Spatial effects in website adoption in the European Regions

Margarita Billón Currás
Departamento de Estructura Económica y Economía del Desarrollo
Facultad de Ciencias Económicas y Empresariales
Universidad Autónoma de Madrid
Campus de Cantoblanco
Carretera de Colmenar Viejo, Km. 15
Madrid 28049
Teléfono: 914978777
Email: margarita.billon@uam.es

Roberto Ezcurra Orayen
Departamento de Economía
Universidad Pública de Navarra
Campus Arrosadía s/n
Pamplona 31006
Teléfono: 948169359
Fax: 948169721
Email: roberto.ezcurra@unavarra.es

Fernando Lera López
Departamento de Economía
Universidad Pública de Navarra
Campus Arrosadía s/n
Pamplona 31006
Teléfono: 948169351
Fax: 948169721
Email: lera@unavarra.es

Abstract

The purpose of this paper is to provide empirical evidence on the neighbouring effects of Internet adoption as measured by the percentage of firms with own website in the European regions. To this end, we apply a set of instruments and techniques commonly used in the spatial econometrics framework to test the hypothesis that proximity matters when explaining internet adoption by firms. In fact, this is the first study that analyzes explicitly the role played by the geographical space in which each region is located when explaining website adoption. We have found that firms in physically adjacent regions register a similar degree of Internet adoption showing positive spatial dependence. Nevertheless, the spatial effects detected are mainly constrained by the national borders. Results also show that GDP per capita, population density, sectoral composition and education are positively related to geographical distribution of internet adoption in the enlarged EU. We have also found that regional disparities in Internet adoption are greater than territorial inequalities in GDP per capita.

Clasificación Código JEL: O18, O33, O52, C21

Introduction

The diffusion of Information and Communication Technologies (ICT) opens up possible new benefits for regional development of the European regions "by favouring the creation and growth of poles of excellence in ICT activities and developing connectivity and networking among enterprises", as the European Union has highlighted (Council of the European Union, 2006: 21). ICT boost economic activity by increasing information transmission at a higher speed and volume, allowing lower transaction and communication costs, improving access to information and knowledge from remote areas and facilitating distance working and coordination of activities between partners previously too distant.

These effects have led to a debate within the academic literature on whether the Internet would mean "the death of distance" (Cairncross, 1997). The question can be raised as follows: Would Internet diffusion mean a change in the role played by geographic closeness and therefore by agglomerations? Or on the contrary, and despite the Internet, would geographic location and proximity still maintain its significance in explaining regional disparities? This paper aims to contribute to this debate by offering empirical evidence on the regional distribution of the Internet in the enlarged EU and on the role played within this context by geographic factors.

Despite its relevance, at the European level studies analysing regional disparities in Internet adoption are still scarce (Milicevic and Gareis, 2003; Demunter, 2005; ESPON, 2006a). This is due to the limited availability of comparable statistical data across regions and to the lack of instruments for systematically quantifying the digital divide. In recent years, various European projects (for example, BISER, 2004 and UNDERSTAND, 2006) have been put into action to analyse Internet adoption at regional level. Unfortunately, neither of them analyses all the European regions nor includes the regions of new member countries. The BISER project considers 28 regions, while UNDERSTAND focuses on the analysis of nine different regions within the context of the EU-15.

In any case, available data about Internet adoption by firms in the European regions conclude that even though Internet adoption is a reality in a great proportion of European firms, regional differences are still significant. For example, according to BISER (2004), Internet use varies from a maximum of 95% in Central Finland to a minimum of 64% in Brittany. The number of firms that have their own website varies from a minimum of 32.7% in Nord/Pas de Calais to a maximum of 73.5% in Salzburg (BISER, 2004).

In general, evidence shows that the spatial pattern of Internet diffusion in the European regions can be largely attributed to the general level of development and to the regional structure in terms of sectoral composition (BISER, 2004; UNDERSTAND, 2006). In addition, while in some countries the spatial distribution tends to be more homogeneous (e.g. Italy, UK), other countries (France, Spain) show remarkable territorial differences (ESPON, 2006a). In any case, little attention has been paid to the study of the spatial characteristics of the digital divide in the European regions.

This paper endeavors to fill this gap by examining the regional distribution of Internet adoption by firms in the enlarged EU, in a quest for empirically well-founded, stylized facts. The main research question is whether Internet adoption by firms, measured by companies with their own website across the European regions, follows a spatial pattern and if geographic proximity still matters. The second question is to what extent Internet adoption by firms at regional level is related to regional variables such as GDP per capita, population density, education level and sectoral composition.

