

# Space is More than Geography: Using Spatial Econometrics in the Study of Political Economy<sup>1</sup>

Nathaniel Beck  
Department of Politics  
New York University

&

Kristian Skrede Gleditsch and Kyle Beardsley  
Department of Political Science  
University of California, San Diego

Draft of September 29, 2004.

<sup>1</sup>Some of this material was presented in a paper by Beck and Gleditsch at the 2003 Joint Sessions of the European Consortium for Political Research, Edinburgh, Mar.-Apr. and the 2003 Political Methodology Meeting, Minneapolis, July. We thank Roger Bivand, Robert Franzese, Jude Hays, Torben Iversen, James Morrow Kelly Pace, Sara McLaughlin Mitchell, Randolph Siverson, and Chris Zorn for helpful discussions and comments, James LeSage for making available his MATLAB spatial econometrics library and answering our questions, and Michael Ward for impressing us on the importance of space. This research was supported by a grant from the National Science Foundation (SES-0214027).

## **Abstract**

While spatial econometrics is being used more frequently in political science, most applications are still based on geographic notions of distance. Here we argue that it is often more fruitful to consider political economy notions of distance, such as relative trade. We also argue that the spatially autoregressive model usually (but not always) should be preferred to the spatially lagged error model. Finally, we consider the role of spatial econometrics in analyzing time-series–cross-section data, and show that a plausible (and testable) assumption allows for the simple introduction of space (however defined) into such analyses. Examples of spatial analyses involving trade and democracy are presented.

# 1 Introduction: The promise of space

The goal of comparative research is invariably to test hypotheses about certain relationships between unit attributes and variation in outcomes of interests. However, as Galton (1889) pointed out over 100 years ago, trying to make inference directly from comparisons across units while assuming that observations are independent can yield misleading conclusions if variation in the outcome of interest stems from diffusion among units rather than functional relationships between the attributes compared. Although diffusion processes may well underlie many of the phenomena studied by political scientists most political economy research still relies on statistical models that assume that the individual observations are independent of one another. Spatial statistical models provide ways to test and accommodate various forms of dependence among observations.

Spatial econometric models have begun to make inroads into the study of political science, and, in particular, the study of international relations. This is evidenced, for example, by the various articles in the special issue of *Political Analysis* (10:3), as well as Gleditsch (2002a) and Gleditsch and Ward (2000). Spatial econometrics has its roots in the study of geography, so naturally these applications have typically used geographic notions of distance in their spatial model specification. However, there is no inherent reason for why spatial distance should need to be limited to geographic or Euclidean distance.

In this paper, we introduce some common spatial statistical models that can be used to estimate specified forms of dependence among observations in political science research. We first show how the basic framework can be used with alternative conceptions of space

other than geographical distance, using the relative importance of geographic neighbors and trading partners in analyses of the distribution of democracy as an example. We then show how spatial statistical models can be used to deal with dependence between dyadic observations where the same units enter into many related dyads. Finally, we extend our discussion to time-series—cross-section. Although it is difficult to estimate time-series—cross-section models with simultaneous spatial dependence, assuming that spatial influences operate with a temporal lag provides a practical alternative that can greatly facilitate estimation.

## 2 The spatial econometric model

In this section, we lay out the basic spatial econometric model. For simplicity, we assume that  $y$  is a continuous variable.<sup>1</sup> While much of our interest is in time-series—cross-section (TSCS) data, where units (nations or dyads) are observed annually over a long time period, we begin with a simple cross-sectional setup to avoid adding the complications of time. Since the standard model is well described elsewhere (Anselin, 1988), our overview will be relatively brief and omit many of the technical details.

As is done in all spatial econometric work, we assume that the structure of dependence between observations is known by the researcher and *is not estimated*. This structure of dependence is given by what is known as the “connectivity matrix,” which specifies the degree of connectivity between any two observations. The assumption that these connectivities

---

<sup>1</sup>If  $y$  is discrete, the spatial setup is much more difficult and we do not pursue it here. See Ward and Gleditsch (2002) for one approach to this problem.

are known *a priori* is both a strong assumption and critical for the methods of spatial econometrics to work. Of course it is no stronger than the typical implicit assumption that all connectivities are zero, that is, all observations are spatially independent. As we shall see, political science can give us insight as to the nature of the connectivities.

We denote the the connectivity matrix by  $\mathbf{W}$ , where a typical element,  $w_{ij}$ , has a value greater than 0 if the observations  $i$  and  $j$  are connected. By convention, units are not considered to be connected to themselves, so any diagonal entry  $w_{ii} = 0$ . In our discussion, we will assume that we are dealing with situations where no observation is an isolate without any ties to other observations (i.e., we rule out that  $\sum \mathbf{w}_i = 0$  for any observation  $i$ ). The connectivity matrix is standardized so that each row vector  $\mathbf{w}_i$  sums to unity. As a consequence, it is not critical to worry about the units to measure connectivity, since  $\mathbf{W}$  is invariant to affine transformations.

Given some specification of connectivity or dependence we can define the “spatial lag ,” or the average value of  $y$  in a state  $i$ 's connected entities, by

$$y_i^{\bar{w}} = \mathbf{w}'_i \mathbf{y}. \tag{1}$$

Row-standardization ensures that the spatial lag  $y_i^{\bar{w}}$  will have the same metric as the original  $y$ .<sup>2</sup>

In a geographic connectivity matrix the notion of observations being “nearby” one another is determined purely by physical distance. Geographically oriented spatial econometricians

---

<sup>2</sup>The are other approaches beyond row standardization that are sensible for particular applications (Cornes and Sandler, 1996, 492–4), but row standardization is far and away most commonly used.

use one of two connectivity measures: either a binary measure of contiguity (or perhaps a binary measure of being closer than a certain specified threshold) or a continuous measure of distance between two units (based on distances between some reference point such as the capital city or the minimum distance between the two closest points on the countries outer' boundaries). However, although “nearby” is usually taken to mean geographical closeness, there is no reason why we cannot use any notion of nearness that makes theoretical sense, so long as this is specified by the analyst and so long as it does not violate any of the assumptions about the connectivity matrix stated above. In the following sections, we will return to this issue and show how spatial methods, with suitably defined non-geographic notions of “distance,” can be used in several applications of interest to political economy scholars.

