GEOGRAPHIC AND NON-GEOGRAPHIC CLUSTERING OF DEMOCRATIC COUNTRIES: A SPATIAL ECONOMETRIC ANALYSIS

by

Dominique Brabant

An essay submitted to the Departments of Economics

in partial fulfillment of the requirements for

the degree of Master of Arts

Queen's University

Kingston, Ontario, Canada

August 2010

Acknowledgments

I would like to thank Professor Huw Lloyd-Ellis for his assistance during my time as a graduate student as well as his comments and suggestions for this paper. I would also like to thank my colleagues who patiently supported me throughout this year and provided valuable help for this paper. All errors and omissions are my own.

Table of Contents

1. Introduction

Political decisions are complex and are often determined by a large number of influences. These influences consist of both domestic issues within a country as well as international interactions, be it by trade, conflict or governmental change. In this paper, I choose to focus on democracy, and how a country's political system is influenced by other countries. I study not only the spread of democracy through time and space, but also how democratic countries are clustered. These clusters can take several forms, reflecting alternative definitions of space.

There is a rich literature on how political and economic decisions are affected by the international conjuncture. The domino theory, a term first used by the United States President Dwight Eisenhower in 1954, claims that if a country becomes organized under a certain ideology, there is better chance that this ideology will spread in surrounding countries (Leeson and Dean 2009). While the idea was then applied to stop the spread of communism, researchers from that time have used it for all kinds of political or economic decisions. Simmons et al. (2006) study a more general phenomenon that they call international policy diffusion, which describes how a country's decision-makers are influenced by choices made in the past by other countries' decision-makers.

The spread of democracy through time and space has been mostly studied following Eisenhower's point of view, an interventionist approach. As I consider democracy as a national decision, influenceable by international factors, the approach I adopt for this paper is closer to the approach adopted by Simmons et al. (2006). In their paper, they describe four mechanisms explaining international policy diffusion: coercion, competition, learning, and emulation:

- Coercion affects the domino theory, because powerful countries are assumed to have a significant influence on weaker countries, indirectly through international organizations or directly by imposing policies.
- Competition relates to economic relationships and presumes that countries compete with each other to find partners. It influences decision-makers, especially in the development of economic policies.
- Learning reflects the fact that decision-makers are aware of other countries' decisions and can observe the results they have. They can then make their own decisions according to their observations. Simmons et al. (2006) report that a successful decision has a greater likelihood of being reproduced elsewhere than an unsuccessful one.
- Emulation assumes that there are large consensuses throughout the world about what a society should aim for and how to attain it. The authors point out, however, that powerful countries can dictate consensuses more easily, and so emulation can dominate coercion.

Although the domino effect has been studied in many papers, very few of them have tried to evaluate it empirically, especially in the case of democracy's diffusion. Recently, however, two papers have attempted to explain the spread of democracy using spatial econometrics. Leeson and Dean (2009) try to evaluate this theory, using data from 1850 to 2000. They find that the domino theory does apply, but only moderately. A 10 point increase in the average democracy index of its neighbours is associated with a 1.1 point increase in a country's democracy index, which confirms that democracy changes are contagious. Beck et al. (2006) also try to evaluate the domino theory by using as a definition of “neighbours” not only geographic neighbours, but also main trading partners. Using cross-sectional data, they obtain positive and significant spatial effects, but they do not push their interpretation any further, with the goal of their paper being to explain how to use spatial econometrics in that context.

In this paper, I aim to build on and extend these analyses, employing panel data and alternative definitions of “neighbourhood”. Following Simmons et al. (2006), I believe that various international relationships may have an impact on political decisions and can therefore be modeled differently than previously attempted. For example, using other variables like trade and immigration can help researchers understand them better than using only geography. Performing an empirical study with panel data, rather than only cross-sectional data, allows me to include dynamic interactions between countries, not only the contemporaneous ones. This strategy has an important impact on the interpretation of the results. I also carry out two different analyses, the first one explaining the level of democracy in a country, the second one explaining the change in the level of democracy. This allows me to capture all the possible interactions between countries. The empirical analysis employs spatial econometrics, a branch of econometrics described later. Estimates are obtained using Maximum-likelihood, through software developed by James P. LeSage and J. Paul Elhorst.

This paper is original in that it studies the spread of democracy using alternative definitions of space, and employs panel data. Also, the idea of associating immigration with the spread of democracy is an empirical research topic that is completely new. The results I obtain in this paper show that not only do geographic neighbours have a significant influence on a country’s level of democracy, but immigration and trading partners can have an even greater influence. These results have to be taken with care, though, as the models used may contain endogeneous regressors. I address this concern as much as possible by considering the sensitivity of the results to various alternative assumptions. The next sections of this paper are organized as follows. Section 2 provides an overview of spatial econometrics, section 3 explores the spread of democracy in the literature, section 4 details the methodology used, section 5 explains the results, section 6 interprets the results, section 7 introduces alternative results, section 8 performs a sensitivity analysis, and section 9 outlines the main conclusions of the study.

2. An overview of Spatial Econometrics

2.1 History

The idea of spatial econometrics first appeared in Europe in the 1970's. However, the broad concept of geographic dependence was proposed more than a century ago by Sir Francis Galton, which gave birth to "Galton's problem" (Franzese and Hays 2008). According to Franzese and Hays (2008), Galton's problem resides in the difficulty of separating two different effects reflected by the data. These two effects are interdependence – also called spatial spillovers in the literature – and common shocks, a particular event having effects on more than one region. Anselin (1999) suggests that the main motivation for developing spatial econometrics is to model data related at different levels of geographic region, such as comparing spatially dependent provinces belonging to different countries. Since its creation, spatial econometrics has been used in many areas of social sciences, beginning with geography and regional science, and later in sociology and political science. It is now increasingly used in several spheres of economics, such as labour, agricultural and environmental economics (Anselin 1999).

Forty years ago, Curry (1970) and Gould (1970) explicitly brought up the problems of modelling, estimating and interpreting data containing spatial relations. However, Hordijk and Paelinck (1976) were the first to lay out the foundations for spatial econometrics in the form that it is used today and described thereafter.

2.2Basic Relationships

In classical econometrics, when one tries to model a linear relationship between two or more variables, assuming there is no form of spatial relationship, he or she usually uses the following regression:

$$y_i = X_i\beta + \varepsilon_i,\ \varepsilon_i \sim N(0,\sigma^2) \tag{1}$$

where y_i is the dependent variable, X_i contains one or more explanatory variables and β their coefficient, and ε_i is a random error term. However, this model becomes inaccurate when the presence of some sort of spatial relationship is suspected. LeSage and Pace (2009) propose a rudimentary attempt to fix this problem by adding the value of the dependent variable of an observation's neighbour, or the average value of all of its neighbours, as an explanatory variable. This allows an observation's neighbour to have an impact on its explained variable. LeSage and Pace (2009) get the following regression:

$$\begin{aligned} y_i &= \alpha_i y_j + X_i\beta + \varepsilon_i,\ \varepsilon_i \sim N(0,\sigma^2) \\ y_j &= \alpha_j y_i + X_j\beta + \varepsilon_j,\ \varepsilon_j \sim N(0,\sigma^2) \end{aligned} \tag{2}$$

It is important to note that, as there is no time index, this model only applies to cross-sectional data. Therefore, every observation must correspond to a different geographical location. The definition of "neighbour" is chosen by the researcher. It is usually related to geography, using common borders or Euclidean distance, but it does not have to be.

The problem with the model proposed in equation (2) is that if an observation has more than one neighbour and the researcher considers a large number of observations, the number of equations in the model will grow rapidly. In order to put some structure in the various possible neighbourhood relationships between the different observations, LeSage and Pace (2009) express their model as showed in equation (3).

$$y_i = \alpha + \rho \sum_{j=1}^{n} w_{ij} y_j + \varepsilon_i,\ \varepsilon_i \sim N(0, \sigma^2) \tag{3}$$

where w_{ij} is element i,j of a spatial n x n weight matrix[1], and ρ describes the spatial interdependence. From the perspective of Galton's problem, regression (3) is capturing the interdependence part of the relationship, as opposed to that due to common shocks.

2.3 Weight Matrices

In spatial econometrics, weight matrices are used to specify the neighbourhood relations in the regression. A weight matrix must correspond as best as possible to existing spatial relationships between different observations. It is a square matrix whose dimensions correspond to the number of spatial observations, and that is usually mainly composed of zeros – a sparse matrix. Even if forming a weight matrix seems simple, the methodology used can have an important impact on the results of the analysis. According to Anselin (1988), it is one of the most important and most controversial issues in spatial econometrics. Any researcher must be very careful when modeling the structure of the spatial dependence, because the estimation and interpretation of the spatial dependence coefficient will depend greatly on the weight matrix that has been used. Anselin (1988) suggests that this structure should reflect concepts such as accessibility and potential

[1] Weight matrices are also referred as connectivity matrices in the literature and in this text.

instead of a possible link observed directly in the data. In other words, he recommends creating a weight matrix based on theory rather than based on observed data. Otherwise, this could lead to a circular reasoning.

In theory, nothing prevents weight matrices to be time varying (Beck et al. 2006). Of course, using a different matrix for each time period makes the estimation far_more complicated computationally. This is why the available software does not yet allow using more than one connectivity matrix per regression. As the usage of alternative definitions of space is recent in political science, the idea of time-varying connectivity matrices has not been empirically exploited yet.