In comparison with the previous literature this work a) offers an analysis on the role played by geographic location in explaining the spatial distribution of Internet adoption; b) covers 239 regions, which is a geographic setting considerably wider than that considered in the scant studies which have previously examined this issue and; c) uses a methodological approach not employed earlier in this literature, based on a set of various spatial econometric techniques (Haining, 1990; Bailey and Gatrell, 1995).

Following on from this introduction, the remainder of this paper is organized as follows. Section two is devoted to a review of the relevant findings obtained from the previous literature regarding the relationships between Internet adoption and geographic proximity and agglomerations. Sections three and four focus on the empirical study of the spatial distribution of website adoption in the European regions. The final section offers the main conclusions and policy implications from our work.

Literature review

The spatial distribution of economic activity can be explained as the result of the tug-of-war between forces promoting geographic concentration of income and productive factors (centripetal forces) and those that tend to oppose it (centrifugal forces) (Fujita *et al.*, 1999). Centripetal forces are associated mainly with the important role played in this context by agglomeration economies and increasing returns to scale. By contrast, centrifugal forces can be explained by the presence of immobile productive factors, non-tradeable goods for consumption, or the increasing relevance of congestion costs and other pure diseconomies linked to excessive agglomeration.

The importance of the previous factors varies according to the theory dealing with space and agglomeration. The traditional and new trade theories explain the differences in spatial location by the development of the comparative advantage. New economic geography models (NEG), initiated by Krugman (1991) in the early 1990s, emphasize cost and demand linkages and the importance of market size as the key agglomeration forces (Ottaviano and Puga, 1998). Unlike most NEG models, other theories, such as urban and spatial economics and diffusion theory, focus their attention on the role played by technology externalities and knowledge spillovers (Fujita and Mori, 2005).

Internet and agglomerations

One would expect that centripetal and centrifugal forces influence firms' location decisions, given that some of the benefits of these processes depend on geographic proximity (Kolko, 2002). However, existing theories, as well as empirical evidence, do not offer conclusive results regarding the links between Internet adoption, concentration and deconcentration activities and proximity. On the one hand, it is argued that the Internet, due to its numerous advantages (e.g. decreasing coordination costs and allowing remote coordination), would reduce the role played by agglomerations. On the other hand, it is claimed that agglomerations and ICT, and particularly the Internet, are complementary and not substitutes, and that geographically-limited "neighborhood" remains relevant.

Following Forman (2003; 2005) and Forman *et al.* (2005a; 2005b) we can distinguish different theoretical approaches to explain Internet adoption by firms. According to global village theory, firms in areas with low population density, such as rural areas, will benefit more from Internet technology due to the reduction of coordination costs within and between firms, among other factors. Consistent with this view, Internet adoption and urban agglomeration would be substitutes.

Following urban density theory, adoption costs will decrease when population size and density increase. Access to ICT infrastructure, skilled labor or knowledge spillovers would favor and complement Internet adoption in urban agglomerations, Internet and agglomerations being complementary. Finally, according to industry composition theory, the location of information and high-tech intensive industries would boost Internet adoption in urban areas.

The economic geography literature has studied the spatial distribution of Internet adoption showing the emergence of the so-called Internet geography. Zook (2006) points out that the Internet gives rise to the creation of a new communication space where the new "geographies of the Internet" (technical, human, political and cultural) emerge. Empirical evidence demonstrates that Internet distribution follows an uneven spatial pattern and that new examples of agglomeration are arising (Giovannetti *et al.*, 2003; Weltevreden *et al.*, 2005; Zook, 2000; 2002; 2003; 2006).

Some studies demonstrate that different indicators of Internet activity, such as bandwidth or domain names, are mainly concentrated in urban areas (Malecki, 2002; Kolko, 2000; Sternberg and Krymalowski, 2002). Zook (2000) finds that the leading centres for Internet content in the US exhibit a high degree of clustering. They seem to be related to existing industrial sectors (information-intensive industries) showing specialization in commercial domain names, located in close cities.

There is also empirical evidence on geographic proximity's role in explaining ICT adoption. Galliano (2005) finds that the intensity of ICT use by French manufacturing firms varies according to the firm's location (urban versus rural). Bell and Song (2004) empirically demonstrate the significance of neighborhood effects in the space-time pattern of trial for an Internet retailing service. Billon *et al.* (2007) find positive spatial dependence when explaining Internet adoption by households in European regions, adjacent regions registering a similar degree of penetration.

