + \\ & -\frac{1}{2\sigma^2}(Y_1-(B-A)^{-1}\beta_X)'(B'B-B'AB^{-1}(B'AB^{-1})')(Y_1-(B-A)^{-1}\beta_X), \end{aligned} \quad (14)$$

where  $\beta_X=(\beta_1+\beta_2)X_1+(\beta_3+\beta_4)WX_1$  and the log of  $|B'B-B'AB^{-1}(B'AB^{-1})'|$  corresponds to the Jacobian term from  $u_1$  to  $Y_1$ .

The determinants of the matrices  $|I-\delta W|=|B|$  in equation (6) and  $|B'B-B'AB^{-1}(B'AB^{-1})'|$  in equation (14) can be expressed in function of the characteristic roots of the spatial weight matrix  $W$ . To see this, we make use of five matrix properties taken from Griffith (1988: 44, table 3.1): (i) if a matrix is multiplied by some scalar constant, then its characteristic roots are also multiplied by this constant; (ii) if  $bI$  is added to a matrix, where  $b$  is a real scalar, then  $b$  is added to each of the characteristic roots of that matrix; (iii) the characteristic roots of a matrix and its transpose are the same; (iv) the characteristic roots of a matrix and its inverse are inverses of each other; and (v) if a positive definite matrix is powered by some real number, each of its characteristic roots is powered by this same real number (see also Greene, 1993: 34-36). Taken together, the characteristic roots of the matrix  $B$  are  $1-\delta\omega_i$  ( $i=1,\dots,n$ ) and of the matrix  $B'B-B'AB^{-1}(B'AB^{-1})'$  are  $(1-\delta\omega_i)^2-(\tau+\eta\omega_i)^2$  ( $i=1,\dots,n$ ).

The unconditional log-likelihood function of the complete sample of size  $T$  is thus seen to be

$$\begin{aligned} \text{Log } f_{Y_T,\dots,Y_1} = & -\frac{1}{2}nT\log(2\pi\sigma^2) + \frac{1}{2}\sum_{i=1}^n \log((1-\delta\omega_i)^2-(\tau+\eta\omega_i)^2) + \\ & + (T-1)\sum_{i=1}^n \log(1-\delta\omega_i) - \frac{1}{2\sigma^2}\sum_{t=2}^T e_t' e_t + \\ & -\frac{1}{2\sigma^2}e_1'((B-A)')^{-1}(B'B-B'AB^{-1}(B'AB^{-1})')^{-1}(B-A)^{-1}e_1, \end{aligned} \quad (15)$$

where  $e_1=(B-A)Y_1-\beta_X=Y_1-\delta WY_1-(\tau+\eta W)Y_1-(\beta_1+\beta_2)X_1-(\beta_3+\beta_4)WX_1$  and  $e_t=Y_t-\tau Y_{t-1}-\delta WY_{t-1}-\eta WY_{t-1}-\beta_1 X_t-\beta_2 X_{t-1}-\beta_3 WX_t-\beta_4 WX_{t-1}$  ( $t>1$ ).

Table 1 reports the log-likelihood functions of the restricted models.

## 2.7 STATIONARITY AND STABILITY

Since a space-time data set has two dimensions, it is possible to consider asymptotic behavior as  $n \rightarrow \infty$ ,  $T \rightarrow \infty$ , or both. Generally speaking, it is easier to increase the cross-section dimension of a space-time data set. If as a result  $n \rightarrow \infty$  is believed to be the most relevant asymptotics, it is not necessary to assume  $|AB^{-1}| < 1$  as long as  $T$  is fixed. By contrast, Nerlove (1999) has pointed out that the cross-section of first observations conveys a great deal of information about the process generating the data since these observations reflect how that process has operated in the past. Thus, conditioning on the cross-section of first observations is an undesirable feature, especially when the time dimension of the space-time data set is short.

When it is assumed that  $|AB^{-1}| < 1$ , the process generating the data is stationary in time. A more detailed description of this condition is given in table 2, which also reports the stationarity conditions for the restricted models. It can be seen that these conditions are also captured by the log-likelihood functions in that these functions are not defined for parameter values that do not satisfy these conditions.

In contrast to stationarity in time, stationarity in space is quite difficult to impose. Quite recently, Kelejian and Prucha (1998, 1999) formulated one necessary condition which must be satisfied: the row and column sums of the spatial weight matrix  $W$  must be bounded uniformly in absolute value as  $n \rightarrow \infty$ . When the spatial weight matrix is a binary contiguity matrix, this condition is satisfied. Normally, no spatial unit is assumed to be a neighbor to more than a given number, say  $q$ , of other spatial units. When the spatial weight matrix is an inverse distance matrix, this condition is not automatically satisfied. Consider an infinite number of spatial units that are linearly arranged. The distance of each spatial unit to its first left and right neighbor is 1, to its second left and right neighbor is 2, and so on. When  $W$  is an inverse distance matrix and all its off-diagonal elements are  $1/d_{ab}$  where  $d_{ab}$  is the distance between two spatial units  $a$  and  $b$ , the sum of each row of  $W$  is  $\sum 2/d$ , a series which is infinite. Note that  $\sum 2/d^\alpha$  would converge if and only if  $\alpha > 1$ . In sum, when  $n$



is fixed and  $T \rightarrow \infty$ , the vector process  $Y$  is stationary, but when  $n \rightarrow \infty$  and  $T$  is fixed the weights formulated as a function of distance between spatial units must decline rapidly; otherwise standard large sample theory may not be applied. In addition, it should be realized that a spatial process, though stable, may not be stationary (see Haining, 1990: 82; Kelejian and Robinson, 1995: 78). This is a separate problem beyond the scope of this paper.