2.4 SAR and SEM models

There are two main kinds of models in spatial econometrics, corresponding to the two different impacts in Galton's problem. The model described in equation (3) corresponds to the interdependence effect and is also called first-order spatial autoregressive process. In matrix notation, it becomes equation (4), corresponding to the Data Generating Process (DGP) from equation (5):

$$y = \rho W y + \varepsilon,\ \ \varepsilon \sim N(0, \sigma^2 I_n) \tag{4}$$

$$y = (I_n - \rho W)^{-1}\varepsilon,\ \ \varepsilon \sim N(0, \sigma^2 I_n) \tag{5}$$

Adding other non-spatial explanatory variables, LeSage and Pace (2009) define the Spatial Autoregressive model (SAR) as follows:

$$y = \rho W y + X\beta + \varepsilon,\ \ \varepsilon \sim N(0, \sigma^2 I_n) \tag{6}$$

where X may contain a constant or not.

The other family of models relates to the second impact in Galton's problem: common shocks. If a particular event, or in the non-temporal case a particular variable, is impossible to measure and has a similar impact on more than one observation, the spatial relationship will be included in the error term instead of being part of the explanatory variables. LeSage and Pace (2009) define the Spatial Error model (SEM) as follows:

$$\begin{aligned} y &= X\beta + u \\ u &= \rho W u + \varepsilon \qquad \varepsilon \sim N(0, \sigma^2 I_n) \end{aligned} \tag{7}$$

where u is the disturbance and ε is a random error for each observation. The spatial relationship is now in the disturbance, as opposed to the explanatory variables.

Links have often been made in the literature between spatial econometrics and time-series econometrics. Even if these similarities can be useful to gain a better understanding of the basic concepts of spatial econometrics, the analogy does not go any further. The timeline used in time-series analysis is usually unambiguous, as the relations between data go in only one direction or two if the expectations are taken into account. In spatial econometrics, these relations can go in several different possible directions and an observation can have a various number of neighbours. If cross-sectional data is used, the influence goes back and forth between two observations until a stable equilibrium is reached. If panel data is used, the neighbourhood relationships can vary with time. Hence, both the estimation and interpretation of spatial econometrics models are very different from time series analysis.

2.5 Estimation of the SAR and SEM models

Researchers have developed several methods to estimate spatial econometrics models, trying to get the best possible coefficients, while fixing, or at least minimizing, all possible issues that could arise. They also have to consider the computational feasibility of these methods. First of all, it is desirable to explain why developing these techniques is required, as the models to be estimated are, at first sight, not much different from the ones usually estimated by Ordinary least squares (OLS). It seems clear that simply estimating a model like the one presented in equation (1) using data containing spatial dependence yields inaccurate results. Franzese and Hays (2008) explain that as there is likely a spatial relationship between the explanatory variables, OLS will overestimate their impacts. However, even if appropriate spatial lags are included in the model, such as in equation (6), OLS estimates will still be biased. As the residual ε is correlated with Wy, the estimated coefficients will be inconsistent and suffer from endogeneity and simultaneity biases.

Franzese and Hays (2004) compare several approaches to estimate spatial econometric models. They obtain the results I just mentioned about OLS and they also discover that Maximum-likelihood estimation (MLE) yields consistent and asymptotically efficient estimates. However, at that time, this method is hard to implement, so they introduce another approach called "Spatial two-stage-least-squares, instrumental variables" (S-2SLS-IV). This approach uses all of the independent variables included in X as an instrument for y in the spatial dependence explanatory term, so that WX replaces Wy in equation (6). As long as the sample is large and the instruments are fully exogenous, it yields as precise estimates as MLE but it is easier (though still difficult) to implement. Since 2004, software using MLE for spatial econometrics has been developed, it is the method I use in this paper and briefly describe here.

For the SAR model from equation (6), LeSage and Pace (2009) define the log-likelihood function to maximize as the following:

$$\ln L = -\left(\frac{n}{2}\right)\ln\left(\pi\sigma^2\right) + \ln\left|I_n - \rho W\right| - \frac{e'e}{2\sigma^2},\; e = y - \rho W y - X\beta \qquad (8)$$
$$\rho \in (\min eigenvalue(W), \max eigenvalue(W))$$

where n is the number of spatial observations. LeSage and Pace (2009) argue that to estimate a SAR model by MLE, the best method is to use a concentrated log-likelihood function. The first step is to differentiate the function with respect to β and σ^2, and to find each their closed-form solution as a function of ρ. These two expressions can then be plugged back in the first function, which can now be optimized with respect to ρ to find $\hat{\rho}$. β and σ can finally be estimated by substitution. This method finds the same solution as "normal" maximum likelihood estimation[2]. LeSage and Pace (2009) derive the following concentrated log-likelihood function:

$$\begin{aligned}
&\ln L = k + \ln\left|I_n - \rho W\right| - \left(\frac{n}{2}\right)\ln\left(e'e\right)\\
&e = e_0 - \rho e_d\\
&e_o = y - X\beta_0\\
&e_d = Wy - X\beta_d\\
&\beta_0 = (X'X)^{-1}X'y\\
&\beta_d = (X'X)^{-1}X'Wy
\end{aligned} \qquad (9)$$

where k is a constant. Finding the determinant of the matrix $|I_n\text{-}\rho W|$ is the part of this approach that is hard to compute for large samples, because this matrix is of dimensions n x n. To facilitate the estimation, researchers create tools using the sparsity of the connectivity matrix.

[2] For further discussion on concentrated log-likelihood functions, see Davidson and MacKinnon (2004).

To estimate a SEM model as defined in equation (7), LeSage and Pace (2009) use a similar approach. The log-likelihood function is presented in equation (10) and the concentrated one in equation (11):

$$\begin{aligned} &\ln L = -\left(\frac{n}{2}\right)\ln(\pi\sigma^2) + \ln\left|I_n - \lambda W\right| - \frac{e'e}{2\sigma^2} \\ &e = (I_n - \lambda W)(y - X\beta) \end{aligned} \qquad (10)$$

$$\begin{aligned} &\ln L(\lambda) = k + \ln\left|I_n - \lambda W\right| - \left(\frac{n}{2}\right)\ln S(\lambda) \\ &S(\lambda) = e(\lambda)'e(\lambda) \end{aligned} \qquad (11)$$

S(λ) is harder to estimate than S(ρ) in the SAR model, but it can be done using numerical methods, like Newton's method[3].

2.6 Goodness-of-Fit of Spatial Models

Evaluating the goodness-of-fit of different models and selecting the best model in spatial econometrics is not an easy task. Unlike classical econometrics, there is no general measure that can be used like R-squared. Anselin (1988) stresses that R-squared is an invalid measure in a context where a spatial model is estimated by Maximum-likelihood. Focussing on models with cross-sectional data, he suggests a few alternatives, including a pseudo R-squared – a squared correlation between fitted and observed values – and the maximized log-likelihood. However, he warns that even these two measures may yield different rankings when comparing the same models.

Elhorst (2010) describes goodness-of-fit measures for spatial models including panel data and estimated by Maximum-likelihood specifically. When panel data is used,

[3] For further explanations on Maximum-likelihood estimation in spatial econometrics, see Anselin (1988) or LeSage and Pace (2009).

the presence of possible spatial and/or temporal fixed or random effects changes the way to evaluate the goodness-of-fit. He mentions a measure, the correlation-squared, which recalls Anselin's (1988) pseudo R-squared, but takes into account the fixed effects. The formula for calculating correlation-squared is in equation (12):

$$corr^2(Y,\hat{Y}) = \frac{\left[\left(Y-\bar{Y}\right)'\left(\hat{Y}-\bar{Y}\right)\right]^2}{\left[\left(Y-\bar{Y}\right)'\left(Y-\bar{Y}\right)\right]\left[\left(\hat{Y}-\bar{Y}\right)'\left(\hat{Y}-\bar{Y}\right)\right]} \quad (12)$$

In the context of a SAR model containing spatial fixed effects, correlation-squared ignores the part of the regression explained by the spatial fixed effects. As fixed effects can be an important determinant of the explained variable, considering them would yield a high goodness-of-fit without it being linked to the specifications of the model tested. Therefore, the difference between R-squared and correlation-squared reflects the importance of the impact of the spatial fixed effects.

2.7 Extensions

LeSage and Pace (2009) introduce a number of other models, all based on SAR and SEM, but exploring other possibilities to model spatial interactions. The SAC model is a general form of the spatial regression model, combining both the interdependence and the common shock effects. It is defined as follows:

$$\begin{aligned} y &= \rho W_1 y + X\beta + u \\ u &= \theta W_2 u + \varepsilon,\ \ \varepsilon \sim N(0,\sigma^2 I_n) \end{aligned} \quad (13)$$

where W_1 and W_2 are two weight matrices that can be the same or not.

The Spatial Durbin model (SDM), also introduced by LeSage and Pace (2009), is motivated by omitted explanatory variables and allows for these omitted variables to be space-dependent. It takes the following form:

$$y = \rho W y + X\beta + WX\gamma + \varepsilon, \ \varepsilon \sim N(0, \sigma^2 I_n) \tag{14}$$

where γ is the spatial dependence coefficient for explanatory variables, and the other variables are as defined before. The third extension presented by the same authors recalls a model introduced earlier by Anselin and Bera (1998). It is the spatial autoregressive moving average model (SARMA) and it is shown below:

$$\begin{aligned} y &= \rho W_1 y + X\beta + u \\ u &= (I_n - \theta W_2)\varepsilon, \ \varepsilon \sim N(0, \sigma^2 I_n) \end{aligned} \tag{15}$$

where θ is the spatial coefficient for the error term. This model uses a moving-average process to model the disturbance, instead of an autoregressive process as used in the SAC model. The SARMA model is used when the researcher suspects that the spatial dependence is local in the error term – only direct neighbours have an influence – and global in the dependent variable.