On the other hand, other authors demonstrate that ICT use lowers the level of industry concentration while boosting convergence, showing that industries and firms are more evenly geographically distributed (Kolko, 2002). Kolko (2000) measuring Internet domain registration demonstrates a positive relationship between city size and Internet adoption. However, the proximity of a city to other cities is negatively related to Internet usage, which would support the death of distance hypothesis. In the same vein, Forman (2003; 2005) and Forman *et al.* in several studies (2005a; 2005b) on the United States find that the Internet has been adopted across many industries which do not show similar geographic distributions. These authors also find that the geographic distribution of the Internet can vary depending on type of Internet application. For the basic use of the Internet for business, such as email or browsing, global village theory is the one that best explains Internet adoption, given that at this level population density or city size are not so relevant. On the other hand, more advanced applications devoted to improve and develop computing processes such as electronic commerce (Forman *et al.*, 2005a), are explained by urban density theory especially in the case of intrafirm technologies.

In short, there is empirical evidence on both concentration and deconcentration activities related to Internet adoption by firms. However, many factors explaining geographic differences, and particularly concentrations, are closely related to the perceived advantages of urban agglomerations. In fact, some authors argue that Internet concentration is mainly registered in those areas where there were agglomerations prior to Internet diffusion. The prior distribution of industries across different geographic locations (Forman, 2003), the presence of ICT-intensive firms and high-tech and knowledge-intensive industries (Zook, 2000; Forman and Goldfarb, 2005; Kolko, 2000), the existence of complementary factors such as the availability of skilled workforce or intermediate inputs (Zook, 2000; Kolko, 2000; 2002) or the type of Internet applications (Forman *et al.*, 2005a; Forman, 2005) are factors contributing to Internet adoption in urban areas.

Internet and diffusion theory

Diffusion theory has been the methodological approach most frequently employed to explain the diffusion rates of new technologies. Some authors analyse information spillovers from users to non-users, which are the core of the "epidemic models" (Karshenas and Stoneman, 1995; Geroski, 2000; Galliano, 2005). According to this view, the greater the number of adopters, the higher the probability of a non-adopter adopting the Internet and being "contaminated" (Greenstein and Prince, 2006). With regard to consumers and according to heterogeneity models (Rosenberg, 1972; Rogers, 1995) some authors attribute existing differences in diffusion rates of the Internet to differences in the observed economic and demographic characteristics of adopters (such as income, location, employment, education, family structure, age, gender and attitude towards technology).

With regard to firms, Internet diffusion within and across firms can be explained by a wide variety of internal and external factors extensively studied in the innovation diffusion literature (e.g. Forman and Goldfarb, 2005; Del Aguila and Padilla, 2006). Internal variables commonly considered are some CEO and firm structural characteristics (size, multinational ownership), perceived benefits from ICT and their characteristics, the role played by human capital, firm internal organization or competitive strategies. External influence considers the environment within which the firm performs its activities: competitive pressure, sectoral specialization, geographic location and government policies.

Within the framework of diffusion theory, according to the knowledge spillovers literature (Fujita and Mori, 2005, Audretsch and Feldman, 2004) applied to ICT diffusion, the key point about Internet diffusion lies in the type of knowledge transmitted. If the knowledge transmitted can be codified and transmitted through ICT and the Internet, then geographic proximity will reduce its role in explaining Internet diffusion. On the other hand, if the knowledge to be transmitted is tacit or non-codified, then geographic proximity will matter and, therefore, closeness between agents will remain significant (Giovannetti *et al.*, 2003). From this latter perspective, face-to-face communication would help to transmit that uncodified knowledge which is not transmittable over distance. The use of the Internet might occur faster in those geographic areas, such as populated areas or cities, where the density of sources of knowledge about such technology is higher and, therefore, similar to what happens with other innovations (Kaufmann *et al.,* 2003; Van Oort *et al.,* 2004) where there is geographic concentration of knowledge spillovers. This underlines the role played in Internet diffusion by both relational proximity (Kaufman *et al.,* 2003) and geographic proximity (Leamer and Storper, 2001; Giovannetti *et al.,* 2003).

In short, we can find mixed empirical evidence on spatial patterns in Internet adoption and the role played by agglomerations and proximity. Within this framework, our hypothesis highlights the relevance of spatial effects, the role of geographic proximity and the existence of clusters when explaining Internet adoption by firms across the European regions.