## 2.8 ON MAXIMIZING THE LOG-LIKELIHOOD FUNCTION

Since direct estimation of the log-likelihood function with respect to the full set of parameters failed, an iterative two-step procedure has been adopted. A problem is that this procedure is not applicable when the spatial weight matrix is asymmetric due to row-normalization.<sup>5</sup>

Divide the parameter set into  $\tau$ ,  $\delta$  and  $\eta$  on the one hand, and  $\beta=(\beta_1,\beta_2,\beta_3,\beta_4)$  and  $\sigma^2$  on the other. First, an initial estimate of  $\beta$  and  $\sigma^2$  is needed. This can be obtained by estimating equation (1) by OLS conditional upon the first observations or, alternatively, upon all observations and by setting  $Y_0=Y_1$  and  $X_0=X_1$ . This initial estimate should satisfy the conditions stated in table 2; otherwise another initial estimate must be chosen. On inserting the maximizing values of  $\beta$  and  $\sigma^2$  into the log-likelihood function, the concentrated log-likelihood function of  $\tau$ ,  $\delta$  and  $\eta$  is obtained, which then is maximized. Next, the first observations must be transformed to obtain the maximum likelihood estimates of  $\beta$  and  $\sigma^2$ , given  $\tau$ ,  $\delta$  and  $\eta$ ; These can be obtained by applying feasible GLS. Finally, these successive steps must be repeated until convergence.

The most difficult step is to maximize the concentrated log-likelihood function of  $\tau$ ,  $\delta$  and  $\eta$ . In most commercial econometric software, the built-in maximum likelihood procedure expresses the log-likelihood function in terms that correspond to each observation. The overall log-likelihood is then simply the sum of these terms,  $\text{LogL}=\sum_t \sum_i \log L_{ti}$ . From equation (15) it can be seen that the characteristic roots of the

---

<sup>5</sup>This problem only holds for the general first-order serial and spatial autoregressive distributed lag model and for the spatial equilibrium correction model.

matrices  $B$  and  $B'B-B'AB^{-1}(B'AB^{-1})'$  can be expressed in this form by introducing them as additional pseudo variables; the order in which these characteristic roots are listed is irrelevant, since they appear separately from the other variables in the log-likelihood (see Anselin and Hudak, 1992: 532). However, the transformation on the vector of first observations is more difficult to express in this form.

When the spatial weight matrix is symmetric,  $W=W'$ , we get  $A=A'$ ,  $B=B'$  and  $B'B-B'AB^{-1}(B'AB^{-1})'=B^2-A^2=(B+A)(B-A)$ . Consequently, the last right-hand side term of equation (15) reduces to

$$-\frac{1}{2\sigma^2}((B-A)Y_1 - \beta_x)'(B-A)^{-1}(B+A)((B-A)Y_1 - \beta_x). \quad (16)$$

As  $(B-A)Y_1 - \beta_x = Y_1 - \tau Y_1 - (\delta + \eta)WY_1 - (\beta_1 + \beta_2)X_1 - (\beta_3 + \beta_4)WX_1$ , the variables  $Y_1$ ,  $WY_1$ ,  $X_1$  and  $WX_1$  must be multiplied by  $P = [(B-A)^{-1}(B+A)]^{1/2}$ . One important property of the matrix  $P$  is that it has the same characteristic vectors as the matrix  $W$  (see Griffith: 44, table 3.1, and Greene, 1993: 34-36). Let  $C$  be a  $n \times n$  matrix whose  $i^{\text{th}}$  column is the characteristic vector  $c_i$  of the spatial weight matrix  $W$ ,  $C = [c_1 \ c_2 \ \dots \ c_n]$ , and let  $\Gamma$  be a diagonal matrix whose  $i^{\text{th}}$  diagonal element is  $r_i$  where  $r_i$  is the  $i^{\text{th}}$  characteristic root of  $P$  that corresponds to  $c_i$ , i.e.,  $r_i$  is the square root of  $(1 - \delta\omega_i)/(\tau + \eta\omega_i)$ . Then we may also use the transformation matrix  $\Gamma C'$  instead of  $P$ . This is because the determinant of  $\Gamma C'$  equals the determinant of  $P$  and  $P'P = (\Gamma C')'\Gamma C' = C\Gamma\Gamma C' = C\Gamma^2 C' = (B-A)^{-1}(B+A)$ . Since the matrix of characteristic vectors  $C$  is independent of  $\tau$ ,  $\delta$  and  $\eta$ , part of the transformation  $\Gamma C'$  on the vector of first observations can be computed in advance by determining the variables  $C'Y_1$ ,  $C'WY_1$ ,  $C'X_1$  and  $C'WX_1$ . When computing the concentrated log-likelihood function of  $\tau$ ,  $\delta$  and  $\eta$ , these variables then need only to be multiplied by the matrix  $\Gamma$ . Since  $\Gamma$  is a diagonal matrix, the  $i^{\text{th}}$  row of  $C'Y_1$ ,  $C'WY_1$ ,  $C'X_1$  and  $C'WX_1$  must in fact be multiplied by  $r_i$ , which can be programmed for each observation separately. The advantage of this solution is that we can use again one of the built-in maximum likelihood procedures of commercial econometric software. The disadvantage is that we also need to know the characteristic vectors of the spatial weight matrix  $W$ , which might be problematic when  $W$  is large.

### 3 NUMERICAL ILLUSTRATIONS

In this section we estimate the relationship between the labor force participation rate and the unemployment rate at regional level. The estimation of this relationship is primarily meant as an empirical illustration and can be seen as a necessary, though not sufficient, step to a complete model of the labor force participation rate. In this respect two remarks are important. First, we have limited the set of regressors to the regional unemployment rate only. We have chosen this variable, as it is part of almost every empirical study on regional labor force participation; Other regressors are less universally adopted (see Elhorst, 1996). Furthermore, its estimate is usually quite robust. Most studies have found that the effect of the unemployment rate on the labor force participation rate is negative. Second, we abstract from a potential simultaneity bias in that the labor force participation rate and unemployment rate in regional labor markets are being determined simultaneously. On the other hand, this is quite common. In an overview paper Elhorst (1996) has pointed out that the majority of studies has treated the unemployment rate as an exogenous variable.

We have used regional data of France, Germany and the UK over the period 1983-1993. These data are harmonized series produced by Eurostat intended to be comparable among EU member states as well as to give a consistent picture of labor force participation and unemployment over time. We have also used Eurostat's regional division on the NUTS2 level, with the exception of the UK, for which Eurostat only collects data on the NUTS1 level (31 regions in Germany excluding the five new German Länder, 22 in France and 11 in the UK). Finally, we have opted for a country-specific analysis to see whether we get similar or different results for each country.