2.8 Future of spatial econometrics

Spatial econometrics, as a new approach developed only in the past four decades, is still incomplete in many areas. Anselin (2007) identifies several challenges for the future. He first mentions that it is necessary to develop new models that will be able to reflect more complex interactions between different locations. He also argues that the models that already exist need more theoretical background. The computation of data and results is another important issue. Even if software products already exist exclusively for the usage of spatial econometrics[4], they do not entirely satisfy researchers' needs. Anselin (2007) mentions the need to develop tools that can handle more data and more complex data. Paelinck (2005) points out that the current software allows researchers to find some estimates, but that these estimates do not always solve their problems. Meticulous researchers have to program their own functions in order to respond to their own needs. This is an important limitation to research.

More generally, Pinkse and Slade (2009) call into doubt the validity of the SAR model. They argue that the normality assumption is implausible, that some relationships maybe endogenous, and that these relationships may not be linear. This critique meets Anselin's call for new models to be developed. They also criticize the "time-series-analysis" roots of spatial econometrics, arguing that stationarity is unlikely, observations are not at equal distance and that if the data set changes, the structure of interdependence may change too, but the connectivity-matrix is fixed in the regression. For the endogeneity problem, it can be fixed theoretically using GMM or IV, but as most software use MLE this issue remains to be addressed for empirical studies (Pinkse and Slade 2009).

3. The spread of democracy in the literature

The possible determinants of democracy have been examined, among others, by Barro (1999). In an empirical study covering over 100 countries and six time periods from 1975 to 1995, he tests several determinants, using their value five years prior to the time where the dependent variable was observed, to allow them the time to have an

[4] Those include LeSage's toolbox for spatial econometrics, using Matlab, Pace and LeSage's toolbox for spatial statistics, also using Matlab, SpaceStat, a recent software developed by Terraseer for spatial statistics and econometrics, and GeoDa, developed by GeoDa Center.

impact and to avoid simultaneity bias. The lagged values are instruments for the contemporaneous ones. He finds that the average number of years attained at the primary level of school and a (smaller) difference in primary attainment between males and females have significant explanatory power for a country's democracy. He finds that per capita GDP has a positive and significant impact on democracy level. He also mentions a negative impact for being an oil exporting country.

The theoretical process by which democracy diffuses through time and space has been widely studied in the literature. However, as extensive and reliable numerical measures of democracy have been developed only recently, the spread of democracy has not been studied much empirically. The first paper using spatial empirical methods to examine this topic is Ward et al. (1997). They try to find answers to three major questions: Is democracy established through domestic or international forces? Does democracy spread through contagion or global process? Does democracy spread wholesale or in parts? Obtaining significant Moran's I values[5], they conclude that democratic countries are indeed clustered in regions. They manage to map this spatial and temporal clustering, but they do not consider any other possible explanatory variables.

Leeson and Dean (2009) also study this topic, but using recent developments of spatial econometrics to obtain more precise results. Their study covers the period from 1850 to 2001. Neighbours are defined as countries sharing a common border. They compare results obtained by SAR and SEM models and conclude that democracy spreads at a modest rate through space and time. Beck et al. (2006) adopt a different approach, using only cross-sectional data, and defining neighbourhood not only by geographic

---

[5]Moran's I is a measure of spatial autocorrelation among a set of observations introduced by Moran (1950). It is usually used as a test for presence of spatial autocorrelation. More information on Moran's I can be found in Anselin (1988).

location, but also by main trading partnerships. They argue that using alternative definitions of space allows one to study different influences, especially relevant in the context of research in political science. They conclude that trading with a more democratic country has a positive impact on a country’s level of democracy. This impact is slightly more important than the one between geographic neighbours.

Two main differences stand out between Leeson and Dean (2009) and Beck et al. (2006). They both rely on assumptions made by the researchers. The first one is the set of data employed for the empirical analysis. As mentioned earlier, using panel data instead of cross-sectional data introduces several changes in modeling and in the interpretation of results. Anselin (1988) mentions that when panel data is available a much wider set of models can be used, increasing the chances the model approximates the “real” pattern. There is another main difference between panel and cross-sectional analysis. In the latter, as time is not considered, it is assumed that the long-run equilibrium or at least a steady state has already been reached. In the context of the spread of democracy across countries, this may not make much sense. Overall, though, the estimation of models using panel data is undertaken using similar methods as estimations using only cross-sectional data.

The second difference reflects a new direction taken by empirical researchers, especially political scientists, using spatial econometrics. As it leaves open some interesting findings about the diffusion of political decisions, the use of new definitions of neighbourhood relationships has gained popularity in recent years. In spatial econometrics, nothing requires that the “distance” between countries be geographic. Isard (1969) was the first to extend the idea of space to a more general definition. However, Anselin (1988) warns that the connectivity matrix is always assumed to be exogenous.

Therefore, defining a new “map of the world” requires a lot of care. Using other variables like trade partnerships to construct the weight matrix can cause endogeneity bias. In the sensitivity analysis, I study the impacts on results of a potential endogeneity bias.

4. Methodology

The software I use for the empirical analysis is the Spatial Econometrics Toolbox, developed by James P. LeSage. It contains several new functions for econometric research and a few models to study spatial data. However, the functions made especially for panel data, which I am using, are newly developed and still incomplete. In the context of this research, it would be interesting to use the SAC model, as it is intuitive that both kinds of impacts from Galton’s problem affect democracy. As the SAC model is not yet available, I had to choose between SAR and SEM models. The four channels explained by Simmons et al. (2006) suggest that the level of democracy itself in a country can influence its neighbours, not only omitted variables in the error term. Therefore, the choice of SAR model seems more reasonable. The functions for this model have been updated in 2008 by J. Paul Elhorst and have not been perfectly tested yet. In order to be able to use them with my data, I had to modify a few of the Matlab’s commands. This reduces the precision of the results I obtain, but the difference is negligible.

The dependent variable of the models I use is the level of democracy in a particular country, in a particular year. For the measure of democracy, I use the “POLITY2” variable from the polity IV project. It is a very large database containing data on political regimes for a large numbers of countries and years[6]. This variable is obtained

[6] It is the most widely used dataset on this topic. http://www.systemicpeace.org/polity/polity4.htm.

by subtracting the country's autocracy score from its democracy score[7]. It can take integer values between -10 and +10, where -10 is strongly autocratic and +10 is strongly democratic. The variable includes non graded observations, where cases of foreign interruption are treated as missing data. I allocate a neutral democracy level of zero to these observations, in order to minimize the impact of this transformation. The "POLITY2" variable is the measure of democracy both Leeson and Dean (2009), and Beck et al. (2006) used. However, unlike Leeson and Dean (2009) who also use panel data, I decide to use not only the change of the level of democracy, but also the level itself. The first difference model assumes that all observations having a value of zero for dependent variable – no change in the political regime – do not have any impact on their neighbours. However, Simmons et al. (2006) explain that through coercion, powerful countries may have an important impact on the diffusion of political decisions. As some powerful countries are politically stable (United States, United Kingdom and China among others), using only the first difference does not take coercion into account. Nonetheless, I also run models where the explained variable is the first difference of the level of democracy, as it allows me to compare my results with the ones of Leeson and Dean (2009). It also has a different, and more intuitive, interpretation. As a sensitivity analysis, Leeson and Dean (2009) also run their models using level of democracy instead of change, and the results they obtain confirm the initial ones: the spatial dependence coefficients stay small and significant.

[7] Autocracy and democracy scores are each obtained by a set of variables computed following a particular scale. "DEMOCRACY" contains four variables, such as openness and competitiveness of executive recruitment and "AUTOCRACY" contains five variables such as regulation and competitiveness of participation.

Considering Barro's (1999) findings, I select a set of three independent variables, besides the spatial influence, to explain democracy. These variables are chosen to explain the level of democracy specifically, and not the change in democracy. The choice of per capita GDP as an explanatory variable follows the widespread idea that wealthier nations tend to be more democratic. The data for per capita GDP is obtained by combining World Bank and UN datasets. It is transformed into constant 2000 $U.S. using the Consumer Price Index historical data released by the U.S. Department of Labor. For countries where sufficient data was only available from the UN, I use the 1970 value for 1960 and 1965 to avoid creating any misleading trends[8].

The second explanatory variable is primary school attainment. It is expected to have positive impacts on democracy, following the idea that a more educated population would want greater participation in a country's administration. The third variable is the difference between male and female primary school attainment. Barro (1999) recalls Tocqueville's (1835) idea that better education opportunities for females are usually present when the social structure is more participatory. The data for primary education variables are available from World Bank databases. I use a measure of the average number of primary school years attained by the total population of 15 years old and above. For the primary education difference between genders, I subtract the females' value from the total value. The resulting variable is therefore not the difference between males and females, but the difference in education between the total population and females.

If it is plausible that GDP, primary education attainment and difference in primary education attainment between genders can have an impact on a country's level of

[8] The details for the data sources can be found in appendix 1.

democracy, the opposite is also possible. This would create a simultaneity bias in the results obtained. In order to minimize this bias, the explanatory variables used are lagged by one time period – generally five years – following Barro's (1999) methodology.