The spatial distribution of the Internet in EU firms

Data

There is a lack of comparable data for studying the spatial distribution of Internet adoption in the European regions. In fact, there is only one relevant study covering from a regional perspective nearly the whole European Union: the ESPON Project 1.2.3 (ESPON, 2006a). Nevertheless, this uses a synthetic territorial index, including skills for ICT use, ICT adoption by households and business, impact on the labor market and ICT patents, and includes only NUTS-2 data for 12 countries out of the ESPON space (29 countries). The previously mentioned BISER and UNDERSTAND projects only consider 28 and nine regions, respectively.

Bearing in mind these limitations, we decided to use an alternative dataset, the ESPON Project Indicators (ESPON, 2006b). Particularly, we have selected the only available indicator associated with Internet adoption within firms at regional level: the proportion of firms with their own website in 2002.

The percentage of firms that have a website is one of the basic indicators for Internet adoption by European firms (European Commission, 2005). It has also been analysed in the BISER and UNDERSTAND projects to measure Internet adoption at regional level. The website adoption indicator is calculated on the basis of data provided by EUROSTAT and National Statistical Offices in 2002. The data cover the NUTS-2(1) level regions in the EU-25 with the exception of Cyprus, Denmark, Luxembourg, Malta, Lithuania, Latvia, Estonia, Slovenia, and the French overseas departments. In this respect, it should be noted that our study covers 239 EU-25 regions, which is a considerably wider geographic setting than that considered in the few previous studies.

Exploratory spatial data analysis

We begin our empirical analysis by examining the spatial distribution of Internet adoption by EU firms measured by website adoption. As can be observed in Figure 1, there are relatively significant disparities in the proportion of firms with their own website across the European regions. Specifically, the regions with the highest share of firms with their own

---

[1] NUTS is the French acronym for "Nomenclature of Territorial Units for Statistics", a hierarchical classification of subnational spatial units established by EUROSTAT. In this classification, NUTS-0 corresponds to country level, while increasing numbers indicate increasing levels of subnational disaggregation.

website tend to be located in some of the most developed EU areas. This is the case, for instance, for Inner London, Brussels, Hamburg or Île de France. By contrast, the lowest values of the study variable are concentrated in the Southern and Eastern periphery of the Union. When assessing the implications to be drawn from Figure 1, it is important to note that this overall picture highlights that the enlargement process of the EU towards Central and Eastern Europe has introduced new spatial patterns in the North-South digital divide identified by the previous literature (European Commission, 2005; ESPON, 2006a). Leaving this aside for the moment, it is worth mentioning that this initial examination suggests that the proportion of firms with their own website is not randomly distributed across space. On the contrary, there seems to be a positive spatial relationship between adjacent areas, neighboring regions registering a similar share of firms using the Internet.

[SEE FIGURE 1]

Some caution is recommended when interpreting the data shown in Figure 1, since the conclusions that might be drawn are highly sensitive to the number and width of the different intervals used to represent the variable of interest. Bearing this in mind, and in order to verify formally the existence of spatial autocorrelation in the distribution under consideration, we proceeded by calculating Moran's I global test (Haining, 1990; Cliff and Ord, 1972) defined as:

$$I = \frac{n\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(y_i - \overline{y})(y_j - \overline{y})}{S_0 \sum_{i=1}^{n}(y_i - \overline{y})^2} \qquad (1)$$

where, in the present context, y_i stands for the percentage of firms with own website in region i, and $\overline{y}$ is the sample average. Likewise, w_{ij} denotes the corresponding element of the spatial weights matrix, W, where $\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij} = S_0$. The spatial weights matrix used in this paper takes into account interactions beyond adjacent regions. In particular, following Le Gallo and Ertur (2003), we considered a row-standardized matrix W based on the ten nearest neighbors, calculated using the geographic distance between the corresponding regional centroids (Pinkse and Slade, 1998). With regard to the interpretation of Moran's I, it should be noted that, after standardization, a significant and positive value of the statistic will indicate the existence of positive spatial autocorrelation, while a significant and negative value of Moran's I will reflect the presence of a pattern of spatial association between dissimilar values.

The result of the global Moran's I test provides us with a standardized value of 25.40, which is significant at the 1% level. This is clear evidence of the existence of a pattern of positive spatial association in this context, which is consistent with the initial impression drawn from Figure 1. We can therefore conclude that, in the European setting, firms located in spatially adjacent regions tend on the whole to exhibit a similar degree of Internet adoption. To further confirm this finding, we also constructed the corresponding Moran's scatterplot for the distribution under analysis. This is a graph on which the standardized values of the variable to be considered (z) are plotted on the horizontal axis and the spatial lag of the same variable (Wz) on the vertical axis. Thus, the four quadrants correspond to different types of spatial association. As can be seen from Figure 2, 86% of the regions considered are located in quadrants I and III. This confirms that the EU is characterized by the presence of spatial clusters of regions with similar levels of Internet adoption while there are relatively few cases in which a region registers a value of the variable under analysis markedly different from the average of its neighbors. Accordingly, the degree of Internet adoption across the European regions follows specific agglomeration patterns, which is in line with the results obtained in other studies (Giovannetti *et al.,* 2003; Weltevreden *et al.*, 2005; Zook, 2000; 2002; 2003; 2006).