The spatial weight matrix used in the estimations is a symmetric normalized inverse distance matrix. Except that the proposed iterative two-step estimation procedure is not applicable if the spatial weight matrix is asymmetric, an inverse distance matrix should not be row-normalized, since scaling the rows so that the weights sum to one would cause this matrix to lose its economic interpretation of distance decay. On the other hand, an inverse distance matrix can still be normalized

so that its largest characteristic root equals unity,  $\omega_{\max}=1$ , just like the largest characteristic root of a row-normalized spatial weight matrix. Let  $\omega_{\max}$  denote the largest characteristic root of an inverse distance matrix that is not yet normalized. Dividing the elements of this spatial weight matrix  $W$  by  $\omega_{\max}$ ,  $W^*=1/\omega_{\max}W$ , has the effect that  $\omega_{\max}^*=1$  and  $-1 \leq \omega_{\min}^* < 0$ .

Table 1 reports the estimation results of all models that have been considered. The following observations are typical for the specific-to-general approach. The starting point of this approach is the static model at the right-hand side of figure 1. The LM test is used to determine whether a model is acceptable on the data or must be extended.<sup>6</sup> It appears that the absence of serial and spatial autocorrelation in the static model must be rejected. Thereupon the static model, corrected for serial or spatial autocorrelation, will be estimated. It appears that the absence of serial autocorrelation in the static model corrected for spatial autocorrelation, as well as the absence of spatial autocorrelation in the static model corrected for serial autocorrelation, must also be rejected. In sum, the static model must be corrected for both serial and spatial autocorrelation. These findings hold for every country.

To consider the static model corrected for both serial and spatial autocorrelation the result of the specific-to-general approach is probably extreme. It could be that an investigator familiar with time series analysis will also consider the serial autoregressive distributed lag model and that an investigator familiar with spatial econometrics will also consider the spatial autoregressive distributed lag model.

From table 1 it appears that the restriction  $\beta_2 = -\tau\beta_1$ , which would simplify the serial autoregressive distributed lag model to a static model with serial autocorrelation, should be rejected. This corroborates that the practice of correcting the static model for serial autocorrelation should not be recommended. At the same time, it appears that the problem of spatial error dependence remains. In other words, by only specifying dynamics in time and not in space, the relationship remains misspecified.

---

<sup>6</sup>If the extension consists of one parameter, the Lagrange Multiplier test then has a chi-squared distribution with 1 degree of freedom. The 5 percent critical value of this distribution is 3.8.

Similarly, it appears that the restriction  $\beta_3 = -\delta\beta_1$ , which would simplify the spatial autoregressive distributed lag model to a static model with spatial autocorrelation, should be rejected. This corroborates that the practice of correcting the static model for spatial autocorrelation should not be recommended. At the same time, it appears that the problem of serial error dependence remains. In other words, by only specifying dynamics in space and not in time, the relationship remains misspecified.

If the two strategies were combined (but to date the number of investigators doing so is rather small) it would appear that the restrictions  $\beta_2 = -\tau\beta_1$  and  $\beta_3 = -\delta\beta_1$  also must be rejected simultaneously. Hence, correcting the static model for both serial and spatial autocorrelation should not be recommended either.

In conclusion we may say that the specific-to-general approach indicates that both serial and spatial effects matter. Only the shape these effects may take is unclear, since this approach may result in three different models that are non-nested and therefore difficult to compare. These models are the combined serial and spatial autoregressive distributed lag model, the serial autoregressive distributed lag model corrected for spatial autocorrelation and the spatial autoregressive distributed lag model corrected for serial autocorrelation.

Under these circumstances it may be helpful to follow the general-to-specific approach. The starting point of this approach is the general serial and spatial autoregressive distributed lag model at the left-hand side of figure 1. The Wald test is used to test for restrictions on the model.<sup>7</sup> From table 1 it can be seen that the restrictions that would reduce this model to a combined serial and spatial autoregressive distributed lag model or a spatial autoregressive distributed lag model corrected for serial autocorrelation must be rejected. Although not reported, the restrictions  $\delta=0$  and  $\beta_3=0$  also must be rejected. This means that a model in which the response of WY and WX is allotted a period in which to take effect is also not acceptable on the data. By contrast, the restrictions that would reduce the general

---

<sup>7</sup>The Wald test has a chi-squared distribution with n degrees of freedom, where n stands for the number of restrictions. The 5 percent critical value of this distribution is 3.8 if n=1, 6.0 if n=2, and 7.8 if n=3. The 1 percent critical value of this distribution is 6.6 if n=1, 9.2 if n=2, and 11.3 if n=3.

serial and spatial autoregressive distributed lag model to a serial autoregressive distributed lag model corrected for spatial autocorrelation are acceptable on the data for France and the UK. They are not for Germany, but if these restrictions are tested on the basis of the LR test, the serial autoregressive distributed lag model corrected for spatial autocorrelation appears to be acceptable also for Germany (though only at 1% significance<sup>8</sup>). This is one of the problems of the general-to-specific approach. Although the Wald test does not require the restricted estimator, it does not always help to find the correct model.<sup>9</sup>

Another model that appears to be acceptable on the data is the spatial equilibrium correction model. The two restrictions  $\delta + \eta = 0$  and  $\beta_3 + \beta_4 = 0$  are acceptable on the data for all countries, on the basis of both the Wald test and the LR test at 1% significance (in three of these six cases also at 5% significance).

A final subject for discussion is whether the static model is really misspecified. On calculating the long-run effect of the unemployment rate on the labor force participation rate for the two models that are acceptable on the data, we obtained the following results: (i) -0.863 for France, -1.231 for Germany, and -0.957 for the UK in the case of the serial autoregressive distributed lag model corrected for spatial autocorrelation; and (ii) -0.891 for France, -1.441 for Germany, and -1.076 for the UK in the case of the spatial equilibrium correction model. These long-run effects are different from the long-run effects in the static model (the coefficient  $\beta_1$ ), but not to a great degree. The explanation might be that the static model stands for a cointegrating relationship between two variables that are nonstationary in levels but stationary after

---

<sup>8</sup>Note that the LR test is the most labor-intensive test, since it can only be carried out after both models have been estimated. The LR test is based on two times the difference between the value of the log-likelihood function in the unrestricted model and the value of the log-likelihood function of the restricted model. This test statistic has a chi-squared distribution with degrees of freedom equal to the number of restriction imposed. In this case we get  $2 * (\log L_{\text{unrestricted}} - \log L_{\text{restricted}}) = 2 * (1034.9 - 1029.5) = 10.8$ . As we have 3 restrictions, the LR test has a chi-squared distribution with 3 degrees of freedom; the 1 percent critical value of this distribution is 11.3.