Using the variables presented above, the general form of the model used for the empirical analysis is the following:

$$\begin{aligned} y_t &= \rho W y_t + X_{t-1}\beta + u \\ u &= \mu_i + \varepsilon,\ \varepsilon \sim N(0,\sigma^2) \end{aligned} \tag{16}$$

where y is the dependent variable, ρ is the spatial dependence coefficient, W is the connectivity matrix, X is the set of explanatory variables including a constant, β is the set of their respective coefficients, and u is the error term. The error term is composed of a spatial fixed effect, μ_i, and a random component, ε, which follows a normal distribution. A model with spatial fixed effects assumes that each country has its own context and it is equivalent to adding a new dummy variable for each country – minus one – in the sample. To assess whether the fixed effects are significant, an LR-test for joint significance is performed. This test compares the log-likelihood of regressions run with and without fixed effects, with the number of countries as degree of freedom.

In this paper, I compare five different models explaining the spread of democracy. They are different in the way spatial dependence is modeled, so they all include a different connectivity matrix. The first one uses a geographic notion of space and connects countries that share a common border[9]. The second connectivity matrix is similar to the first one, but the countries are weighted by their 2009 populations, where population is used as an instrument for the country size. The hypothesis behind this idea is that larger countries could have a greater influence on their neighbours than smaller

[9] See appendix 2 for the list of countries with their geographic neighbours.

ones. Leeson and Dean (2009) perform the same transformation and obtain no different results when weighting or not weighting the countries. The data I use for population is from the U.S. Census Bureau.

These two connectivity matrices both contain the problem of having to deal with islands. These countries are effectively assumed to have no external influences, which is an unrealistic assumption. However, as they do not share a border with any other country, it is reasonable to assume that they are harder to influence, and so that giving them "normal" neighbours may not be realistic. Also, it would require determining subjectively who these neighbours are. I test for possible bias caused by the existence of islands later in the sensitivity analysis.

The advantage of constructing a connectivity matrix using something other than just geography is that it allows the dependences between countries to be asymmetric. The third connectivity matrix I use relates countries to each other by their trading partnerships. Even if countries tend to trade more with partners geographically closer to them, a trading partnership reflects another kind of relationship, one that could have a different, and maybe greater, influence on a country's policies. I assume that a country can be more influenced by another country from which it is importing. In order to form the matrix, I express the imports of country A from country B as a percentage of country A's GDP. When the value obtained is higher than one percent, I put the value into the matrix. When the value is smaller than one percent, I put a zero. The imports data are from the "Direction of Trade" database created by the International Monetary Fund. I use data from 2008 as this year is available for the most countries. Other years are also used for smaller subsamples in the sensitivity analysis. The 2008 GDP data is from the World

Bank database, except for Iraq where it is the value from 2003 (same source) and from Cuba, where the value is an estimation from Index Mundi[10].

The fourth and fifth connectivity matrices are based on population movements. Migrations between countries refer to another kind of relationship. They reflect, among other things, sociological similarities – such as language, religions, and ideologies – that are not easily measured. Assuming the four mechanisms proposed by Simmons et al. (2006) to explain international policy diffusion and assuming that the population has some degree of power in determining what kind of regime hold power in the country, it makes sense to believe that migrations can influence the spread of democracy. To form the fourth connectivity matrix, I divide the number of people – as a stock – who have emigrated from one country to another, by the origin country's total population. I then only keep the numbers over 0.001, leaving zero values for other observations. This procedure assumes that a country is more influenced by those countries to which its citizens move.

The fifth connectivity matrix is also based on population movements but the influence is inverted. The idea is that if a host country has a particularly large fraction of its population that originated from another country, then this part might have enough power to influence the host country's political regime. This connectivity matrix is formed by transposing the original raw numbers matrix, then dividing by the population of the host country. I keep only values greater than 0.001, while other observations are given zeros. The migration data is from the Development Research Centre on Migration,

[10] Index Mundi is a website combining diverse databases to put together detailed country profiles, available at http://indexmundi.com/.

Globalisation and Poverty based at the University of Sussex[11]. To form the matrix, the researchers use a combination of both place of birth and citizenship as a definition of "migrant". They use data collected from each country's censuses. The complete methodology used by the centre is detailed in Parsons et al. (2005) [12]. As the data comes separately from countries, it does not apply to the exact same year for every country, but it is all close to 2000. The population data was also from the U.S. Census Bureau for the year 2000.

All the connectivity matrices used are row-normalized, which means that the matrix is transformed so that all elements of each row sum to one. This step is particularly important when one wants to compare several matrices; it is giving equal importance to the connectivity matrices in all regressions run, regardless of the units chosen. It is also allowing equal exterior influence to each observation, regardless of the number of neighbours it has.

As mentioned previously, it is very important that the connectivity matrices are exogenous. When the connectivity matrix uses only geography, perfect exogeneity is obvious, as countries are not responsible for their geographic location. However, when other variables are used, endogeneity problems may appear. One could think that a country's choice of trade partners is related to its partners' democracy level. A more democratic country could trade more with a more democratic partner, through ideological or sociological reasoning, or through free-trade agreements. This would create an endogeneity bias. In the sensitivity analysis, I compare the results obtained from models where the connectivity matrix is formed using the trading partnerships at the beginning

[11] http://www.migrationdrc.org/
[12] I use the fourth version of their matrix. As a greater number of values are estimated, it is less accurate, but it has the advantage of being more complete.

and at the end of the subsample time period. This should detect the presence of an endogeneity bias if there is one.

For the immigration case, the problem appears to be similar. One could suppose that people usually move from less democratic to more democratic countries to increase their quality of life. However, especially for regional migration, one could also think that migrants move to countries with the highest growth rate. With the relationship between growth and democracy not clearly defined, the existence of an endogeneity bias is ambiguous. Due to insufficient data, I cannot perform the sensitivity analysis for endogeneity in this case[13].

Due to availability of the data and historical reasons, I divide the data into two subsamples, one going from 1965 to 2005, the other one from 1995 to 2005, with each time period separated by five years. The first subsample allows me to study relationships between countries on a longer time period, but with fewer countries available. The second subsample of 130 countries allows me to analyze a larger number of countries. Lots of countries were just created around 1990 after the fall of Soviet Union. Using the smaller sample allows me to also study the progression of democracy in countries passing from communism to democracy, something not possible on a longer time period[14].

As mentioned earlier, estimations are made using MLE. Elhorst (2003) uses the following log-likelihood function to solve the model with a spatially lagged dependent variable and spatial fixed effects (see equation (16)):

---

[13] The bilateral immigration data is currently only available for one time period. However, Professor Chris Parsons is working on a new bilateral dataset which contains bilateral immigration stocks for the 1960-2000 period. It should be available on the World Bank website in September.

[14] As for several countries of this subsample data is only available from 1991, the lag of the first period's dependent variables is only four years. For the first-difference regressions, the first period's values also refer to the difference between 1991 and 1995 values.

$$L = -\frac{NT}{2}\ln(2\pi\sigma^2) + T\sum_{i=1}^{N}\ln(1-\delta\omega_i) - \frac{1}{2\sigma^2}\sum_{t=1}^{T} e_t{}' e_t \qquad (17)$$

$$e_t = (I - \delta W)(Y_t - \bar{Y}) - (X_t - \bar{X})\beta$$
$$\bar{Y} = (\bar{Y}_1, \ldots, \bar{Y}_N)'$$
$$\bar{X} = (\bar{X}_1{}', \ldots, \bar{X}_N{}')'$$

This is the function that is maximized to obtain the estimates discussed below.

5. Results

The results obtained are presented in Tables 1 and 2. Table 1 contains the results for the first subsample, with data going from 1965 to 2005, and Table 2 contains the results for the second subsample, using data from 1995 to 2005. Each table contains results for five different models; every model is composed of a different connectivity matrix.

At first sight, a few similarities can be observed between the five models. The first similarity is that, as opposed to what Barro (1999) finds, the impact of per capita GDP is insignificant. This is even more surprising considering that Beck et al. (2006) and Leeson and Dean (2009) both use per capita GDP as other explanatory variable in their models. As they do not include any variables controlling for education in their respective models, and as GDP can be correlated with education, it is possible that the significant per capita GDP coefficients they find are in fact capturing the effect of primary education. The second similarity is that in all models, the general level of primary education attained is always significant at the one percent level. The difference between genders in primary education attainment is also significant, but the coefficients are all positive. This implies that a greater difference has a positive impact on a country’s level of democracy. This is a very surprising result, which opposes the findings of Barro (1999). Overall, though, primary education appears to be a key determinant of democracy, confirming an already widely shared view.

Table 1: Subsample 1965-2005

	Model 1	Model 2	Model 3	Model 4	Model 5
n	90	90	90	89	89
Constant	-7.804 (-2.732e-009)	-7.895 (-2.766e-009)	-10.222 (-3.541e-009)	-11.465 (-3.889e-009)	-9.054 (-3.074e-009)
ρ	0.300*** (8.693)	0.310*** (9.662)	0.630*** (19.353)	0.400*** (10.543)	0.260*** (5.932)
β_1	-0.000009 (-0.289)	-0.000007 (-0.238)	-0.000015 (-0.482)	-0.000026 (-0.823)	-0.000034 (-1.096)
β_2	2.332*** (10.864)	2.347*** (11.007)	2.341*** (10.752)	2.796*** (13.087)	2.690*** (12.549)
β_3	1.914** (2.021)	1.839* (1.943)	1.895** (1.980)	1.608* (1.650)	1.871* (1.919)
R-squared	0.768	0.768	0.763	0.758	0.758
Corr-squared	0.252	0.247	0.248	0.263	0.252
Maximized log-likelihood	-2213.260	-2212.040	-2216.993	-2197.815	-2196.236
SFE significant	YES	YES	YES	YES	YES
Average SFE (in abs. value)	3.729	3.744	3.959	3.998	3.979

Notes: Model 1: Geographic connectivity matrix. Model 2: Geographic, weighted by population. Model 3: By trade partners. Model 4: By immigration, destination. Model 5: By immigration, origin. n is the number of countries, β_1 is coefficient for per capita GDP, β_2 is coefficient for primary schooling and β_3 is coefficient for difference in primary schooling between total population and females. SFE is spatial fixed effects. In parenthesis is the asymptotic t-statistic.* is significant at 10% level, ** is significant at 5% level and *** is significant at 1% level.