We start with the standard linear regression model and show how spatial dependence will give rise to violations of the classical regression assumptions. Let  $y_i$  represent some dependent variable of interest, and, as usual, we assume it is a linear function of covariates,  $\mathbf{x}_i$ , and some unmeasured variables, “the error,”  $\varepsilon_i$ , so that

$$y_i = \mathbf{x}_i\beta + \varepsilon_i. \tag{2}$$

Linearity here is purely for ease of notation, and any non-linear extensions available in typical econometric models are also available in the spatial framework.

The most critical assumption, other than that the specification of the model is correct, is that the covariates are independent of the error process. We also begin with that assumption, but then weaken the assumption that the error process (and hence the  $y_i|\mathbf{x}_i$ ) are independent

across observations. The basic spatial insight is that “errors”  $\epsilon_i$  (best thought of as omitted or unmeasured variables) in unit  $i$  are related to the “errors”  $\epsilon_j$  in nearby units  $J$ , where  $j \neq i$ .

## Spatially lagged errors

Letting  $\mathbf{w}^i$  be a vector for how “close” the other  $J$  observations ( $j \neq i$ ) are to unit  $i$  (so that  $w_j^i = W_{i,j}$  and hence  $w_i^i = 0$ ) and letting  $\boldsymbol{\epsilon}$  be the vector of all errors, we get the “spatially lagged error” (or what Anselin 1988 calls the “spatial error”) model:

$$y_i = \mathbf{x}_i\beta + \epsilon_i + \lambda\mathbf{w}_i'\boldsymbol{\epsilon}. \quad (3)$$

If  $\lambda = 0$ , this reduces to the standard non-spatial linear regression model.  $\lambda$  is the only spatial parameter in this setup that is actually estimated.

If  $\lambda \neq 0$ , OLS is still consistent, but the reported standard errors will be wrong and the estimated  $\hat{\beta}$  will be inefficient. This can be fixed by typical GLS reasoning, although complications require a full maximum likelihood estimation.<sup>3</sup>

---

<sup>3</sup>The log-likelihood for the spatial model is complicated since it involves the log of the determinant  $|I - \lambda\mathbf{W}|$ , which is an  $n$ th order polynomial that can be very time consuming to evaluate. Ord (1975) showed that this determinant could be written as a function of the product of the eigenvalues  $\omega_i$  of the connectivity matrix  $\mathbf{W}$ ,  $|I - \lambda\mathbf{W}| = \prod (1 - \lambda\omega_i)$ . The eigenvalues  $\omega_i$  can be determined prior to optimization, which allows writing the likelihood for the model in a form that can be easily be estimated. For further econometric details on the spatially lagged error model, the reader should consult Anselin’s (1988) text. Taking advantage of the sparseness of connectivity matrices, Pace and Barry (1997) have developed a more efficient computational procedure for approximating the log determinant that is considerably faster than the

The spatially lagged error model corresponds to the model in time series analysis where the errors show some temporal correlation process. The analogy to the time series serially correlated error model is useful. This analogy tells us that the only way that observations are interdependent is through unmeasured variables that are correlated, in this case across space.

The spatially lagged error model is odd (at least in many applications), in that space matters in the “error process” but not in the substantive portion of the model. Moreover, if we add a new variable to the model, so that we move it from the “error” to the substantive portion of the model, the spatially lagged error model assumes that this variable no longer has a spatial impact of nearby observations. This assumption seems to us hard to defend in many applications, although we believe that this model may be appropriate for interconnectedness among observations in certain applications.

To see why we consider the spatially lagged error model odd, at least for political economy models, let  $y$  be some economic output, say growth. Growth in one country probably depends on growth in nearby countries (where we discuss the meaning of “nearby” below). In the spatially lagged error model, the only way that nearby countries have an impact is through the interrelated error terms; the error for some country is related to all the errors for all the other nearby countries. But remember that the “errors” are just the variables that we either chose not to measure, or did not measure. In particular, they are errors from the perspective of the analyst, not the perspective of policy makers in the country. Thus if Germany grew

---

Ord approach and helpful for working with large data sets.



more quickly due to some variable not included in the specification, that growth would affect all other countries. But if Germany grew more quickly because it had a left government, and if that variable were included in the specification, then this extra German growth would have no impact on the growth in other countries. Since the growth of Germany's trading partners depends on overall German growth, and not just the portion we as analysts treat as the "error" term, we find that for most political economy applications (and probably most applications), that the spatially lagged error model is not appropriate. This does not mean that it is never appropriate, or that one should not think about what is the appropriate model, but that the spatially lagged error model is not the one that would come to mind first.

## **Spatial autoregressive model**

The spatially lagged error model corresponds to the time-series serially correlated errors model. The spatial autoregressive model (also known as the "spatial lag" model) corresponds to the time-series lagged dependent variable model. In this model, the dependent variable is affected by the values of the dependent variable in nearby units, with "nearby" suitably defined. It differs from the spatially lagged errors model in that both the error term and the covariates in nearby units impact the current unit (since it is the current values of the other  $y$ 's, suitably weighted, that impacts the dependent variable). Thus, for example, let the dependent variable be the level of democracy in a country. It is likely that this is partly a function of democracy in nearby countries, rather than just being related to common

unmeasured variables in nearby countries.