<sup>9</sup>In this case the correct model is unknown, but see Florax et al. (1998) for a comparison of the specific-to-general and general-to-specific approaches applied to linear models with spatial effects (not serial effects) based on Monte Carlo experiments.

being first differenced.<sup>10</sup> Engle and Granger (1987) pointed out that the OLS estimator of the regression coefficients in a linear regression model between two level variables that are integrated of the same order is a superconsistent estimator. Consequently, the OLS (=ML) estimator of the regression coefficients in the static model may be quite accurate. The spatial analogue of this finding can be found in Fingleton (1999). Nevertheless, there is evidence that superconsistent estimates may have substantial finite-sample biases and that autoregressive distributed lag models produce better estimates of the long-run coefficients (Banerjee et al., 1993, Ch.7).

#### 4 CONCLUSIONS

There are a number of important implications for econometric modeling of space-time data to be drawn from the analysis in this paper. As both serial and spatial effects are likely to be present in the analysis of space-time data, economic relationships to be estimated with the help of space-time data are better not formulated in static form. Admittedly, this has only been investigated for the relationship between the labor force participation rate and the unemployment rate at regional level. But under normal circumstances, i.e. during periods without wars or climatic catastrophes, economic variables observed at spatial units hardly change over time (or only gradually). The same holds for their spatial distributions. Consequently, it is not unreasonable to suspect serial and spatial effects to be present in the analysis of space-time data in broad terms.

Re-estimating relationships formulated in static form but corrected for first-order serial autocorrelation, first-order spatial autocorrelation, or both, should not be recommended either. One may better start from a first-order autoregressive lag model

---

<sup>10</sup>It is beyond the scope of this paper to determine the order of integration of the labor force participation rate and the unemployment rate. One reason is that a widely accepted panel data unit root test to determine the order of integration of space-time data is not yet available (see Maddala and Kim, 1998, for an overview of panel data unit root tests).

in both space and time and then try to narrow it down by looking for simplifications that are acceptable on the data, better known as the general-to-specific approach.

As both serial and spatial effects are likely to be present in the analysis of space-time data, a relationship formulated as a first-order autoregressive lag model in space or time only will generally still be misspecified. This implies that the investigator familiar with time series analysis should not be blind to spatial dependence between the observations on the spatial units at each point in time, and that the investigator familiar with spatial econometrics should not be blind to serial dependence between the observations on each spatial unit over time.

When looking for a relationship that captures both serial and spatial effects, it might be that a spatial autoregressive distributed lag model corrected for serial autocorrelation, or a serial autoregressive distributed lag model corrected for spatial autocorrelation, is acceptable on the data. Another model that might be acceptable on the data is the spatial equilibrium model which up to now has not been used in empirical research. Tests on the relationship between the labor force participation rate and the unemployment rate at regional level have shown that the last two models are acceptable on the data, indicating that in this particular case serial dynamic effects are more important than spatial dynamic effects.

A subject for further research is the estimation of spatiotemporal models specified as a typical panel data model. The two most widely used panel data models are the fixed effects and random effects models, which have in common that the intercept is assumed to be variable. The reason these models have not yet been considered is that estimation becomes complicated if both a spatial and a serial lagged dependent variable are part of the explanatory variables. This is because the serial lagged dependent variable causes the fixed effects estimator to lose its property of consistency if  $T$  is fixed, no matter how large  $n$  is, and the random effects estimator to lose its property of consistency even if both  $n$  and  $T$  tend to infinity (see Hsiao, 1986, Ch.4; Baltagi, 1995, Ch.8). Consequently, estimation by maximum likelihood is not straightforward. The alternative, first differencing the model to eliminate the variable intercept and then estimation by the generalized method-of-moments (GMM) using a set of appropriate instruments, is not straightforward either. The problem is that the



GMM procedure described in the literature to estimate serial dynamic panel data models (e.g. Baltagi, 1995: 126-128) does not take into account the conditions on the parameters implied in the Jacobian term. A similar problem occurs if a spatially lagged dependent variable is added to the regression. If the restriction  $1/\omega_{\min} < \delta < 1/\omega_{\max}$  implied in the Jacobian term  $|I-\tau W|$  is not included, the spatial autoregressive parameter  $\delta$  tends to be overestimated (Anselin, 1980: 88).

## REFERENCES

- Anselin L. (1988) *Spatial Econometrics: Methods and Models*. Kluwer, Dordrecht.
- Anselin L., Florax R.J.G.M. (1995, eds.) *New Directions in Spatial Econometrics*. Springer, Berlin.
- Anselin L., Hudak S. (1992) Spatial Econometrics in Practice; A Review of Software Options. *Regional Science and Urban Economics* 22: 509-536.
- Baltagi B.H. (1995) *Econometric Analysis of Panel Data*. Wiley, Chichester.
- Banerjee A., Dolado J.J., Galbraith J.W., Hendry D.F. (1993) *Cointegration, Error-Correction, and the Econometric Analysis of Non-Stationarity Data*. Oxford University Press, Oxford.
- Bartels C.P.A. (1979) Operational Statistical Methods for Analysing Spatial Data. In: Bartels C.P.A., Ketellapper R.H. (eds.) *Exploratory and Explanatory Statistical Analysis of Spatial Data*. Martinus Nijhoff, Boston.
- Bennett R.J. (1979) *Spatial Time Series*. Pion, London.
- Bivand R.S. (1984) Regression Modeling with Spatial Dependence: An Application of Some Class Selection and Estimation Methods. *Geographical Analysis* 16: 25-37.
- Blommestein H., Nijkamp P. (1986) Testing the Spatial Scale and the Dynamic Structure in Regional Models (A Contribution to Spatial Econometric Specification Analysis). *Journal of Regional Science* 26: 1-17.
- Bronars S.G., Jansen D.W. (1987) The Geographic Distribution of Unemployment Rates in the U.S. *Journal of Econometrics* 36: 251-279.
- Burridge P. (1981) Testing for a Common Factor in a Spatial Autoregression Model. *Environment and Planning A* 13: 795-800.
- Cliff A.D., Haggett P., Ord J.K., Bassett K.A., Davies R.B. (1975) *Elements of Spatial Structure*. Cambridge University Press, Cambridge.
- Cressie N.A.C. (1991) *Statistics for Spatial Data*. John Wiley, New York.
- Elhorst J.P. (1996) Regional Labour Market Research on Participation Rates. *Tijdschrift voor Economische en Sociale Geografie [Journal of Economic and Social Geography]* 87: 209-221.
- Engle R.F., Granger C.W.J. (1987) Cointegration and Error Correction: Representation, Estimation and Testing. *Econometrica* 55: 251-276.
- Fingleton B. (1999) Spurious Spatial Regression: Some Monte Carlo Results with a Spatial Unit Root and Spatial Cointegration. *Journal of Regional Science* 39: 1-19.
- Fischer M.M., Getis A. (1997, eds.) *Recent Developments in Spatial Analysis*. Springer, Berlin.
- Florax R., Folmer H. (1992) Specification and Estimation of Spatial Linear Regression Models. *Regional Science and Urban Economics* 22: 405-432.
- Florax R.J.G.M., Folmer H., Rey S. (1998) The Relevance of Hendry's Econometric Methodology: Experimental Simulation Results for Linear Spatial Process Models.