Table 2: Subsample 1995-2005

	Model 1	Model 2	Model 3	Model 4	Model 5
n	130	130	126	128	128
Constant	-3.456 (-9.171e-009)	-3.556 (-9.433e-009)	-5.163 (-1.340e-008)	-3.041 (-8.474e-009)	-3.304 (-9.201e-009)
ρ	0.020 (0.323)	-0.009 (-0.157)	0.430*** (5.921)	-0.060 (-0.810)	-0.016 (-0.228)
β_1	-0.000006 (-0.102)	-0.000005 (-0.090)	-0.000008 (-0.127)	-0.000002 (-0.031)	-0.000002 (-0.030)
β_2	1.538*** (3.766)	1.582*** (3.87)	1.430*** (3.344)	1.576*** (3.818)	1.562*** (3.774)
β_3	2.371** (2.322)	2.346** (2.298)	2.729*** (2.661)	2.327** (2.266)	2.304** (2.242)
R-squared	0.921	0.921	0.920	0.917	0.917
Corr-squared	0.047	0.047	0.055	0.046	0.046
Maximized log-likelihood	-785.932	-785.965	-762.596	-776.130	-776.268
SFE significant	YES	YES	YES	YES	YES
Average SFE (in abs value)	4.645	4.699	4.319	4.641	4.589

Notes: Model 1: Geographic connectivity matrix. Model 2: Geographic, weighted by population. Model 3: By trade partners. Model 4: By immigration, destination. Model 5: By immigration, origin. n is the number of countries, β_1 is coefficient for per capita GDP, β_2 is coefficient for primary schooling and β_3 is coefficient for difference in primary schooling between total population and females. SFE is spatial fixed effects. In parenthesis is the asymptotic t-statistic.* is significant at 10% level, ** is significant at 5% level and *** is significant at 1% level.

The third resemblance is that the spatial interdependence coefficient is always positive and significant for the first subsample. This suggests that democracy countries are indeed clustered, and that this clustering may be not only geographic. However, for the second subsample, the spatial coefficient is only significant in Model 3. This result suggests that, except when clusters are defined as trade partnerships, democratic countries have not been clustered during the 1995-2005 period. This may be related to the spatial fixed effects, which are always significant and more important in the second subsample than in the first one on average. It suggests that a country's own context has a greater influence on its democracy level in the past two decades, leaving less explanatory power to spatial interdependence. This result also confirms Leeson and Dean's (2009) findings. When they include year and country fixed effects, the spatial dependence coefficient the authors obtain decreases with time. This same coefficient increases when no fixed effects are included, corroborating the hypothesis that fixed effects have gained importance in the past years.

As explained earlier, evaluating the goodness-of-fit of spatial models is not as straightforward as in the non-spatial case. For this reason, I will not attempt to rank the five models here, but a few remarks can be pointed out of Tables 1 and 2. First of all, for each subsample, all models yield similar measures of goodness-of-fit. Therefore, the definition of space used to construct the connectivity matrix has no impact on the model's goodness-of-fit. According to the maximized log-likelihood, the models are better specified for the second subsample, but the correlation-squared values suggest the opposite. The difference between R-squared and correlation-squared is a good indicator of the significance of the spatial fixed effects. Confirming the hypothesis made above, the spatial fixed effects have a much stronger impact in the second subsample than in the first one.

For the reasons just mentioned, it is not possible to compare the different models with certainty. Interestingly, however, the models yield a wide range of values as spatial interdependence coefficients (ρ) for the first subsample. This is ruling out the hypothesis that the connectivity matrices are all similar. Models 1 and 2, using a geographic definition of neighbourhood, yield the lowest spatial dependence coefficients, around 0.30, and confirm Leeson and Dean's (2009) findings about the similarity of the results obtained with these two connectivity matrices. The coefficient for Model 5 is the lowest at 0.26. Model 4 yields a higher spatial coefficient of 0.40. The fact that both immigration models yield a significant spatial dependence coefficient confirms the hypothesis such that countries sharing common sociological traits influence each other's political decisions.

Model 3 is the one generating the highest spatial dependence coefficient, with 0.63. This implies that when taking political decisions concerning democracy, a country is strongly influenced by its main trading partners. According to the results obtained for the second subsample, it is also the only factor, among all those tested in this paper, that still had an influence in the past 15 years.

As mentioned earlier, Beck et al. (2006), using cross-sectional data from 1998, find that the spatial dependence is similar when using geographic distance or main trading partners to form the connectivity matrix, but the relationship is slightly stronger in the latter case. The gap between the results of Model 1 and Model 3 is very surprising. It is however important to remember that these models differ from the ones tested by Beck et al. (2006). At first sight, the spatial coefficients seem to imply that democratic countries do cluster a lot, especially through trade and immigration partnerships.

6. Interpretation of the results

In light of the results presented above, it is possible to draw several conclusions about the influence populations are subject to when making their democracy-related decisions. First of all, it is evident that countries' levels of democracy are connected to each other, but finding exactly how remains difficult. As their numerical values do not have a precise meaning, the spatial interdependence coefficients are hard to interpret. As the explained variable is the level of democracy, a higher coefficient implies that the average degree of democracy amongst a country's neighbours has a greater influence on its own level of democracy. This implies a stronger clustering of democratic countries.

The aim of this paper is precisely to shed light on this question, and some additional conclusions can be drawn. The most important one is that the spatial dependence between countries is not only geographic, confirming the new direction recently explored by political scientists. Political decisions are complex and definitely related between countries. Politicians discovered a long time ago the importance of the influence geographic regions and neighbours could have. Today, other variables like trade partnerships and immigration destinations are shown to also have some bearing on political decisions' diffusion.

The significant spatial fixed effects also yield interesting results, although not very surprising. As decisions related to the level of democracy are important, they usually take a long time to happen. Hence, most countries' level of democracy does not change drastically from one period to the other, making the countries' own context a significant explanatory variable. This "own context" has however become more influential during the 1995-2005 period. Combined with the smaller spatial dependence, this suggests a hypothesis of interest. From the data it appears as though political regimes fluctuated more frequently earlier in the 20th Century, and have been more stable in the past 15 years.

7. Alternative results

In order to obtain a different perspective on the spatial interdependence present in the data, and in order to be able to compare my results with previous findings, I run the five models again, but this time using the first difference of the level of democracy as a dependent variable. This new modeling is measuring another kind of interaction, namely how a *change* in the level of democracy of a particular country can affect the *change* in its neighbours' level of democracy. Instead of studying the clustering of democratic countries, these alternative models can be seen as explaining the spread of democracy as such. If all countries remain politically stable, not much spatial dependence will be observed.

The new results obtained are presented in Appendix 3. The first thing to be pointed out is that the alternative results, although different in magnitude, follow the same general trends. For the first subsample, spatial dependence coefficients are still significant (except for Model 4), but lower. For the second subsample, only Model 3 yields a significant spatial coefficient.

The primary education variables have different impacts on the spread of democracy and on the clustering of democratic countries. The new results show that during the 1995-2005 period, none of the education variables have a significant impact on the change of democracy. For the 1965-2005 period, the general average attainment is strongly significant, but the difference between genders is insignificant.

As a comparison, when including year and country fixed effects, Leeson and Dean (2009), who use only geographic connectivity matrices, obtain similar spatial coefficients. Differently specified, their model yields spatial coefficients of 0.046 and 0.014 for the subsamples covering the years 1951-2001 and 1991-2001 respectively. I obtain 0.09 for the 1965-2005 period, and the one for the 1995-2005 period is insignificant.

One main difference between the two sets of results is in regards to the spatial fixed effects. Each country's own context does not have any significant impact on its behaviour – change or stability – about democracy when the first difference is studied. This is also confirmed by the closeness of R-squared and correlation-squared. As it can be clearly observed that some countries are more stable than others, this probably means that the direction and the magnitude of the changes taken by the unstable ones are random.

8. Sensitivity analysis

In order to assess the accuracy of the results I obtain, I made slight changes to the data used for the different models examined[15]. As mentioned earlier, the geographic location of islands can be an issue when using geographic connectivity matrices in spatial econometrics. To avoid this possible bias, I run Models 1 and 2 a second time, excluding data from countries that are islands. Leeson and Dean (2009) perform the same robustness verification and obtain no notably different results when excluding or not the islands from the samples, as long as fixed effects are included. My results also remain generally consistent.

In order to avoid including observations that would have an extreme influence on other observations, I run Models 3 and 4 a second time eliminating data from China for Model 3 and from the United States for Model 4. As a very large proportion of the countries included in the sample have China as a main trading partner, China is assumed to have an important influence on the world's level of democracy. Running Model 3 again produces slightly smaller spatial dependence coefficient: 0.61 instead of 0.63 for the first subsample and 0.39 instead of 0.43 for the second one. This difference seems to imply that China does have an important influence on the democracy level of other countries. When China is not included in the regression, some dependence relationships are not considered, therefore the spatial interdependence is weaker. It is important to recall that the positive spatial coefficient does not mean that a trade partnership with

[15] The results obtained as part of the sensitivity analysis can be found in Appendix 4.

China increases a country's level of democracy. It is the strength of the relationship that is higher, and the positive sign means that the influence goes in the same direction as the neighbours.