[SEE FIGURE 2]

It is important to keep in mind, however, that Moran's I is calculated on a global basis for the whole of the sample. Hence, we do not know whether, irrespective of the overall dependence pattern, there are clusters of regions in which the concentration of high and low percentages of firms with their own website is significantly greater than what would be predicted in a homogeneous spatial distribution. It is also impossible with this test to check for the presence of groupings of regions with dissimilar values of the variable under analysis, that is, regions with a degree of Internet adoption significantly different from their neighbors. To overcome these shortcomings, we decided to calculate the local Moran's I, local Moran's I, I_i, by means of the following expression (Anselin, 1995):

$$I_i = \frac{n(y_i - \overline{y})}{\sum_{i=1}^{n} (y_i - \overline{y})^2} \sum_{j \in J_i} w_{ij} (y_j - \overline{y}) \tag{2}$$

where J_i denotes the set of regions neighbouring i. After standardization, a significant positive (negative) value of I_i will indicate a clustering around region i of regions with similar (dissimilar) values of the variable under consideration.

As in the global Moran's I test considered earlier, the usual practice in the case of I_i is to assume a normal asymptotic distribution when calculating the corresponding significance levels. Anselin (1995), however, has shown that the first and second order moments used in the standardization of the I_i statistic are obtained under the null hypothesis of no global spatial autocorrelation. However, this contrasts with the results obtained previously. Therefore, following the proposal of Anselin (1995), we calculated the corresponding pseudo-significance levels by means of an empirical distribution derived from 10,000 random permutations.

Figure 3 shows the significant regional groupings detected, and indicates whether or not they concentrate a similar percentage of firms with their own website. It can be seen that the conclusions to be drawn from this analysis are consistent with the results obtained earlier. Thus, the various regional clusters derived from the calculation of I_i are mostly made up of regions with a similar proportion of firms with their own website. In particular, it is possible to observe how the concentrations of high values of the variable under analysis are situated in Western Germany, the Netherlands, Belgium, most of the United Kingdom and the North of France, Italy, Sweden and Finland. On the other hand, the groupings of regions characterized by a low proportion of firms with a presence on the Internet are located in Portugal, Greece, Poland, Slovakia, a good part of Spain, the Czech Republic and Hungary, as well as the South of Italy. Additionally, it is interesting to note that Basse Normandie, Luxembourg (Belgium), Flevoland, and Cornwall are the only regions in our sample showing a significantly lower value of the variable examined than their neighbors. Likewise, as can be observed, Madrid registers a significantly greater presence of firms on the Internet than the adjacent areas. In any event, all other considerations notwithstanding, it is clear that a comparison of Figure 3 with the regional distribution of GDP per capita within the enlarged EU suggests the possible influence of regional development on Internet adoption at the regional level.

[SEE FIGURE 3]

The analysis so far is useful in describing the spatial distribution of website adoption in EU firms, but it seems unsuitable for quantifying the magnitude of regional differences in the variable that concerns us. To this end, one possibility would be to calculate some dispersion measure. Nevertheless, it is evident that none of the dispersion statistics proposed in the literature gives an accurate description of the entire distribution (Quah, 1996a; 1996b). As a consequence, we opted to estimate the density function of the distribution under consideration. Following common practice in the literature on economic growth, we have used non-parametric estimation techniques, thus avoiding the need to specify any particular functional form beforehand.

In particular, we employed an adaptive kernel method with flexible bandwidths (Silverman, 1986). This approach is especially advisable in the present context, since the possibility of varying the bandwidth along the support of the distribution allows a reduction in the variance of the estimates in areas characterized by the presence of few observations, and decreases the bias of the estimates in areas with many observations. In particular, we have used the adaptive two-stage estimator proposed by Abramson (1982) and given by:

$$\hat{f}(y) = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{h\,l_i}K\left(\frac{y-y_i}{h\,l_i}\right) \tag{3}$$

In the above expression K is a kernel function and $l_i = \sqrt{\frac{g}{\tilde{f}(y_i)}}$, where g is the geometric average over all i of the pilot density estimate $\tilde{f}(y)$. The pilot density estimate is a standard fixed bandwidth kernel density estimate obtained with h as a bandwidth. In this study Epanechnikov kernel functions were used, while the value of h was selected following Silverman (1986: 48).