Again, we start with the simple cross-sectional model. Using the same notation as above, and letting  $\mathbf{y}$  be the vector of values for  $y$ , the spatial lag model has the form

$$y_i = \mathbf{x}_i\beta + \kappa\mathbf{w}'_i\mathbf{y} + \varepsilon_i. \quad (4)$$

Unlike the spatially lagged error model, OLS is inconsistent for the spatial autoregressive model. The problem lies in that the expected value of the product of the spatial lag term and the error term is non-zero.<sup>4</sup> This spatial autoregressive model is difficult to estimate, but it can be done by complicated maximum likelihood.<sup>5</sup>

The traditional spatial autoregressive model presumes that there is only one form of dependence, which can be represented in a single connectivity matrix  $\mathbf{W}$ . As above, we are primarily interested in the potential for using the model with weighting vectors that are more politically inspired than the distance or proximity weights used by geographers. However, in many cases there may be several possible networks or forms of dependence. It is possible to generalize the spatial autoregressive model to two distinct connectivity matrices  $\mathbf{W}_1$  and  $\mathbf{W}_2$  and estimate separate parameters  $\kappa_1$  and  $\kappa_2$  for the relative impact of each, by

$$y_i = \mathbf{x}_i\beta + \kappa_1\mathbf{w}'_{i1}\mathbf{y} + \kappa_2\mathbf{w}'_{i2}\mathbf{y} + \varepsilon_i. \quad (5)$$

Lacombe (2004), for example, uses such a model to estimate parameters distinguishing

---

<sup>4</sup>The consistency of OLS depends on  $E[(\mathbf{W}y_t)'\varepsilon] = 0$ . However,

$$\text{plim}N^{-1}[(\mathbf{W}y_t)'\varepsilon] = \text{plim}N^{-1}\varepsilon'\mathbf{W}(\mathbf{I} - \kappa\mathbf{W})^{-1}\varepsilon.$$

Unless  $\kappa = 0$ , this probability limit will not equal zero.

<sup>5</sup>Again, we refer to Anselin (1988) and Pace and Barry (1997) for details on estimation of the model.

within-state unit and between-state unit effects of welfare programs on female labor force participation. The expanded spatial autoregressive model is even more complicated to estimate than the standard spatial autoregressive model, but the same ML estimator can be generalized to this case, provided the two matrices are sufficiently different and do not contain entirely overlapping information.<sup>6</sup>

### 3 Beyond Euclid: non-geographic notions of space

Although most applications of spatial statistics in the social sciences have use geographic distance, there is nothing in the basic framework that requires that the connectivity matrix  $\mathbf{W}$  must be based on geographic distance *per se*.

In political science, we usually have interesting networks or linkages defined by political or social phenomena. For example, one might envision that observations are influenced not by geographically proximate units, but rather by historical shared ties (such as language or colonial history) or high levels of interactions. Deutsch and Isard (1961) make this point in an early paper that precedes the spatial econometric models discussed in this paper. Sociologists sometimes speak of social distance or Blau-space where distances between individuals are based on coordinates that are not geographical locations.

Although researchers have often suggested that dependence may be due to distances that are not necessarily geographical, we have found very few actual examples of connectivity

---

<sup>6</sup>Whereas the standard spatial autoregressive MLE requires calculating the Jacobian determinant  $|I - \kappa\mathbf{W}|$ , in the generalized spatial autoregressive models we must consider  $|I - \kappa_1\mathbf{W}_1 - \kappa_2\mathbf{W}_2|$ .

matrices based upon things other than Euclidean distance. Dow, Burton, White and Reitz (1984) consider dependence from geographical distance as well as language similarity in an application to the diffusion of gambling. They estimate separate models for each matrix, and the relative influence on one type of space when the other is taken into account is not examined empirically. Conley (1999) uses a measure of the transportation costs for physical capital between countries in a study of economic growth. In a study of environmental degradation, Lofdahl (2002) considers a measure of trade with other states relative to the size of GDP to estimate the environmental impacts of globalization or economic openness. Despite the prominence of the concept of social space and the clear analogies between graph theory and the spatial statistical models, the existing literature has paid very little attention to the potential for applications of spatial statistics to social distances. Anselin's monograph, for example, does not contain a single example of non-geographical distance metrics.

In this paper we use a variety of non-geographic measures (democracy, trade and common dyadic membership) in various spatial analyses. We focus our attention on the spatial autoregressive model here since this model seems more useful to us in this particular application. However, the same alternative measures of space could in principle be used with the spatially lagged error model.

## 4 The Requisites of Democracy

Following Lipset's 1960 social requisites hypothesis, an extensive literature has examined how social and economic attributes influence the likelihood that countries will be demo-

cratic. However, there are many reasons to suspect that the level of democracy in one country could be influenced by the level of democracy in other states.<sup>7</sup> Previous analyses have considered diffusion in the context of relations between countries that are geographic neighbors. However, there is no particular reason why connections between states must be limited to geographic distance, and much causal evidence suggesting that the nearest or most relevant actors are not necessarily the geographic neighbors. Moreover, we have no theory which specifies *a priori* that one particular distance measure is correct. So here we estimate a model with two measures of distance, and allow the data to suggest the relative importance of the two. Our measure of democracy is taken from the Polity 4 data.<sup>8</sup> We use the full 21 point institutionalized democracy scale suggested by Jagers and Gurr (1995).

## Defining connectivity and space

We use two plausible definitions of connectivity between states. Clearly, diffusion between countries may be likely to occur between nearby countries in a geographical sense. This is the basis for our first connectivity criterion, using the geographical distance between states. We use the minimum distance data from Gleditsch and Ward (2001) to define countries as connected if they are within 500 km of one another. This yields a binary connectivity matrix where each entry  $w_{ij}$  is 1 if state  $i$  and state  $j$  are within 500 km from each other. Each neighboring country is given equal weight in the row for country  $i$ . Here, and elsewhere in

---

<sup>7</sup>See Gleditsch (2002a) for a more extended discussion.

<sup>8</sup>We use the modified version of Polity 4 data with estimates for countries not included in the Polity data based on the Freedom House data, available at <http://weber.ucsd.edu/~kgledits/Polity.html>.

this section, we normalize the matrix so that each row sums to 1.