- Paper presented at the 38th European Regional Science Association Congress, Vienna, Austria.
- Gilbert C.L. (1986) Professor Hendry's Econometric Methodology. *Oxford Bulletin of Economics and Statistics* 48: 283-307.
- Greene W.H. (1993) *Econometric Analysis*, 2nd edition. Macmillan, New York.
- Griffith D.A. (1988) *Advanced Spatial Statistics*. Kluwer, Dordrecht.
- Griffith D.A., Amrhein C.G., Huriot J.-M. (eds.) *Econometric Advances in Spatial Modelling and Methodology: Essays in Honour of Jean Paelinck*. Kluwer, Dordrecht.
- Griffiths W.E., Hill R.C., Judge G.G. (1993) *Learning and Practising Econometrics*. Wiley, New York.
- Haining R. (1990) *Spatial Data Analysis in the Social and Environmental Sciences*. Cambridge University Press, Cambridge.
- Hendry D.F. (1995) *Dynamic Econometrics*. Oxford University Press, Oxford.
- Hendry D.F., Mizon G.E. (1978) Serial Correlation as a Convenient Simplification, not a Nuisance: A Comment on a Study of the Demand for Money by the Bank of England. *Economic Journal* 88: 549-563.
- Hsiao C. (1986) *Analysis of Panel Data*. Cambridge University Press, Cambridge.
- Kelejian H.H., Robinson D.P. (1995) Spatial Correlation; A Suggested Alternative to the Autoregressive Model. In: Anselin R., Florax R.J.G.M. (eds.) *New Directions in Spatial Econometrics*. Springer, Berlin.
- Kelejian H.H., Prucha I.R. (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: 99-121.
- Kelejian H.H., Prucha I.R. (1999) A Generalized Moments Estimator for the Autoregressive Parameter in a Spatial Model. *International Economic Review* 40: 509.
- Maddala G.S., Kim I.-M. (1998) *Unit Roots, Cointegration and Structural Change*. Cambridge University Press, Cambridge.
- Mizon G.E. (1995) A Simple Message for Autocorrelation Correctors: Don't. *Journal of Econometrics* 69: 267-288.
- Nerlove M. (1999) Properties of Alternative Estimators of Dynamic Panel Models: An Empirical Analysis of Cross-Country Data for the Study of Economic Growth. In: Hsiao C., Lahiri K., Lee L.-F., Pesaran M.H. (eds.) *Analysis of Panels and Limited Dependent Variable Models*. Cambridge University Press, Cambridge.
- Odland J. (1988) *Spatial Autocorrelation*. Sage Publications, Newbury Park, California.
- Ord K. (1975) Estimation Methods for Models of Spatial Interaction. *Journal of the American Statistical Association* 70: 120-126.
- Ord J.K. (1981) Towards a Theory of Spatial Statistics: A Comment. *Geographical Analysis* 13: 91-93.

- Pace R.K, Barry R., Clapp J., Rodriguez M. (1998) Spatiotemporal Autoregressive Models of Neighborhood Effects. *Journal of Real Estate Finance and Economics* 17: 15-33.
- Pfeifer P.E., Deutsch S.J. (1980) A Three-Stage Iterative Procedure for Space-Time Modeling. *Technometrics* 22: 35-47.
- Ramanathan R. (1995) *Introductory Econometrics*, 3th edition. Dryden Press, Forth Worth.
- Robinson G.M. (1998) *Methods and Techniques in Human Geography*. John Wiley, Chichester.
- Upton G.J.G., Fingleton B. (1985) *Spatial Data Analysis by Example*, volume 1. Wiley, Chichester.
- Wickens M.R., Breusch T.S. (1988) Dynamic Specification, the Long-Run and the Estimation of Transformed Regression Models. *The Economic Journal* 98: 189-205.
- Wrigley N., Bennett R.J. (1981, eds.) *Quantitative Geography*. London, Routledge & Kegan Paul.