Just like China is dominating trading partnerships, the United States is dominating the emigrants' destinations[16]. Running Model 4 without the United States data also yields smaller spatial dependence coefficients – 0.25 instead of 0.40, indicating the country really has an important influence on other countries' level of democracy. Even though those two countries play a big role in the clustering of democratic states, the spatial coefficients obtained when eliminating them are still significant. Therefore, democratic countries are still found to be clustered without them.

As mentioned earlier, ensuring the exogeneity of the connectivity matrix is crucial in spatial econometrics. This can be an issue for the models containing trade and immigration data in their connectivity matrix. A country can choose its trading partners according to their level of democracy. The trade partnerships at the end of the period – close to what I use initially – might then be the results of the behaviour of the level of democracy during the sample's period, which might create endogeneity.

In order to evaluate the impact this potential simultaneity bias has on the results, I compare the results obtained by Model 3 with two different connectivity matrices. The first one contains trade data from the beginning of the subsample's time period, and the second one contains the same measure but at the end of the period. As the Direction of Trade data is only available from 1980, the first subsample covers the 1980-2005 period. Both subsamples also contain a smaller number of countries. Obtaining similar results for

[16] For the second subsample, only 5 out of 125 countries do not have trade as a main trading partner, and 33 out 127 countries do not have the United States as main destination country.

the two different connectivity matrices would rule out the possibility of an important endogeneity bias.

For each subsample, the results obtained with the different connectivity matrices are very similar. The main difference is in the spatial dependence coefficient obtained for the second subsample which is higher – 0.19 compared to 0.16 – when trade partnerships of the end of the period are used. These results allow the conclusion that the main findings found earlier are not biased in an important way due to this particular form of endogeneity. However, they also suggest the existence of a bias in the second subsample. A similar sensitivity analysis is not possible for testing endogeneity of the connectivity matrices using immigration patterns due to unavailability of more complete data.

9. Conclusion

The goal of this paper was to study what determines a country's degree of democracy. More precisely, I was interested in the international side of these determinants. Starting from four mechanisms through which democracy is theorized to spread internationally – following ideas proposed by Simmons et al. (2006) – I attempted to discover if a country's geographic location, immigration patterns and trading partnerships have an impact on its level of democracy. In order to perform this empirical study, I used a spatial autoregressive model, with spatial fixed effects. Following the recent literature, I extended the concept of space beyond geography. Estimations were performed using Maximum likelihood methods. My models also contained other explanatory variables: per capita GDP, average primary education attainment, and the difference in primary education between genders.

The first result I obtained confirmed Barro's (1999) findings: primary education is a key determinant of democracy. Conditional on the two primary education variables, the other explanatory variable – per capita GDP – was not significant. I also concluded that there exists a spatial dependence in the clustering of democratic countries. Democracy seems to follow not only geographic borders, but also the path taken by immigrants and, to a greater extent, traded goods. Alternative models, explaining the *change* in the country's level of democracy, confirmed these conclusions. Democracy is spreading through time and space, following not only geographic channels: immigration and trade partnerships also play a part. These two phenomena – clustering of democratic states and spreading of democracy – were very strong during the second half of the 20th Century. In the past decade, countries seem to have followed a different path that still remains to be explained.

A first sensitivity analysis eliminating possibly problematic data allowed me to confirm these results. A second sensitivity analysis allowed me to conclude that there exists an endogeneity bias when trade partnerships are used to define neighbourhood, but that it is not important enough to invalidate the results obtained. The two steps were important to validate the robustness of the results originally obtained. As they both give positive results, it can be asserted that democratic countries cluster, not only on a geographic basis, but also through trade and immigration partnerships.

This paper however contains several weaknesses. Due to the low availability of software for spatial econometrics using panel data, I was limited in how I could model both the general relationship and the spatial dependence. These two steps are critical for obtaining valid results. Another potential problem concerns the endogeneity of the connectivity matrix. Although I find that this endogeneity bias does not invalidate the results for the model based on trade partnerships, I also find that this bias does exist. The results for the models based on immigration also need to be validated.

The two main weaknesses of this paper can be corrected in future research. Using a model different from the basic SAR and SEM can influence results. The appropriate Matlab functions can be developed starting from the ones for cross-sectional data included in LeSage’s econometrics toolbox. This toolbox also already contains the procedures for estimating a few spatial models using Generalized Method of Moments (GMM). Pinkse and Slade (2009) assert that endogeneity biases can usually be fixed using GMM estimation methods and an appropriate variance-covariance matrix estimator. They mention a non-parametric heteroscedasticity and autocorrelation consistent (HAC) estimator suggested by Kelejian and Prucha (2007). To my knowledge, this estimator is not used by the currently existing spatial econometrics software; it is thus something that could be improved in the future.

Overall, the results obtained in this paper have demonstrated that clustering of democratic countries occurs differently than suggested by previous studies. The diffusion of political decisions is influenced not only by geographic borders, but also by trade and immigration partnerships. Furthermore, with new tools available in the near future, empirical researchers will gain further insight into this topic. This could be a valuable aid for organizations concerned with democratization.

References

Anselin, Luc. 1988. *Spatial Econometrics: Methods and Models*. Boston: Kluwer Academic Publishers.

Anselin, Luc. 1999. Spatial Econometrics. School of Social Sciences, University of Texas at Dallas.

Anselin, Luc. 2007. Spatial Econometrics in RSUE: Retrospect and Prospect. *Regional Science and Urban Economics* 37: 450-56.

Anselin, Luc and Anil K. Bera. 1998. Spatial Dependence in Linear Regression Models with an Introduction to Spatial Econometrics. In *Handbook of Applied Economic Statistics*, A. Ullah and D.E. Giles (eds). New York: Marcel Dekker, 237-89. (to be checked)

Barro, Robert J. 1999. Determinants of Democracy. *The Journal of Political Economy* Vol. 107, No. 6: S158-83.

Beck, Nathaniel, Kristian Skrede Gleditsch, and Kyle Beardsley. 2006. Space Is More than Geography: Using Spatial Econometrics in the Study of Political Economy. *International Studies Quarterly* 50: 27-44.

Curry, Leslie. 1970. Univariate Spatial Forecasting. *Economic Geography* 46: 241-58.

Davidson, Russel and James G. MacKinnon. 2004. *Econometric Theory and Methods*. Oxford: Oxford University Press.

Elhorst, J. Paul. 2003. Specification and Estimation of Spatial Panel Data Models. International Regional Science Review 26, 3: 244-68.

Elhorst, J. Paul. 2010. Spatial Panel Data Models. In *Handbook of Applied Spatial Analysis*, edited by Fischer, M. M., and A. Getis. Berlin: Springer.

Franzese, Jr, Robert J., and Jude C. Hays. 2004. Empirical Modeling Strategies for Spatial Interdependence: Omitted-Variables vs. Simultaneity Bias. University of Michican, Ann Arbour.

Franzese, Jr, Robert J., and Jude C. Hays. 2008. Interdependence in Comparative Politics: Substance, Theory, Empirics, Substance. *Comparative Political Studies* 41: 742-80.

Gould, Peter. 1970. Is Statistix Inferens the Geographical Name for a Wild Goose? *Economic Geography* 14: 439-48.

Hordijk, Leen, and Jean Paelinck. 1976. Some principles and results in spatial econometrics. *Recherches Économiques de Louvain* 42: 175-97.

Isard, Walter. 1969. *General Theory: Social, Political, Economic, and Regional, With Particular Reference to Decision-Making Analysis*. Cambridge: M.I.T. Press.

Kelejian, Harry H., and Ingmar R. Prucha. 2007. HAC Estimation in a Spatial Framework. *Journal of Econometrics 140*: 131-154.

Leeson, Peter T., and Andrea M. Dean. 2009. The Democratic Domino Theory: An Empirical Investigation. *American Journal of Political Science* Vol. 53, No. 3: 533-51.

LeSage, James P., and R. Kelley Pace. 2009. *Introduction to Spatial Econometrics*. Boca Raton: Chapman & Hall/CRC

Moran, P. A. P. 1950. Notes on Continuous Stochastic Phenomena. *Biometrika* 37: 17-23.

Paelinck, Jean H. P. 2005. Spatial Econometrics: History, State-of-the-art and Challenges Ahead. Prepared for the Workshop on Spatial Econometrics, Kiel Institute for World Economics.

Parsons, Christopher R., Ronald Skeldon, Terrie L. Walmsley, and L. Alan Winters. 2005. Quantifying the International Bilateral Movements of Migrants. Issued by the Development Research Centre on Migration, Globalisation and Poverty.

Pinkse, Joris, and Margaret E. Slade. 2009. The Future of Spatial Econometrics. Prepared for the 50th anniversary symposium of the Journal of Regional Science held at the New York Federal Reserve.

Simmons, Beth A., Frank Dobbin, and Geoffrey Garrett. 2006. Introduction: The International Diffusion of Liberalism. *International Organization* 60 (Fall): 781-810.

Tocqueville, Alexis de. 1835. *De la Démocratie en Amérique*. Bruxelles: L. Hauman et Compe.

Ward, Michael D., John O'Loughlin, Michaeal Shin, Corey L. Lofdahl, Kristian S. Gleditsch, and Jordin S Cohen. 1997. The Spatial and Temporal Diffusion of Democracy, 1946-1994. Presented at the 37th Annual Conference of the International Studies Association, San Diego, California.