Our second specification of connectivity is based on the volume of trade flows, taken from the Expanded GDP and Trade data (version 4.1) described in Gleditsch (2002b). A country is considered connected to all other countries that it has some trade with. However, countries tend to be more dependent or influenced by their major trading partners, where the bilateral trade flows are large relative to a country's total trade. The trade connectivity matrix differs from the previous distance matrix in two notable ways. First, whereas the distance matrix assign equal weights to any geographical neighbor, the trade matrix consists of weights where the importance of another state  $j$  to state  $i$  is given by the volume of the dyadic trade flow between  $i$  and  $j$  as a proportion of country's  $i$  total trade. This weights large trading partners much more heavily than smaller trading partners. Moreover, in the distance matrix, any neighbor of  $i$  must always have  $i$  as a non-trivial neighbor. In the trade matrix, however, it will often be the case that one country, say El Salvador, has another country, say the United States, as its major trading partner, yet in turn is a relatively small and trivial trading partner to the other country.<sup>9</sup>

Clearly, the bases for each of the specifications of connectivity matrices are quite different, and the spatial lag measures for democracy based on geography and trade look quite different from one another. Although trading trading patterns are also geographically clustered and the two final spatial lag vectors are positively correlated with one another (0.616),

---

<sup>9</sup>More generally, social distances violate the so-called triangle inequality defining metric spaces, since it is no longer generally the case that the sum of the distances from A to B and from B to C must be greater or equal to the distance from A to C.

the matrices based on the two definitions of connectivity are not so similar as to be indistinguishable from one another. Whereas the spatial lag of democracy defined over distance has a bimodal density function, much like the density of the democracy variable itself, the spatial lag of democracy defined over the trade flow matrix has a single peak, with a mean much higher than the median (or even mean) of the democracy variable. One interesting feature of the trade matrix is that some more-open developing countries have the bulk of their trade with large, wealthier countries, which more often tend to be democratic. As a result, these developing countries that have greater openness and trade will tend to have a higher “spatial lag” or average democracy score among its trading partners. This suggests that trade may identify a very different set of pull factors than geographical proximity or influence from neighbors.

## **Cross-sectional analyses**

To explore the importance of spatial linkages between observations in the distribution, we start by a cross sectional analysis of data for the year 1998. We work with a very simple social requisites model, which explains democracy by one variable, the (natural) log of GDP per capita (in constant 1996 US dollars). The results are shown in Table 1. As can be seen, the OLS results in the second column suggest a strong positive relationship between GDP per capita and level of democracy.

However, as previously discussed, the OLS coefficient estimate of the log of GDP per capita may well be biased if values on the democracy cluster spatially beyond what can

Table 1: Democracy and Social Requisites, 1998

Variable	OLS,1998	Spatial autoregressive estimates		
Constant	-20.84 (3.66)	-13.78 (3.49)	-20.41 (3.51)	-14.34 (3.20)
Ln(GDPPC)	2.78 (0.43)	1.81 (0.42)	2.33 (0.42)	1.27 (0.38)
$\kappa_1$ (distance)		0.47 (0.09)		0.39 (0.13)
$\kappa_2$ (trade)			0.50 (0.14)	0.80 (0.11)
N	170	170	170	170

be accounted for by GDP per capita. The third column of Table 1 contains estimates of a spatially autoregressive model with a lagged term defined by geographic distances. The first thing to note is that the results in the second column of Table 1 show clear evidence of spatial clustering. The estimate of  $\kappa$  indicates positive spatial clustering in democracy levels across neighboring countries, and the clustering is statistically significant. Moreover, these results also indicate that the coefficient estimate and standard error of the OLS model assuming independent observations may display substantial upward bias. More specifically, we find that the coefficient estimate for the natural log of GDP per capita is reduced to about 60% of its original size.

The results for the trade connectivities matrix in the third column of Table 1 look substantially different from the geographical connectivity matrix. The estimated  $\kappa$  indicates significant clustering, much the same as “neighbors” defined by geographical distance, or that countries tend to see a push towards the average level of democracy among their trading partners. Stated differently, the results suggest that countries that trade more with



democracies are more likely to be democratic, over and beyond what one would expect based on their wealth or income. However, the coefficient estimate for the natural log of GDP per capita changes much less relative to the OLS coefficient estimate than was the case for the geographical connectivity specifications. This implies that the clustering of democracies with respect to trading partners does not affect our estimate of the effect of GDP per capita to the same extent as the clustering with respect to geographic proximity.

So far we have assumed a single spatial dimensions in each regression. To see the relative contribution of each definition of “space” once others are taken into account, we now turn to cross-sectional analyses where we estimate different parameters for each of the different “spatial” metrics. The right column of Table 1 shows the results from a spatial autoregressive model with *both* distance and trade connectivities.

The first thing to note is that the estimated  $\kappa$  coefficients are all positive, and statistically significant, suggesting that distance and trade indicate substantively different influences on level of democracy. However, whereas both geographical clustering and clustering over trading partners seem to similarly matter when estimated separately, the estimated impact of spatial clustering over distance is considerably reduced once clustering over trading partners is taken into account; the estimated  $\kappa_1$  for distance is only about half the size of the estimated  $\kappa_2$  for trade. These results suggest that both neighboring democracies and democratic trade partners can pull countries toward democracy, but that trade matters more. The estimated effect of income is cut in half once these two forms of spatial dependence are taken into account. Thus it is important to both take spatial dependencies into account, and by using

multiple forms of spatial dependencies we can make interesting inferences about the relative importance of various types of spatial dependencies. Our next application is rather different, showing that spatial methods can solve some technical problems related to interdependent dyads. This is one of the few cases where we feel that the spatially lagged error model is appropriate.

## 5 Dyadic Dependence

Much IR analysis is based on the dyad-year design. Most of the attention to the problem of interdependence among dyads has centered on whether the successive annual observations on the same dyad are independent (Beck and Katz, 1996). However, time-series–cross-section analysts have also worried about whether observations on different units at the same time point are independent (Beck and Katz, 1995). Here we consider one type of interdependence that arises in dyadic data: two dyads that contain a common member are unlikely to be independent. Perhaps even more seriously, data sets often contain the directed dyads AB and BA; these two dyads are particularly unlikely to be independent.<sup>10</sup> We use some spatial econometrics to help with this issue. But rather than use a geographic notion of closeness, we posit that two dyads are close if they share a common member, and are especially close if they are the reverse of each other.

---

<sup>10</sup>The only research on this that we know of is by Mansfield and Bronson (1997) which adjoins to each model two dummy variables, one to represent each state in the dyad. As discussed in greater detail in Beck and Katz (2001), these fixed effects are not ideal in IR data.