[SEE FIGURE 4]

The results obtained are shown in Figure 4. The horizontal axis represents the proportion of firms with their own website normalized according to the sample average in order to facilitate comparisons, while the associated density appears on the vertical axis. As can be observed, most of the probability mass tends to be concentrated around the sample average. Nevertheless, there is also a second mode, located at the lower end of the distribution and formed by a relatively large number of regions with values of the study variable below 80% of the sample average. This finding may be potentially important, since it raises the possibility that the distribution under study will fragment in the long term into two separate groups of regions differentiated according to the proportion of firms with their own website.

In order to weigh up these results correctly, we estimated in the way just described the density function of the distribution of GDP per capita in our 239 regions. The results obtained are also presented in Figure 4. If we compare the external shape of both distributions, our estimates show the existence of relevant differences between them. Specifically, the analysis carried out reveals that existing territorial imbalances in GDP per capita are clearly greater than regional disparities in Internet adoption, confirming the available evidence (e.g. ESPON, 2006a). This suggests that there are other factors, in addition to GDP per capita, contributing to explain the observed differences in the proportion of firms with their own website across the EU regions. This issue will be discussed in greater detail in the next section.

Explanatory analysis: The influence of spatial effects

Selection of variables

Similar to previous innovations, empirical studies point up the role played by economic development in the diffusion of the Internet (Martin and Robinson, 2004; Andonova, 2006). GDP per capita is the indicator commonly used, as data are available at national and regional level. GDP per capita generally predicts the likelihood of adoption and the extent of use of the Internet because, among other reasons, it is related to communication infrastructure. When specialized infrastructures for adopting innovations are required, it is expected that innovations will be located in those areas with a higher level of infrastructure. Consequently, it could be predicted that the higher the level of GDP, the higher the level of funding invested in telecommunications and information technology (Fink and Kenny, 2003), and as a result, the higher the diffusion rate of the Internet. For the reasons mentioned, the GDP per capita (GDPpc) of the sample regions was included in the list of regressors of the model to be estimated in this section.

Following previous literature mentioned in Section two, population density becomes a relevant variable when explaining Internet adoption according to both urban and spatial models as well as diffusion theories. Here we investigate the possible influence of population density (DEN) on the proportion of firms with their own website.

As previously pointed out, along with the level of economic development and population density, other resources, such as human capital, have proven to be of critical importance as complementary factors in explaining Internet adoption. The educational level of a region can be used as a proxy of the availability of skilled labor available to firms adopting the Internet and as a means of knowledge creation and diffusion. As no data on skilled labor are available for European regions, we decided to consider the percentage of population aged 15 and over with completed tertiary education (EDU). In fact, education has been recognized as a key variable in empirical studies (Milicevic and Gareis, 2003; Martin and Robinson, 2004; Demoussis *et al.*, 2004).

Finally, we examined the role played in this context by regional specialization. The influence of the sectoral composition of economic activity on existing territorial imbalances in the EU is well known in the literature (Neven and Gouyette, 1995; European Commission, 1999). In fact, empirical evidence cited in Section two confirms the role of industry composition in Internet adoption (Forman and Goldfarb, 2005; Zook, 2000). Hence, we found it relevant to consider the possibility of a relationship between the productive structure of the different regions and the proportion of firms with a website. For this reason, our model includes the regional employment shares in agriculture (AGR), industry (IND) and market services (SER). Note, moreover, that despite the process of convergence in regional productive structures that has characterized the European economy during recent decades (Hallet, 2002), considerable differences persist in the patterns of regional specialization across the EU regions (Bode *et al.*, 2004).

Empirical evidence

Accordingly, the econometric model finally proposed to explain the proportion of firms with their own website in region i during year 2002 can be expressed as:

$$y_i = b_0 + b_1 GDPpc_i + b_2 DEN_i + b_3 EDU_i + b_4 AGR_i + b_5 IND_i + b_6 SER_i + e_i \qquad (4)$$

where e is the corresponding disturbance term. Before estimating this model, one should bear in mind the possibility of endogeneity problems in this context, given that some of the independent variables may be contemporaneously correlated with the error term. In order to overcome this potential problem, we opted to use lagged values of the independent variables in all the estimations (Rodríguez-Pose and Vilalta-Bufi, 2005). Specifically, we employed the average values of the various regressors over the period 1999-2001, which also allows us to minimize the possible effects of the economic cycle on the results of the analysis. Finally, it is worth mentioning that the data on the set of explanatory variables were drawn from EUROSTAT and Cambridge Econometrics.