APPENDIX 1

Countries	Included in subsample 1 (1960-2005)	Included in subsample 2 (1991-2005)	GDP data from UN database	GDP data from World Bank database
Afghanistan	X	X	X	
Albania	X	X	X	
Algeria		X		X
Argentina	X	X		X
Armenia		X		X
Australia	X	X		X
Austria	X	X		X
Bahrain		X		X
Bangladesh		X		X
Belgium	X	X		X
Benin	X	X		X
Bolivia	X	X		X
Botswana		X		X
Brazil	X	X		X
Bulgaria	X	X	X	
Burundi		X		X
Cambodia	X	X	X	
Cameroon	X	X		X
Canada	X	X		X
Central African Republic	X	X		X
Chile	X	X		X
China	X	X		X
Colombia	X	X		X
Congo, Brazzaville	X	X		X
Congo, Kinshasa	X	X		X
Costa Rica	X	X		X
Croatia		X		X
Cuba	X	X	X	
Cyprus	X	X	X	
Denmark	X	X	X	
Dominican Republic	X	X	X	
Ecuador	X	X	X	
Egypt	X	X	X	
El Salvador	X	X	X	
Fiji		X	X	
Finland	X	X		X
France	X	X		X
Gabon	X	X		X
Gambia		X		X
Germany[17]	X	X		X
Ghana	X	X		X
Greece	X	X		X
Guatemala	X	X		X
Guyana		X		X

[17] Before 1991, data from West Germany is used.

Haiti	X	X	X	
Honduras	X	X		X
Hungary	X	X	X	
India	X	X		X
Indonesia	X	X		X
Iran	X	X		X
Iraq	X	X		X
Ireland	X	X		X
Israel	X	X		X
Italy	X	X		X
Ivory Coast	X	X		X
Jamaica	X	X		X
Japan	X	X		X
Jordan	X	X		X
Kazakhstan		X		X
Kenya		X		X
Korea, South	X	X		X
Kuwait		X		X
Kyrgyzstan		X		X
Laos	X	X	X	
Latvia		X		X
Lesotho		X		X
Liberia	X	X		X
Libya	X	X	X	
Lithuania		X		X
Malawi		X		X
Malaysia	X	X		X
Mali	X	X		X
Mauritania	X	X		X
Mauritius		X		X
Mexico	X	X		X
Moldova		X		X
Mongolia	X	X	X	
Morocco	X	X		X
Mozambique		X		X
Myanmar	X	X	X	
Namibia		X		X
Nepal	X	X		X
Netherlands	X	X		X
New Zealand	X	X		X
Nicaragua	X	X		X
Niger	X	X		X
Norway	X	X		X
Pakistan	X	X		X
Panama	X	X		X
Papua New Guinea		X		X
Paraguay	X	X		X
Peru	X	X		X
Philippines	X	X		X
Poland	X	X	X	
Portugal	X	X		X
Qatar		X		X
Romania	X	X	X	
Russia		X	X	
Rwanda		X		X
Saudi Arabia	X	X	X	

Senegal	X	X	X
Sierra Leone		X	X
Singapore	X	X	X
Slovenia		X	X
South Africa	X	X	X
Spain	X	X	X
Sri Lanka	X	X	X
Sudan	X	X	X
Swaziland		X	X
Sweden	X	X	X
Switzerland	X	X	X
Syria	X	X	X
Tajikistan		X	X
Tanzania		X	X
Thailand	X	X	X
Togo		X	X
Trinidad and Tobago		X	X
Tunisia	X	X	X
Turkey	X	X	X
Uganda		X	X
Ukraine		X	X
United Arab Emirates		X	X
United Kingdom	X	X	X
United States	X	X	X
Uruguay	X	X	X
Venezuela	X	X	X
Vietnam		X	X
Yemen		X	X
Zambia		X	X
Zimbabwe		X	X

APPENDIX 2

Countries	Neighbours
Afghanistan	China, India, Iran, Pakistan, Tajikistan
Albania	Greece
Algeria	Libya, Mali, Mauritania, Morocco, Niger, Tunisia
Argentina	Bolivia, Brazil, Chile, Paraguay, Peru, Uruguay
Armenia	Azerbaijan, Iran, Turkey
Australia	-
Austria	Germany, Hungary, Italy, Slovenia, Switzerland
Bahrain	-
Bangladesh	India, Myanmar
Belgium	France, Germany, Netherlands
Benin	Niger, Togo
Bolivia	Argentina, Brazil, Chile, Paraguay, Peru
Botswana	Namibia, South Africa, Zambia, Zimbabwe
Brazil	Argentina, Bolivia, Colombia, Guyana, Paraguay, Peru, Uruguay, Venezuela
Bulgaria	Greece, Romania, Turkey
Burundi	Congo (Kinshasa), Rwanda, Tanzania
Cambodia	Laos, Thailand, Vietnam
Cameroon	Central African Republic, Congo (Brazzaville), Gabon
Canada	United States
Central African Republic	Cameroon, Congo (Kinshasa), Congo (Brazzaville), Sudan
Chile	Argentina, Bolivia, Peru
China	Afghanistan, India, Kazakhstan, Kyrgyzstan, Laos, Mongolia, Myanmar, Nepal, Russia, Tajikistan, Vietnam
Colombia	Brazil, Ecuador, Panama, Peru, Venezuela
Congo, Brazzaville	Angola, Cameroon, Central African Republic, Congo (Kinshasa), Gabon
Congo, Kinshasa	Angola, Burundi, Central African Republic, Congo (Brazzaville), Rwanda, Sudan, Tanzania, Uganda
Costa Rica	Nicaragua, Panama
Croatia	Hungary, Slovenia
Cuba	-
Cyprus	-
Denmark	Germany
Dominican Republic	Haiti
Ecuador	Colombia, Peru
Egypt	Israel, Libya, Sudan
El Salvador	Guatemala, Honduras
Fiji	-
Finland	Norway, Russia, Sweden
France	Belgium, Germany, Italy, Spain, Switzerland
Gabon	Cameroon, Congo (Brazzaville),
Gambia	Senegal
Germany	Austria, Belgium, Denmark, France, Netherlands, Poland, Switzerland
Ghana	Ivory Coast, Togo
Greece	Albania, Bulgaria, Turkey
Guatemala	El Salvador, Honduras, Mexico
Guyana	Brazil, Venezuela

Haiti	Dominican Republic
Honduras	El Salvador, Guatemala, Nicaragua
Hungary	Austria, Croatia, Romania, Slovenia, Ukraine
India	Afghanistan, Bangladesh, China, Myanmar, Nepal, Pakistan
Indonesia	Malaysia, Papua New Guinea
Iran	Afghanistan, Armenia, Azerbaijan, Iraq, Pakistan, Turkey
Iraq	Iran, Jordan, Kuwait, Saudi Arabia, Syria, Turkey
Ireland	United Kingdom
Israel	Egypt, Jordan, Syria
Italy	Austria, France, Slovenia, Switzerland
Ivory Coast	Ghana, Liberia, Mali
Jamaica	-
Japan	-
Jordan	Iraq, Israel, Saudi Arabia, Syria
Kazakhstan	China, Kyrgyzstan, Russia
Kenya	Sudan, Tanzania, Uganda
Korea, South	-
Kuwait	Iraq, Saudi Arabia
Kyrgyzstan	China, Kazakhstan, Tajikistan
Laos	Cambodia, China, Myanmar, Thailand, Vietnam
Latvia	Lithuania, Russia
Lesotho	South Africa
Liberia	Ivory Coast, Sierra Leone
Libya	Algeria, Egypt, Niger, Sudan, Tunisia
Lithuania	Latvia, Poland, Russia
Malawi	Mozambique, Tanzania, Zambia
Malaysia	Indonesia, Thailand
Mali	Algeria, Ivory Coast, Mauritania, Niger, Senegal
Mauritania	Algeria, Mali, Morocco, Senegal
Mauritius	-
Mexico	Guatemala, United States
Moldova	Romania, Ukraine
Mongolia	China, Russia
Morocco	Algeria, Mauritania
Mozambique	Malawi, South Africa, Swaziland, Tanzania, Zambia, Zimbabwe
Myanmar	Bangladesh, China, India, Laos, Thailand
Namibia	Angola, Botswana, South Africa, Zambia, Zimbabwe
Nepal	China, India
Netherlands	Belgium, Germany
New Zealand	-
Nicaragua	Costa Rica, Honduras
Niger	Algeria, Benin, Libya, Mali
Norway	Finland, Russia, Sweden
Pakistan	Afghanistan, India, Iran
Panama	Colombia, Costa Rica
Papua New Guinea	Indonesia
Paraguay	Argentina, Bolivia, Brazil
Peru	Bolivia, Brazil, Chile, Colombia, Ecuador
Philippines	-
Poland	Germany, Lithuania, Russia, Ukraine

Portugal	Spain
Qatar	Saudi Arabia
Romania	Bulgaria, Hungary, Moldova, Ukraine
Russia	Azerbaijan, China, Finland, Kazakhstan, Latvia, Lithuania, Mongolia, Norway, Poland, Ukraine
Rwanda	Burundi, Congo (Kinshasa), Tanzania, Uganda
Saudi Arabia	Iraq, Jordan, Kuwait, Qatar, United Arab Emirates, Yemen
Senegal	Gambia, Mali, Mauritania
Sierra Leone	Liberia
Singapore	-
Slovenia	Austria, Croatia, Hungary, Italy
South Africa	Botswana, Lesotho, Mozambique, Namibia, Swaziland, Zimbabwe
Spain	France, Portugal
Sri Lanka	-
Sudan	Central African Republic, Congo (Kinshasa), Egypt, Kenya, Libya, Uganda
Swaziland	Mozambique, South Africa
Sweden	Finland, Norway
Switzerland	Austria, France, Germany, Italy
Syria	Iraq, Israel, Jordan, Turkey
Tajikistan	Afghanistan, China, Kyrgyzstan
Tanzania	Burundi, Congo (Kinshasa), Kenya, Malawi, Mozambique, Rwanda, Uganda, Zambia
Thailand	Cambodia, Laos, Malaysia, Myanmar
Togo	Benin, Ghana
Trinidad and Tobago	-
Tunisia	Algeria, Libya
Turkey	Armenia, Bulgaria, Greece, Iran, Iraq, Syria
Uganda	Congo (Kinshasa), Kenya, Rwanda, Sudan, Tanzania
Ukraine	Hungary, Moldova, Poland, Romania, Russia
United Arab Emirates	Saudi Arabia
United Kingdom	Ireland
United States	Canada, Mexico
Uruguay	Argentina, Brazil
Venezuela	Brazil, Colombia, Guyana
Vietnam	Cambodia, China, Laos
Yemen	Saudi Arabia
Zambia	Angola, Botswana, Congo (Kinshasa), Malawi, Mozambique, Namibia, Tanzania, Zimbabwe
Zimbabwe	Botswana, Mozambique, Namibia, South Africa, Zambia

APPENDIX 3

This appendix presents the alternative results obtained with the first difference models.