In principle, we could estimate a standard spatially lagged error model with a weighting vector that considers another dyad near a given dyad AB if they share a common member (i.e., either A or B). This has the disadvantage, however, of not distinguishing between the dyads that include either A or B (usually, this will be a large number) and the reverse dyad BA, composed of the same two members but looking at the reverse flow from B to A, which is likely to be particularly influential.

One possibility would be to assign the weight for the reverse dyad some value  $v$  and estimate the relative weight for the reverse dyad relative to other connected dyads. In a previous paper, we estimated the value of the weight  $v$  for the reverse dyad by a search based on overall model fit (Beck and Gleditsch, 2003). Another alternative which we pursue here is to estimate a spatially lagged error model with two matrices, where the first  $\mathbf{W}_1$  includes connectivities to all common member dyads, save for the reverse dyad, and the second matrix  $\mathbf{W}_2$  includes connectivities to all reverse dyads. Thus it is critical to be able to estimate spatial models with two distinct connectivity matrices, with the particular weight on each to be estimated.

## **Example: The politics of dyadic trade**

Because  $y$  must be continuous, we chose a data set dealing with the political determinants of trade (viewed as exports directed from A to B). We estimate a standard political economy model of trade flows, similar to one common in the IPE literature.<sup>11</sup> The non-political deter-

---

<sup>11</sup>It is very similar to the basic model of Morrow, Siverson and Tabares (1999), except they only look at major powers.

minants of trade are taken to be as in a standard gravity model, where exports are regressed on the GDP and population of both exporter and importer and the distance between the two (all values are logged).<sup>12</sup> As for the political variables, we look at whether the dyad is in a Militarized Interstate Dispute (MID), the similarity of their alliance portfolios (S), and the lower of the Polity democracy scores of the two states in the dyad (DEM).<sup>13</sup> To avoid the added complication of time we have here limited ourselves to a cross-section, using only data from 1998.<sup>14</sup> We include all independent states in the international systems, using data from Gleditsch (2002b); other data sources are standard.

While one can analyze very large models using sparse matrix routines, the number of non-zero entries in a connectivity matrix for dyads with common members increases very rapidly in the number of observations  $N$ . For a given  $N$ , a particular state  $A$  will be a member of  $N(N - 1)$  other directed dyads. Since  $B$  likewise enters into  $N(N - 1)$  dyads, it follows that for a given dyad  $AB$  there will be  $2[N(N - 1) - 1]$  connected dyads, not including the reverse dyad  $BA$ . For a cross section of  $N = 180$ , for example, we would then get a

---

<sup>12</sup>All values were logged because of the nature of the gravity model of trade used by Morrow, et al. This leads to non-standard coding for dummy variables and such, but since our interest is only on the changes in coefficients and standard errors, we need not go into this here.

<sup>13</sup>We take the natural log of the minimum democracy value, after adding 11, so that the variable before transformation has possible values ranging from 1 to 21. We also log the S score, which has values between 0 and 1 in our data (not including 0). Finally, we take the log of the natural log of the MID variable, which has values of 1 (no war) or 2 (war).

<sup>14</sup>The method extends totally straightforwardly to time-series-cross-section data, but that introduces other issues discussed in the next section. These other issues overwhelm the discussion of dyadic dependence we wish to focus on here.

connectivity matrix with  $2[180(180 - 1) - 1] = 64,438$  entries for each row  $i$ , which in turn implies a matrix with almost 2.1 billion non-zero weights for the  $180(180 - 1) = 32,220$  directed dyads.

To avoid excessive computational demands we analyze two smaller regional sub-samples. We consider two empirical examples that we believe will illustrate common situations in applied work: 1) a European sample including only “high-quality” observations that we are relatively confident in, and 2) an African sample including many estimates of more questionable nature. Whereas most of the European data are “observed” or reported in standard sources, an inspection of the African data reveals that many of the dyadic trade flows for the African states are based on imputations. Since some of the imputations are based on assuming that flows are similar to the reverse flows or that there is no trade between countries (see Gleditsch 2002b for a discussion of the problem of missing data in the IMF directions of trade data and possible imputation methods), they may tend to exaggerate the similarity between exports from A to B and the exports from B to A. Dropping the most contentious estimates, however, leaves very few remaining observations for African dyads. While we believe that such strategies for addressing problems of missing data are likely to be better than simply applying listwise deletion and analyzing the remaining sample as if the observations were missing at random, many imputation techniques are known to generate serial correlation among observations (and by implication the errors). Imputation methods based on the similarity of reverse dyads can exacerbate the problem of serially correlated errors. However, the spatially lagged error model provides a relatively simple way to address

this potential problem.

The OLS and spatially lagged error model results are displayed in Tables 2 and 3. The parameter  $\lambda_1$  refers to the estimate for the reverse dyad (i.e., dyad BA for a dyad AB) while  $\lambda_2$  denotes the impact of all the other common member dyads.

Table 2: Directed Export Flows, Europe, 1998

Variable	OLS	Spatially lagged error estimates
Constant	-32.70 (0.67)	-34.74 (1.55)
Ln democracy	0.38 (0.06)	0.48 (0.15)
Ln population A	0.86 (0.02)	0.91 (0.04)
Ln population B	0.75 (0.02)	0.79 (0.04)
Ln GDP A	1.54 (0.04)	1.59 (0.09)
Ln GDP B	1.01 (0.04)	1.06 (0.09)
Ln S	0.33 (0.05)	0.35 (0.05)
Ln distance	-0.34 (0.01)	-0.34 (0.02)
Ln MID	-1.94 (0.27)	-1.33 (0.37)
$\lambda_1$		0.29 (0.01)
$\lambda_2$		0.94 (0.01)
N	1500	1500

As can be seen from Tables 2 and 3, the coefficient estimates for  $\lambda_1$  and  $\lambda_2$  are both positive and significant, indicating that there is considerable similarity in the trade flows of