**Table 1** The log-likelihood functions ( $1_n$  is a  $n \times 1$  vector of unit elements)

---

Static	$-\frac{1}{2}nT \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{t=1}^T e_t' e_t, \quad e_t = Y_t - \beta_0 - \beta_1 X_t$
Spat aut	$-\frac{1}{2}nT \ln(2\pi\sigma^2) + T \sum_{i=1}^n \log(1 - \delta\omega_i) - \frac{1}{2\sigma^2} \sum_{t=1}^T e_t' e_t, \quad e_t = Y_t - \beta_0 - \beta_1 X_t - \delta(WY_t - \beta_0 W1_n - \beta_1 WX_t)$
Ser aut	$-\frac{1}{2}nT \ln(2\pi\sigma^2) + \frac{1}{2}n \ln(1 - \tau^2) - \frac{1}{2\sigma^2} (1 - \tau^2) e_1' e_1 - \frac{1}{2\sigma^2} \sum_{t=2}^T e_t' e_t,$ $e_1 = Y_1 - \beta_0 - \beta_1 X_1$ $e_t = Y_t - \beta_0 - \beta_1 X_t - \tau(Y_{t-1} - \beta_0 - \beta_1 X_{t-1}) \quad t > 1$
Spat/Ser aut	$-\frac{1}{2}nT \ln(2\pi\sigma^2) + T \sum_{i=1}^n \log(1 - \delta\omega_i) + \frac{1}{2} \sum_{i=1}^n \ln\left(1 - \frac{\tau^2}{(1 - \delta\omega_i)^2}\right) - \frac{1}{2\sigma^2} e_1' (I - \tau^2 (I - \delta W))' (I - \delta W) e_1 +$ $-\frac{1}{2\sigma^2} \sum_{t=2}^T e_t' e_t, \quad e_1 = Y_1 - \beta_0 - \beta_1 X_1 - \delta(WY_1 - \beta_0 W1_n - \beta_1 WX_1)$ $e_t = Y_t - \beta_0 - \beta_1 X_t - \delta(WY_t - \beta_0 W1_n - \beta_1 WX_t) - \tau(Y_{t-1} - \beta_0 - \beta_1 X_{t-1}) \quad t > 1$

$$\text{Spat lag} \quad -\frac{1}{2}nT \ln(2\pi\sigma^2) + T \sum_{i=1}^n \log(1 - \delta\omega_i) - \frac{1}{2\sigma^2} \sum_{t=1}^T e_t' e_t, \quad e_t = Y_t - \delta W Y_t - \beta_0 - \beta_1 X_t - \beta_3 W X_t$$

$$\text{Ser lag} \quad -\frac{1}{2}nT \ln(2\pi\sigma^2) + \frac{1}{2}n \ln(1 - \tau^2) - \frac{1}{2\sigma^2} \frac{1+\tau}{1-\tau} e_1' e_1 - \frac{1}{2\sigma^2} \sum_{t=2}^T e_t' e_t,$$

$$e_1 = Y_1 - \tau Y_1 - \beta_0 - (\beta_1 + \beta_2) X_1$$

$$e_t = Y_t - \tau Y_{t-1} - \beta_0 - \beta_1 X_t - \beta_2 X_{t-1} \quad t > 1$$

$$\text{Spat lag/Ser aut} \quad -\frac{1}{2}nT \ln(2\pi\sigma^2) + T \sum_{i=1}^n \log(1 - \delta\omega_i) + \frac{1}{2}n \ln(1 - \tau^2) - \frac{1}{2\sigma^2} (1 - \tau^2) e_1' e_1 - \frac{1}{2\sigma^2} \sum_{t=2}^T e_t' e_t$$

$$e_1 = Y_1 - \delta W Y_1 - \beta_0 - \beta_1 X_1 - \beta_3 W X_1$$

$$e_t = Y_t - \delta W Y_t - \beta_0 - \beta_1 X_t - \beta_3 W X_t - \tau(Y_{t-1} - \delta W Y_{t-1} - \beta_0 - \beta_1 X_{t-1} - \beta_3 W X_{t-1}) \quad t > 1$$

$$\text{Ser lag/Spat aut} \quad -\frac{1}{2}nT \ln(2\pi\sigma^2) + T \sum_{i=1}^n \log(1 - \delta\omega_i) + \frac{1}{2}n \ln(1 - \tau^2) - \frac{1}{2\sigma^2} \frac{1+\tau}{1-\tau} e_1' e_1 - \frac{1}{2\sigma^2} \sum_{t=2}^T e_t' e_t$$

$$e_1 = Y_1 - \tau Y_1 - \beta_0 - (\beta_1 + \beta_2) X_1 - \delta(W Y_1 - \tau W Y_1 - \beta_0 W 1_n - (\beta_1 + \beta_2) W X_1)$$

$$e_t = Y_t - \tau Y_{t-1} - \beta_0 - \beta_1 X_t - \beta_2 X_{t-1} - \delta(W Y_t - \tau W Y_{t-1} - \beta_0 W 1_n - \beta_1 X_t - \beta_2 W X_{t-1}) \quad t > 1$$

$$\begin{aligned}
\text{Spat equilibrium cor.} \quad & -\frac{1}{2}nT \log(2\pi\sigma^2) + \frac{1}{2}n \log(1 - \tau^2) + (T-1) \sum_{i=1}^n \log(1 - \delta\omega_i) - \frac{1}{2\delta^2} \frac{1}{(1-\tau)^2} e_1' \\
& [(I - \delta W)'(I - \delta W) - (I - \delta W')(\tau I - \delta W)(I - \delta W)^{-1}(I - \delta W')^{-1}(\tau I - \delta W')(I - \delta W)]^{-1} e_1 + \\
& -\frac{1}{2\sigma^2} \sum_{t=2}^T e_t' e_t, \quad e_1 = Y_1 - \tau Y_1 - \beta_0 - (\beta_1 + \beta_2)X_1 \\
& e_t = Y_t - \tau Y_{t-1} - \delta W \Delta Y_t - \beta_0 - \beta_1 X_t - \beta_2 X_{t-1} - \beta_3 W \Delta X_t \quad t > 1 \\
\text{Spat/Ser lag combined} \quad & -\frac{1}{2}nT \log(2\pi\sigma^2) + T \sum_{i=1}^n \log(1 - \delta\omega_i) + \frac{1}{2} \sum_{i=1}^n \log\left(1 - \frac{\tau^2}{(1 - \delta\omega_i)^2}\right) - \frac{1}{2\sigma^2} e_1'(I - \delta W' - \tau I)^{-1} \\
& [(I - \delta W)'(I - \delta W) - \tau^2(I - \delta W')(I - \delta W)^{-1}(I - \delta W')^{-1}(I - \delta W)]^{-1}(I - \delta W - \tau I)^{-1} e_1 - \frac{1}{2\sigma^2} \sum_{t=2}^T e_t' e_t, \\
& e_1 = Y_1 - \tau Y_1 - \delta W Y_1 - (\beta_1 + \beta_2)X_1 - \beta_3 W X_1 \\
& e_t = Y_t - \tau Y_{t-1} - \delta W Y_t - \beta_1 X_t - \beta_2 X_{t-1} - \beta_3 W X_t \quad t > 1 \\
\text{Spat/Ser lag general} \quad & -\frac{1}{2}nT \log(2\pi\sigma^2) + T \sum_{i=1}^n \log(1 - \delta\omega_i) + \frac{1}{2} \sum_{i=1}^n \log\left(1 - \frac{(\tau + \eta\omega_i)^2}{(1 - \delta\omega_i)^2}\right) - \frac{1}{2\sigma^2} e_1'(I - \delta W' - \tau I - \eta W')^{-1} \\
& [(I - \delta W)'(I - \delta W) - (I - \delta W')(\tau I + \eta W)(I - \delta W)^{-1}(I - \delta W')^{-1}(\tau I + \eta W')(I - \delta W)]^{-1} \\
& (I - \delta W - \tau I - \eta W)^{-1} e_1 - \frac{1}{2\sigma^2} \sum_{t=2}^T e_t' e_t, \\
& e_1 = Y_1 - \tau Y_1 - (\delta + \eta)W Y_1 - (\beta_1 + \beta_2)X_1 - (\beta_3 - \beta_4)W X_1 \\
& e_t = Y_t - \tau Y_{t-1} - \delta W Y_t - \eta W Y_{t-1} - \beta_1 X_t - \beta_2 X_{t-1} - \beta_3 W X_t - \beta_4 W X_{t-1} \quad t > 1
\end{aligned}$$