Table 3: Subsample 1965-2005

	Model 1	Model 2	Model 3	Model 4	Model 5
n	90	90	90	89	89
Constant	-2.437 (-8.428e-010)	-2.418 (-8.363e-010)	-2.391 (-8.295e-010)	-2.675 (-9.249e-010)	-2.455 (-8.517e-010)
ρ	0.090** (2.273)	0.090** (2.418)	0.290*** (5.317)	0.080 (1.634)	0.130*** (2.68)
β_1	-0.000004 (-0.135)	-0.000004 (-0.134)	-0.000005 (-0.172)	-0.000005 (-0.148)	-0.000007 (-0.231)
β_2	0.715*** (3.590)	0.709*** (3.560)	0.698*** (3.516)	0.774*** (3.915)	0.714*** (3.617)
β_3	1.094 (1.142)	1.118 (1.167)	1.121 (1.174)	1.198 (1.253)	1.124 (1.180)
R-squared	0.070	0.070	0.075	0.066	0.072
Corr-squared	0.019	0.019	0.021	0.021	0.020
Maximized log-likelihood	-2211.904	-2211.670	-2209.786	-2178.911	-2176.539
SFE significant	NO	NO	NO	NO	NO
Average SFE (in abs. value)	1.154	1.143	1.124	1.220	1.141

Notes: Model 1: Geographic connectivity matrix. Model 2: Geographic, weighted by population. Model 3: By trade partners. Model 4: By immigration, destination. Model 5: By immigration, origin. n is the number of countries, β_1 is coefficient for per capita GDP, β_2 is coefficient for primary schooling and β_3 is coefficient for difference in primary schooling between total population and females. SFE is spatial fixed effects. In parenthesis is the asymptotic t-statistic.* is significant at 10% level, ** is significant at 5% level and *** is significant at 1% level.

Table 4: Subsample 1995-2005

	Model 1	Model 2	Model 3	Model 4	Model 5
n	130	130	126	128	128
Constant	7.871 (8.245e-009)	7.852 (8.226e-009)	0.929 (1.073e-009)	8.560 (9.776e-009)	7.861 (8.988e-009)
ρ	0.030 (0.485)	0.040 (0.693)	0.004 (0.038)	-0.030 (-0.406)	0.070 (1.022)
β_1	0.000006 (0.105)	0.000007 (0.119)	-0.000099* (-1.873)	0.000002 (0.026)	0.000004 (0.054)
β_2	-1.860*** (-2.70)	-1.859*** (-2.70)	0.270 (0.436)	-1.990 (-2.864)	-1.850*** (-2.672)
β_3	4.135** (2.481)	4.135** (2.481)	-2.610* (-1.759)	4.132 (2.462)	4.133** (2.466)
R-squared	0.268	0.268	0.405	0.268	0.270
Corr-squared	0.033	0.033	0.017	0.034	0.033
Maximized log-likelihood	-948.548	-948.511	-874.32	-936.431	-936.008
SFE significant	NO	NO	YES	NO	NO
Average SFE (in abs. value)	2.841	2.840	1.766	2.987	2.826

Notes: Model 1: Geographic connectivity matrix. Model 2: Geographic, weighted by population. Model 3: By trade partners. Model 4: By immigration, destination. Model 5: By immigration, origin. n is the number of countries, β_1 is coefficient for per capita GDP, β_2 is coefficient for primary schooling and β_3 is coefficient for difference in primary schooling between total population and females. SFE is spatial fixed effects. In parenthesis is the asymptotic t-statistic.* is significant at 10% level, ** is significant at 5% level and *** is significant at 1% level.

APPENDIX 4

This appendix presents the results obtained in the sensitivity analysis. Tables 5 and 6 contain the results obtained when the possibly distorting observations are removed. Table 7 contains the results obtained to test the possible impacts of an endogeneity bias.

Table 5: Subsample 1965-2005

	Model 1 (with no islands)	Model 2 (with no islands)	Model 3 (with no China)	Model 4 (with no U.S.)
n	81	81	89	88
Constant	-8.362 (-3.325e-009)	-8.619 (-3.424e-009)	-11.294 (-3.868e-009)	-10.267 (-3.588e-009)
ρ	0.280*** (7.949)	0.270*** (8.136)	0.610*** (18.949)	0.250*** (5.720)
β_1	-0.000001 (-0.029)	0.000001 (0.034)	-0.000012 (-0.40)	-0.000014 (-0.446)
β_2	2.472*** (10.772)	2.545*** (11.138)	2.344*** (11.141)	2.794*** (12.877)
β_3	1.811 (1.785)	1.702* (1.676)	1.835* (1.908)	1.647* (1.665)
R-squared	0.756	0.755	0.761	0.750
Corr-squared	0.264	0.258	0.248	0.260
Maximized log-likelihood	-2011.174	-2009.797	-2193.754	-2182.583
SFE significant	YES	YES	YES	YES
Average SFE (in abs. value)	3.699	3.732	3.892	3.996

Notes: Model 1: geographic connectivity matrix. Model 2: geographic, weighted by population. Model 3: by trade partners. Model 4: by immigration, destination. n is the number of countries, β_1 is coefficient for per capita GDP, β_2 is coefficient for primary schooling and β_3 is coefficient for difference in primary schooling between total population and females. SFE is spatial fixed effects. In parenthesis is the asymptotic t-statistic.* is significant at 10% level, ** is significant at 5% level and *** is significant at 1% level.

Table 6: Subsample 1995-2005

	Model 1 (with no islands)	Model 2 (with no islands)	Model 3 (with no China)	Model 4 (with no U.S.)
n	117	117	125	127
Constant	-3.656 (-1.092e-008)	-3.756 (-1.122e-008)	-5.686 (-1.460e-008)	-3.182 (-9.166e-009)
ρ	0.020 (0.323)	-0.010 (-0.173)	0.390*** (5.440)	-0.040 (-0.555)
β_1	-0.000004 (-0.061)	-0.000003 (-0.046)	-0.000007*** (-0.124)	-0.000002 (-0.025)
β_2	1.582*** (3.557)	1.630*** (3.664)	1.441*** (3.375)	1.562*** (3.769)
β_3	2.416** (2.231)	2.393** (2.210)	2.820 (2.727)	2.334** (2.263)
R-squared	0.911	0.911	0.919	0.916
Corr-squared	0.046	0.046	0.051	0.046
Maximized log-likelihood	-725.439	-725.458	-757.834	-771.613
SFE significant	YES	YES	YES	YES
Average SFE (in abs. value)	4.560	4.622	4.295	4.631

Notes: Model 1: geographic connectivity matrix. Model 2: geographic, weighted by population. Model 3: by trade partners. Model 4: by immigration, destination. n is the number of countries, β_1 is coefficient for per capita GDP, β_2 is coefficient for primary schooling and β_3 is coefficient for difference in primary schooling between total population and females. SFE is spatial fixed effects. In parenthesis is the asymptotic t-statistic.* is significant at 10% level, ** is significant at 5% level and *** is significant at 1% level.

Table 7: W by trade

	W begin 1980	W end 2005	W begin 1995	W end 2005
n	58	58	101	
Constant	-14.427	-11.663	-5.197	-4.810
	(-6.211e-009)	(-5.091e-009)	(-1.245e-008)	(-1.154e-008)
ρ	0.610***	0.610***	0.160*	0.190*
	(13.655)	(12.803)	(1.866)	(2.087)
β_1	-0.000005	0.000001	0.000003	0.000001
	(-0.147)	(0.028)	(0.043)	(0.015)
β_2	3.185***	2.728***	1.704***	1.646***
	(8.521)	(7.265)	(3.679)	(3.528)
β_3	2.968*	2.957*	3.067***	3.080***
	(1.778)	(1.796)	(2.810)	(2.823)
R-squared	0.817	0.822	0.913	0.913
Corr-squared	0.224	0.225	0.064	0.067
Maximized log-likelihood	-869.248	-864.395	-616.304	-615.960
SFE significant	YES	YES	YES	YES
Average SFE (in abs. value)	4.11	3.97	4.22	4.23

Notes: n is the number of countries, β_1 is coefficient for per capita GDP, β_2 is coefficient for primary schooling and β_3 is coefficient for difference in primary schooling between total population and females. SFE is spatial fixed effects. In parenthesis is the asymptotic t-statistic.* is significant at 10% level, ** is significant at 5% level and *** is significant at 1% level.