[SEE TABLE 1]

As can be observed in Table 1, the unknown coefficients were estimated initially by Ordinary Least Squares (OLS). However, before discussing the results obtained, it is worth recalling that the analysis carried out in the preceding section clearly revealed the presence of positive spatial dependence between the European regions in this context. Therefore, we proceeded by performing various spatial dependence tests based on the residuals provided by the OLS estimations. Specifically, we calculated the Moran's I test (Cliff and Ord, 1972), the Lagrange multiplier tests for the spatial error and the spatial lag model proposed, respectively, by Burridge *et al.* (1980) and Anselin (1988a), plus their robust versions (Anselin *et al.*, 1996). Table 1 reveals that the results of these tests lead in all cases to the rejection of the null hypothesis of absence of residual spatial dependence. Specifically, the positive sign of the Moran's I provides evidence of the presence of a pattern of spatial association between similar values of the OLS residuals. Indeed, according to Anselin and Rey (1991), the values of the various Lagrange multiplier tests calculated suggest the need to include a spatial autoregressive structure in the error term (spatial error model). That is,

$$e = l We + u = (I - lW)^{-1}u \tag{5}$$

where l is the spatial autoregressive parameter and $u \sim N(0, s^2 I)$. Therefore, model (4) should be rewritten as:

$$y_i = b_0 + b_1 GDPpc_i + b_2 DEN_i + b_3 EDU_i + b_4 AGR_i + b_5 IND_i + b_6 SER_i + (I - lW)^{-1}u \tag{6}$$

The effects of spatial residual autocorrelation on the properties of the OLS estimator, which are similar to those present in the time series framework, indicate the inefficiency of the parameter estimates, due to the non-diagonal structure of the disturbance variance matrix (Anselin, 1988b). For this reason, following standard practice in the literature, model (2) was estimated by Maximum Likelihood (ML).

As can be seen in Table 1, the spatial autoregressive parameter is statistically significant. Indeed, the likelihood ratio test performed leads to the rejection of the null hypothesis of $l = 0$, thus confirming the internal coherence of the spatial error model in this context. This implies that a random shock introduced into a specific region will not only affect the proportion of firms with their own website in that region but, through the spatial transformation $(I - lW)^{-1}$, will influence the value of website adoption in the remaining regions.

Before commenting in detail on our findings on the role played in this context by the set of explanatory variables considered, it should be noted that the degree of collinearity among the regressors is relatively moderate, which enables us to be more confident in the estimates of single coefficients. In particular, the variance inflation factor provides an average value of 2.13, which is within the acceptable limits established by Belsley *et al.* (1980). That said, it is worth noting that the coefficient on the GDP per capita is positive and statistically significant, therefore confirming our previous results. Nevertheless, this is not the only factor that contributes to explain the spatial distribution of the Internet across EU firms. Specifically, population density and the share of population with tertiary education completed are positively correlated with the dependent variable. Likewise, our estimate reveals that regional differences in the sectoral composition of economic activity influence the degree of Internet adoption. Thus, the presence of a relatively large agricultural sector has a negative impact on the proportion of firms with their own website, which is not particularly surprising taking into account the specific features of farming markets. However, the rest of the variables included in the model to control for the industry mix are not statistically significant.

The analysis performed so far has highlighted the importance of spatial effects when it comes to explaining the spatial distribution of the Internet across EU firms. Nevertheless, when assessing the implications to be drawn from this result, it should be borne in mind that the approach adopted in the preceding pages was based on the EU as a whole. Accordingly, we have not considered the possibility that common national characteristics could play a relevant role in the transmission of spatial effects within the regions belonging to different countries. In order to investigate this issue, we constructed within-countries and between-countries spatial weights matrices. In the first matrix we set equal to one only the weights corresponding to regions which belong to the same country. In the second case the weights for regions sharing a border but belonging to different countries are those equal to one.

[SEE TABLE 2]

Table 2 shows the results obtained when these two matrices are used to calculate the various tests for spatial dependence from the OLS estimation of model (4). As can be observed, in the case of the within-countries matrix the analysis confirms the presence of spatial residual dependence. Although this conclusion is in line with our previous estimates, it needs to be said that the values of all the tests are greater than those reported in Table 1 using a spatial weights matrix based on the ten-nearest neighbors. However, the situation is different when the between-countries matrix is employed in the analysis. In this case, the value of the Moran's I test and the two Lagrange multiplier tests calculated for the spatial error model are not statistically significant, while the two Lagrange multiplier tests for the spatial lag model are only significant at 10%. Therefore, this analysis suggests that the spatial effects observed in this context are mainly constrained by political borders between the various EU member states, which highlights the important role played by the national dimension in explaining the presence of the Internet in EU firms.