Table 3: Directed Export Flows, Africa, 1998

<b>Variable</b>	OLS	Spatially lagged error estimates
Constant	-7.41 (0.33)	-7.51 (0.98)
Ln democracy	-0.04 (0.04)	0.01 (0.08)
Ln population A	0.26 (0.01)	0.26 (0.03)
Ln population B	0.23 (0.01)	0.23 (0.03)
LN GDP A	0.38 (0.02)	0.38 (0.05)
Ln GDP B	0.31 (0.02)	0.32 (0.05)
Ln similarity	3.41 (0.40)	3.46 (0.65)
Ln distance	-0.17 (0.01)	-0.17 (0.01)
Ln MID	-0.71 (0.18)	-0.36 (0.23)
$\lambda_1$		0.43 (0.01)
$\lambda_2$		0.73 (0.01)
N	2500	2500

a dyad AB and its associated common member dyads, and that this is not fully accounted for by the other covariates in the model. Moreover, the estimates from the spatially lagged error model differ notably from the OLS results that assume that the observations are independent of one another. In particular, we find quite large differences in the estimates for the political variables. The coefficient estimate for the democracy coefficient in the European sample increases by about 25% once we take into account the spatially lagged error struc-

ture. Likewise, the estimate for the impact of MIDs in the spatially lagged error model is reduced by over 30% from its original size in the OLS results, and would not be deemed to be statistically significant in a standard one tailed test. This may reflect that MIDs reduce not only bilateral trade flows but trade with other parties in general, so that countries in MIDs will have generally lower trade with other dyads, not just their antagonist. Clearly, many inferences about how much political factors influence trade flows seem to depend considerably on whether we are willing to assume that all observations are independent of one another, at least in the European context. Our results suggests that this assumption seems rather suspect in the analysis of trade flows.

In the case of Africa, the coefficient estimates show generally smaller differences between the OLS estimates assuming independent observations and the spatially lagged error model. However, we note that the coefficient estimates for MIDs is reduced to about 50% of its original size in the spatially lagged error model. Likewise, the proportional size of  $\lambda_1$  relative to  $\lambda_2$  is much higher in the African sample than the European sample.

We also note that although the estimated standard error of the regression is smaller for the spatially lagged error model than the OLS estimate, the standard error estimates for the coefficient estimates are generally much larger for the spatially lagged error model, in some cases over twice the size of the OLS standard errors. In particular, in the African sample where the coefficient estimates do not differ so much, the more realistic spatially lagged error model SEs are generally more than twice the size of the OLS SEs. The large difference in the SEs likely reflect how imputations based on presuming that reverse dyads are similar are



likely to induce spatially correlated errors. The spatial statistical model, however, suggests a simple way to deal with this problem in applied settings.

Of course, this does not address the question of dynamics. In a time-series–cross-section model we would expect the spatial association to matter less, once the autoregressive component in the dependent variable is taken into account. We turn to this issue in the next section.

## 6 Time-Series–Cross-Section Models

The spatial autoregressive model can be generalized for time-series–cross-section (TSCS) data as

$$y_{i,t} = \mathbf{x}_{i,t}\beta + \kappa\mathbf{w}'_i\mathbf{y}_{i,t} + \varepsilon_{i,t} \quad (6)$$

where the notation implies that the weighting vector is time invariant (although it is easy to allow it to vary over time, so long as it is known *ex ante*), with a similar obvious generalization for our less preferred spatially lagged error model.

Spatial econometricians have primarily analyzed single cross-sections. While there is recent unpublished work by Kelejian and his students, this is for panel, not TSCS data.<sup>15</sup> There is also recent work by Franzese and Hayes (2004), though their interests are orthogonal to ours.

---

<sup>15</sup>We will not go into the differences between these two types of data here, referring the reader to Beck (2001), other than to note that the Kelejian papers deal primarily with how to combine spatial methods with random effects; random effects models are of little interest to students of political economy.

Before beginning our analysis, it should be noted that the spatially lagged error model would have solved many of the problems of the Kmenta-Parks (Kmenta, 1986; Parks, 1967) method as discussed in Beck and Katz (1995). The basic problem with that method is that it allows for contemporaneous errors to be correlated with an arbitrary error structure; spatial econometricians would have specified that with a connectivity matrix. Thus spatial econometricians would have done GLS with only one extra parameter, which works well; Parks-Kmenta, by not specifying the connectivity matrix, were forced to estimate an inordinate number of extra parameters. While most TSCS analysts no longer use Kmenta-Parks, those who are considering this method should clearly prefer the spatial autoregressive model. We do not pursue this further here.

Since TSCS models normally show temporal dynamics, we can add a temporal lag of  $y$  to the model, yielding

$$y_{i,t} = \mathbf{x}_{i,t}\beta + \phi y_{i,t-1} + \kappa \mathbf{w}'_i \mathbf{y}_{i,t} + \varepsilon_{i,t}. \quad (7)$$

Without the spatial term  $\kappa \mathbf{w}'_i \mathbf{y}_{i,t}$ , this equation would be easy to estimate provided the error process shows no temporal correlations, so that the lagged  $y$  is independent of the error process. Equation 7 is often reasonable in practice, although of course it *must* be tested via a Lagrange multiplier test (Beck and Katz, 1996), to make sure that the errors are temporally independent. However, the presence of the lagged dependent variable makes the the Jacobian of the transformation in the ML estimator for the spatial autoregressive model very complex, and, so far as we know, no one has come up with a satisfactory estimate for this model.

TSCS data, however, allow an alternative specification of the spatial autoregressive model which is simple to estimate. Of course we choose specifications which are theoretically sound, not because they are easy to estimate, but if this specification is theoretically plausible, then the ease of estimation should not be sneered at.<sup>16</sup> Suppose that we continue to maintain that  $y_{i,t}$  is related to the neighboring  $y$ 's, but we believe that this impact occurs with a one-period lag. This is often at least as plausible as spatial lags having an instantaneous effect, though of course this plausibility varies by what is being modeled and the theory available to the researcher. For example, for political economy models, we would need to decide whether we expect it to be more likely that the growth of neighboring countries immediately affects growth or that it affects growth with a temporal lag.<sup>17</sup> The latter case yields the model

$$y_{i,t} = \mathbf{x}_{i,t}\beta + \phi y_{i,t-1} + \kappa \mathbf{w}'_i \mathbf{y}_{i,t-1} + \varepsilon_{i,t}. \quad (8)$$

If the errors are temporally independent, this model is easy to estimate via OLS.<sup>18</sup>

---

<sup>16</sup>Ease of estimation, that is being able to do OLS, allows the analyst to model many other features of the data. Committing to a complicated spatial model makes it almost impossible to solve many other problems, problems which may be more important. Thus ease of estimation is not simply to be preferred by lazy analysts.