---

**Table 2** Conditions on the models to ensure stationarity in time<sup>1</sup>

---

Static	-		
Spatial autocorrelation	-		
Serial autocorrelation	$ \tau  < 1$		
Spatial and serial autocorrelation	(i) $ \tau  < 1 - \delta\omega_{\max}$	if $\delta \geq 0$ ,	
	(ii) $ \tau  < 1 - \delta\omega_{\min}$	if $\delta < 0$	
Spatial lag	-		
Serial lag	$ \tau  < 1$		
Spatial lag and serial autocorrelation	$ \tau  < 1$		
Serial lag and spatial autocorrelation	$ \tau  < 1$		
Spatial equilibrium correction	$ \tau  < 1$		
Spatial and serial lag combined	(i) $ \tau  < 1 - \delta\omega_{\max}$	if $\delta \geq 0$ ,	
	(ii) $ \tau  < 1 - \delta\omega_{\min}$	if $\delta < 0$	
Spatial and serial general	(i) $\tau < 1 - (\delta + \eta)\omega_{\max}$	if $\delta + \eta \geq 0$ ,	
	(ii) $\tau < 1 - (\delta + \eta)\omega_{\min}$	if $\delta + \eta < 0$ ,	
	(iii) $-1 + (\delta - \eta)\omega_{\max} < \tau$	if $\delta - \eta \geq 0$ ,	
	(iv) $-1 + (\delta - \eta)\omega_{\min} < \tau$	if $\delta - \eta < 0$	

---

<sup>1</sup> In addition to these restrictions  $1/\omega_{\min} < \delta < 1/\omega_{\max}$  for all models except the static model, the serial autocorrelation model and the serial distributed lag model.



**Table 3** Estimation results of the labor force participation rate regressed on the unemployment rate (*T*-values between parentheses)

Model	France	Germany	the UK
10.	$\beta_0 = 0.742$ (102.29)	$\beta_0 = 0.745$ (141.45)	$\beta_0 = 0.820$ (100.95)
Static	$\beta_1 = -0.863$ (-12.72)	$\beta_1 = -1.202$ (-16.07)	$\beta_1 = -0.917$ (-12.42)
LM.8	118.5 Rej.	180.2 Rej.	43.4 Rej.
LM.9	43.5 Rej.	110.5 Rej.	32.3 Rej.
LogL	544.1	700.8	296.4
9.	$\beta_0 = 0.695$ (47.72)	$\beta_0 = 0.803$ (84.65)	$\beta_0 = 0.817$ (85.27)
Spat aut	$\beta_1 = -0.791$ (-11.89)	$\beta_1 = -0.908$ (-12.09)	$\beta_1 = -0.916$ (-12.56)
	$\delta = 0.892$ (27.27)	$\delta = 0.952$ (78.22)	$\delta = 0.622$ (8.03)
LM.7	93.5 Rej.	81.3 Rej.	51.2 Rej.
LogL	591.2	773.6	308.1
8.	$\beta_0 = 0.695$ (91.73)	$\beta_0 = 0.702$ (81.77)	$\beta_0 = 0.757$ (45.11)
Ser aut	$\beta_1 = -0.364$ (-9.87)	$\beta_1 = -0.460$ (-6.57)	$\beta_1 = -0.385$ (-3.72)
	$\tau = 0.902$ (41.85)	$\tau = 0.941$ (66.89)	$\tau = 0.940$ (36.14)
LM.7	26.9 Rej.	85.2 Rej.	19.2 Rej.
LogL	681.3	961.6	359.4
7.	$\beta_0 = 0.617$ (38.07)	$\beta_0 = 0.781$ (55.69)	$\beta_0 = 0.751$ (40.51)
Spat aut	$\beta_1 = -0.443$ (-11.22)	$\beta_1 = -0.529$ (-9.20)	$\beta_1 = -0.490$ (-5.09)
Ser aut	$\tau = 0.772$ (28.73)	$\tau = 0.803$ (35.24)	$\tau = 0.834$ (17.39)
	$\delta = 0.227$ (8.35)	$\delta = 0.196$ (8.73)	$\delta = 0.161$ (3.33)
LogL	697.7	991.0	366.6