<sup>17</sup>For annual data, this lag would have to be annual, so we only need compare instantaneous neighbor effects and those that take a year to set in. For data measured quarterly or monthly we have more choices, but such data are rare in the study of political economy.

<sup>18</sup>Note that, as with the non-spatial TSCS lagged dependent variable model, the assumption of temporally independent errors can be tested with a Lagrange multiplier test of the OLS residuals, so researchers can do more than just assume that the lagged dependent variable causes the remaining errors to be independent.

## TSCS analysis of major power trade

We return to a similar model of trade to that analyzed in the previous section. Here, however, we work with only the dyads made up of the seven major powers (as in Morrow, et al., 1999) for the period 1907–90. We believe that the dependence between dyadic trade flows can be modelled with a one year time lag rather than instantaneously to be defensible on substantive grounds. Adjustment is likely to occur based on observed flows, or recorded past flows, rather than anticipated current flows. As in our previous trade examples, the dependent variable is the natural log exports of one country to another (in constant dollars). We follow Morrow et al. more closely in this model. The democracy measure is a binary indicator of whether both members of each dyad are democracies, as indicated by whether the states score 6 or greater on the Polity institutionalized democracy scale. We also use the Tau-b measure of alliance portfolio similarity. Finally, we include two dummy variables indicating if the states in each dyad are members of an alliance during bipolarity or multipolarity. As before, we take the natural log of each variable for the gravity model specification. We follow Morrow et al in distinguishing between multipolar and bipolar time periods.<sup>19</sup>

Model 1 in Table 4 shows the OLS results for a specification assuming that all 42 dyads are independent of each other. Model 2 adds the temporal lag of the spatial lag. For both models, Lagrange multiplier tests indicate that there is a small, but statistically significant,

---

<sup>19</sup>Whereas 1999, 655, fn 13. apply a transformation to the first observation after a missing period in each country time series, we are not persuaded that this was a good way to proceed and have simply omitted the first observation after a missing series. These differences in how we treat missing periods imply that our results are not fully comparable to Morrow et al.

amount of serial correlation of the errors. This is not uncommon given the large sample sizes for TSCS data. There is no easy answer here as to how to proceed; we are pretty sure there is some serial correlation, but we are also pretty sure there is only a small amount of serial correlation (and the seriousness of the problem varies with the amount of serial correlation). For the purposes of this paper, we choose to continue with OLS, being aware that OLS is not perfect here (but also believing it to be superior to other alternative estimation strategies).

These results suggests although there is spatial dependence among the observations, the magnitude of the spatial association appears relatively modest once the previous influences have been taken into account by a lagged dependent variable. While the coefficient estimates in the second column of Table 4 are different from the estimates in the first column, these differences are small and nowhere near the differences we saw in the cross-sectional case.

We conjecture that for TSCS data with a lagged dependent variable the spatial effects will often matter less. The reason for this is that the lagged dependent variable already contains any prior spatial effects, and hence the spatial lag provides much less information in the TSCS model with a lagged dependent variable than it does in the cross-sectional context. We stress that the work is being done by the lagged dependent variable; spatial lags may have a strong impact in TSCS models which do not include the lagged dependent variable in the specification.

But even in our example, it is still worthwhile to include the spatial lag (the estimates with the spatial lag included must be superior to those which assume that there is no spatial effect, and if the spatial lag is insignificant, one can always then go back to the simpler

Table 4: Directed Export Flows, Major Powers, 1907–90

<b>Variable</b>	Coef (Ass. SE)	Coef (Ass. SE)
Constant	-0.25 (0.10)	0.17 (0.11)
Ln GNP A	0.03 (0.01)	0.02 (0.01)
Ln GNP B	0.04 (0.01)	0.03 (0.01)
Ln population A	0.02 (0.02)	0.04 (0.02)
Ln population B	0.02 (0.02)	0.03 (0.02)
Ln distance	-0.03 (0.01)	-0.04 (0.01)
Ln dau b	0.13 (0.06)	0.11 (0.06)
Ln democracy	0.13 (0.03)	0.14 (0.03)
Ln MID	-0.20 (0.04)	-0.20 (0.04)
Ln multipolar	-0.30 (0.05)	-0.28 (0.05)
Ln Bipolar	-0.06 (0.05)	-0.04 (0.05)
Ln $y_{t-1}$	0.92 (0.01)	0.91 (0.01)
<b>W</b> (ln $y_{t-1}$ )	–	0.02 (0.01)
N	2565	2565

model). If our substantive problem makes it plausible to assume that spatial effects also occur with a temporal lag, and if tests indicate that the remaining errors appear serially independent (or otherwise small enough to ignore), then it is easy enough to include the spatial lag in any model.<sup>20</sup>

This idea is not unknown to students of political economy, though they have usually not been explicit in their use of spatial econometrics. To take but two prominent examples, both Garrett (1998) and Iversen (1999) model economic outcomes and policies in the OECD nations (using TSCS data) as partly determined by economic performance in the other OECD nations. But Iversen (1999, 65) used only the simple OECD average of the dependent variable as a covariate, without any discussion of either the spatial or temporal lag structure. His use of simple averages implies that all connectivities are identical, and it is unclear whether the dependent variable for any country was included in the OECD average. Garrett also proceeded intuitively, using the trade weighted average of the growth of GDP in the OECD countries in all his regressions (both for growth and for other economic outcomes such as inflation and unemployment). Note that this is different from the spatial approach, which used the spatial lag of the dependent variable in any regression. So while the approaches of prominent political economists make intuitive sense, their analyses would benefit from more formal use of the spatial econometrics literature.

---

<sup>20</sup>We conjecture that spatial econometricians have not studied our proposed method because it is both econometrically trivial and because it does not lead to enormous changes in estimated coefficients.