6.	$\delta = -0.030$ (-1.02)	$\delta = -0.040$ (-2.02)	$\delta = 0.030$ (1.33)
Spat lag	$\beta_0 = 0.651$ (44.25)	$\beta_0 = 0.811$ (67.14)	$\beta_0 = 0.801$ (56.03)
	$\beta_1 = -0.666$ (-10.16)	$\beta_1 = -0.730$ (-8.09)	$\beta_1 = -0.851$ (-9.77)
	$\beta_3 = 0.937$ (5.91)	$\beta_3 = -1.180$ (-8.45)	$\beta_3 = -0.096$ (-0.76)
WALD.9	317.4 Rej.	817.1 Rej.	46.7 Rej.
LM.4	138.1 Rej.	133.2 Rej.	46.0 Rej.
LogL	596.5	796.4	314.0
5.	$\tau = 0.852$ (31.92)	$\tau = 0.915$ (54.16)	$\tau = 0.866$ (18.95)
Ser lag	$\beta_0 = 0.109$ (5.61)	$\beta_0 = 0.067$ (5.16)	$\beta_0 = 0.113$ (3.00)
	$\beta_1 = -0.304$ (-8.37)	$\beta_1 = -0.390$ (-5.13)	$\beta_1 = -0.343$ (-2.98)
	$\beta_2 = 0.177$ (4.22)	$\beta_2 = 0.236$ (2.94)	$\beta_2 = 0.198$ (1.60)
WALD.8	9.1 Rej.	13.3 Rej.	3.7 Not Rej.
LM.3	60.2 Rej.	97.4 Rej.	36.9 Rej.
LogL	690.8	975.0	365.2
4.	$\delta = 0.098$ (2.86)	$\delta = 0.572$ (9.01)	$\delta = 0.146$ (1.92)
Spat lag	$\beta_0 = 0.589$ (32.32)	$\beta_0 = 0.312$ (6.95)	$\beta_0 = 0.654$ (13.13)
Ser aut	$\beta_1 = -0.477$ (-13.39)	$\beta_1 = -0.444$ (-5.47)	$\beta_1 = -0.305$ (-1.71)
	$\beta_3 = 0.572$ (4.71)	$\beta_3 = 0.276$ (1.98)	$\beta_3 = -0.006$ (-0.03)
	$\tau = 0.843$ (32.84)	$\tau = 0.991$ (368.47)	$\tau = 0.959$ (38.58)
LogL	724.2	1023.9	367.1
3.	$\tau = 0.867$ (37.07)	$\tau = 0.905$ (57.51)	$\tau = 0.865$ (19.12)
Ser lag	$\beta_0 = 0.097$ (5.82)	$\beta_0 = 0.074$ (5.95)	$\beta_0 = 0.111$ (3.01)
Spat aut	$\beta_1 = -0.505$ (-19.06)	$\beta_1 = -0.477$ (-7.14)	$\beta_1 = -0.407$ (-3.05)
	$\beta_2 = 0.391$ (12.30)	$\beta_2 = 0.361$ (5.02)	$\beta_2 = 0.278$ (1.95)
	$\delta = 0.867$ (37.07)	$\delta = 0.784$ (15.76)	$\delta = 0.624$ (7.89)
LogL	730.9	1029.5	381.8

2.	$\tau = 0.886$ (38.95)	$\tau = 0.941$ (78.64)	$\tau = 0.911$ (22.96)
Spat	$\delta = 0.031$ (0.28)	$\delta = 0.428$ (6.05)	$\delta = 0.325$ (3.15)
eq.cor.	$\beta_0 = 0.084$ (5.03)	$\beta_0 = 0.046$ (4.80)	$\beta_0 = 0.075$ (2.35)
	$\beta_1 = -0.515$ (-15.78)	$\beta_1 = -0.524$ (-7.52)	$\beta_1 = -0.478$ (-2.73)
	$\beta_2 = 0.413$ (10.88)	$\beta_2 = 0.439$ (5.97)	$\beta_2 = 0.382$ (2.15)
	$\beta_3 = 0.735$ (5.92)	$\beta_3 = 0.346$ (2.59)	$\beta_3 = 0.272$ (1.21)
LogL	732.2	1031.4	381.1
1.	$\tau = 0.768$ (26.96)	$\tau = 0.851$ (38.84)	$\tau = 0.848$ (15.76)
Combined	$\delta = 0.002$ (0.16)	$\delta = 0.013$ (1.82)	$\delta = 0.019$ (1.61)
Spat lag	$\beta_0 = 0.151$ (7.94)	$\beta_0 = 0.119$ (6.49)	$\beta_0 = 0.123$ (2.91)
Ser lag	$\beta_1 = -0.347$ (-8.76)	$\beta_1 = -0.218$ (-2.04)	$\beta_1 = -0.284$ (-2.32)
	$\beta_2 = 0.193$ (4.66)	$\beta_2 = 0.115$ (1.17)	$\beta_2 = 0.147$ (1.16)
	$\beta_3 = 0.168$ (2.27)	$\beta_3 = -0.331$ (-6.07)	$\beta_3 = -0.105$ (-1.68)
WALD.7	14.6 Rej.	81.5 Rej.	9.4 Rej.
LogL	698.9	994.7	368.0
0.	$\tau = 0.836$ (35.32)	$\tau = 0.879$ (49.96)	$\tau = 0.865$ (19.87)
General	$\delta = 0.605$ (7.22)	$\delta = 0.682$ (10.70)	$\delta = 0.602$ (7.17)
Spat lag	$\eta = -0.600$ (-7.17)	$\eta = -0.685$ (-10.38)	$\eta = -0.594$ (-6.87)
Ser lag	$\beta_0 = 0.107$ (6.83)	$\beta_0 = 0.100$ (6.85)	$\beta_0 = 0.110$ (3.20)
	$\beta_1 = -0.519$ (-19.10)	$\beta_1 = -0.493$ (-6.97)	$\beta_1 = -0.460$ (-2.76)
	$\beta_2 = 0.410$ (13.30)	$\beta_2 = 0.375$ (4.92)	$\beta_2 = 0.335$ (1.97)
	$\beta_3 = 0.642$ (5.33)	$\beta_3 = 0.353$ (2.68)	$\beta_3 = 0.311$ (1.40)
	$\beta_4 = -0.553$ (-4.56)	$\beta_4 = -0.502$ (-3.60)	$\beta_4 = -0.345$ (-1.55)
WALD.1	102.4 Rej.	116.2 Rej.	51.5 Rej.
WALD.2	9.1 Not Rej-1%	13.5 Rej.	0.7 Not Rej.
WALD.3	7.7 Not Rej.	12.2 Rej.	4.8 Not Rej.
WALD.4	16.1 Rej.	39.1 Rej.	12.1 Rej.
LogL	733.3	1034.9	383.7

Hypotheses are tested at 5% significance, unless otherwise stated

**Figure 1** Restricted econometric models subsumed by the general serial and spatial autoregressive distributed lag model