## 7 Conclusions

Spatial econometric techniques are now starting to be used by political scientists. Because of the geographic heritage of these models, their primary application has been to incorporate physical notions of space (distance) into political models, and, particularly, to argue that geographically nearby units are linked together (or, less usefully, that their error terms are linked together). While this approach is highly promising, we think it can be made much more fruitful if we allow for interconnectivities that go beyond geography. In the end, our message is theoretical, not technical. In international relations and comparative political economy, we would expect units to be affected by the what takes place in other units. We would expect the connectivity of units to be a function of political and social, as well as geographic, variables. While the heritage of spatial econometrics is geographic, there is no reason to limit spatial econometric models to geographic modes of thinking.

With a richer view of what may constitute the interconnections between units, and the ability to test which of two (or perhaps more) interconnections are more important, the spatial autoregressive model should become an important tool in the kit of the political economist (whether in comparative politics or international relations). We do not think of nations as isolates and there is no reason that our models should either treat nations as isolations, or study interactions in a non-systematic manner.



## References

- Anselin, Luc. 1988. *Spatial Econometrics: Methods and Models*. Dordrecht: Kluwer.
- Beck, N. 2001. "Time-Series–Cross-Section Data: What Have We Learned in the Past Few Years?" *Annual Review of Political Science* 4:271–93.
- Beck, Nathaniel and Jonathan N. Katz. 1995. "What To Do (and Not To Do) with Time-Series Cross-Section Data." *American Political Science Review* 89:634–47.
- Beck, Nathaniel and Jonathan N. Katz. 1996. "Nuisance vs. Substance: Specifying and Estimating Time-Series–Cross-Section Models." *Political Analysis* 6:1–36.
- Beck, Nathaniel and Jonathan N. Katz. 2001. "Throwing Out the Baby With the Bath Water: A Comment on Green, Kim and Yoon." *International Organizations* 55:487–95.
- Beck, Nathaniel and Kristian Skrede Gleditsch. 2003. "Geography is More than Space." Paper presented at the Workshop on Geography, Conflict, and Cooperation, Joint Sessions Meeting of the European Consortium for Political Research, Edinburgh, UK, 28 March - 2 April.
- Conley, T.G. 1999. "GMM Estimation with Cross Sectional Dependence." *Journal of Econometrics* 92:1–45.
- Cornes, Richard and Todd Sandler. 1996. *The Theory of Externalities, Public Goods, and Club Goods*. 2nd ed. New York: Cambridge University Press.
- Deutsch, Karl W. and Walter Isard. 1961. "A Note on a Generalized Concept of Effective Distance." *Behavioral Science* 6:308–311.
- Dow, Malcolm M., Michael L. Burton, Douglas R. White and Karl P. Reitz. 1984. "Galton's Problem as Network Autocorrelation." *American Ethnologist* 11:754–770.
- Franzese, Robert and Jude C. Hayes. 2004. "Empirical Modeling Strategies for Spatial Interdependence: Omitted-Variable vs. Simultaneity Biases." Paper Presented at the 2004 Political Methodology Summer Conference, Palo Alto, CA, 29-31 July.
- Galton, Francis. 1889. "Comment on E.B. Tylor 'On a Method of Investigating the Development of Institutions: Applied to Laws of Marriage and Descent.'" *Journal of the Royal Anthropological Institute* 18:268–9.
- Garrett, Geoffrey. 1998. *Partisan Politics in the Global Economy*. New York: Cambridge University Press.
- Gleditsch, Kristian Skrede. 2002a. *All Politics is Local: The Diffusion of Conflict, Integration, and Democratization*. Ann Arbor, MI: University of Michigan Press.

- Gleditsch, Kristian Skrede. 2002b. "Expanded Trade and GDP Data, 1946-99." *Journal of Conflict Resolution* 46:712–724.
- Gleditsch, Kristian Skrede and Michael D. Ward. 2000. "War and Peace in Time and Space: The Role of Democratization." *International Studies Quarterly* 44:1–29.
- Gleditsch, Kristian Skrede and Michael D. Ward. 2001. "Measuring Space: A Minimum Distance Database and Applications to International Studies." *Journal of Peace Research* 38:749–768.
- Iversen, Torben. 1999. *Contested Economic Institutions: The Politics of Macroeconomics and Wage Bargaining in Advanced Democracies*. New York: Cambridge University Press.
- Jagers, Keith and Ted R. Gurr. 1995. "Tracking Democracy's 'Third Wave' with the Polity III data." *Journal of Peace Research* 32:469–82.
- Kmenta, Jan. 1986. *Elements of Econometrics*. Second ed. New York: Macmillan.
- Lacombe, Donald. 2004. "Does Econometric Methodology Matter? An Analysis of Public Policy Using Spatial Econometric Techniques." *Geographical Analysis* TBA:forthcoming.
- Lipset, Seymour M. 1960. *Political Man: The Social Bases of Politics*. Anchor.
- Lofdahl, Corey L. 2002. *Environmental Impacts of Globalization and Trade: A Systems Study*. Cambridge, MA: MIT Press.
- Mansfield, Edward D. and Rachel Bronson. 1997. "Alliances, Preferential Trading Arrangements, and International Trade." *American Political Science Review* 91:94–107.
- Morrow, James D., Randolph M. Siverson and Tressa E. Tabares. 1999. "The Political Determinants of International Trade: The Major Powers, 1907-90." *American Political Science Review* 92:649–661.
- Ord, J. Keith. 1975. "Estimation Methods for Models of Spatial Interactions." *Journal of the American Statistical Association* 70:120–126.
- Pace, Kelley and Ronald Barry. 1997. "Quick Computation of Spatial Autoregressive Estimators." *Geographical Analysis* 29:232–246.
- Parks, Richard. 1967. "Efficient Estimation of a System of Regression Equations When Disturbances Are Both Serially and Contemporaneously Correlated." *Journal of the American Statistical Association* 62:500–509.
- Ward, Michael D. and Kristian Skrede Gleditsch. 2002. "Location, Location, Location: An MCMC Approach to Modeling the Spatial Context of War and Peace." *Political Analysis* 10:244–260.