CONCLUSIONS AND POLICY IMPLICATIONS

This study offers empirical evidence on regional distribution of Internet adoption by European firms within the framework of the on-going debate on whether or not the Internet means the “death of distance”. Our results highlight the significance of spatial effects in explaining regional disparities in Internet adoption. For the first time in this literature, we have used a set of methodological tools from spatial econometrics that allows us to capture the spatial characteristics of the data and the influence of geographic proximity in shaping Internet diffusion within the EU.

Our findings show that there are important differences in the share of firms with their own website across the regions of the enlarged EU. Nevertheless, when assessing the magnitude of regional disparities within this framework, it should be noted that the various density functions estimated indicate that the level of dispersion of our study variable is more reduced than that registered by GDP per capita.

In addition, the empirical evidence obtained reveals the presence of positive spatial autocorrelation. This implies that the presence of the Internet in EU firms is not randomly distributed across space. On the contrary, physically adjacent regions tend, on the whole, to exhibit a similar proportion of firms with their own website. In fact, we have detected the existence of several clusters of regions with similar values of the study variable distinguishing them from the neighboring zones. Specifically, the groupings of regions with a significantly high share of firms with presence on the Internet are situated in Western Germany, the Netherlands, Belgium, the United Kingdom, and the North of France, Italy, Sweden and Finland. On the other hand, the clusters characterized by a low proportion of firms with their own website are located in the Southern and Eastern periphery of the Union. These results are potentially important, since they suggest that the enlargement process of the EU towards Central and Eastern Europe has introduced new spatial patterns in the North-South digital divide identified by the previous literature on the subject.

In order to complete these results, we have carried out a causal analysis of the observed regional differences. Taking into consideration that the presence of spatial autocorrelation in this context affects negatively the results obtained from standard regression analysis, we have estimated an econometric model incorporating a spatial autoregressive structure in the error term. Our estimates indicate that the share of firms with their own website in the regions of the enlarged EU is positively correlated with the level of GDP per capita, population density and the percentage of population with tertiary education. On the other hand, the existence of a relatively large agricultural sector has a negative impact on the dependent variable.

In any event, it is important to note that the results obtained show clearly the importance of spatial effects when it comes to explaining the spatial distribution of the Internet across EU firms. Nevertheless, the empirical evidence provided in this paper indicates that the spatial effects observed in this context are mainly constrained by political borders between the various EU member states. By contrast, the transmission of spatial spillovers across regions belonging to different countries is less relevant, thus confirming the important role played by the national dimension in explaining the presence of the Internet in the EU firms. From a policy perspective, this empirical evidence confirms the results obtained by other regional studies (e.g. ESPON, 2006a) and emphasizes the need for national ICT policies rather than region-specific ones to foster technological convergence in the enlarged EU.

The main limitation of this study is the lack of data at regional level, not only in relation to Internet adoption but also to possible explanatory regional variables such as human capital, industrial composition or institutional variables. In addition, the availability of data for several years would have allowed us to study the evolution over time of regional disparities in Internet adoption across European regions. Further research should also include the analysis and comparison of possible spatial effects generated by the adoption of various types of ICT by households, firms and governments.

References

Abramson, I. S. (1982) On bandwidth variation in kernel estimates - a square root law. *Annals of Statistic* 10, 1217-23.

Andonova, V. (2006) Mobile phones, the Internet and the institutional environment. *Telecommunications Policy* 30, 29-45.

Anselin, L. and S. J. Rey (1991) Properties of tests for spatial dependence in linear regression models. *Geographical Analysis* 23, 112-31.

Anselin, L. (1988a) Lagrange multiplier test diagnostics for spatial dependence and spatial heterogeneity. *Geographical Analysis* 20, 1-17.

Anselin, L. (1988b) *Spatial econometrics: Methods and models*. Kluwer Academic, Dordrecht.

Anselin, L. (1995) Local indicators of spatial association-LISA. *Geographical Analysis* 27, 93-115.

Anselin, L., A. Bera, R. Florax and M. Yoon (1996) Simple diagnostic tests for spatial dependence. *Regional Science and Urban Economics* 26, 77-104.

Audretsch, D.B. and M.P. Feldman. (2004) Knowledge spillovers and the geography of innovation. In V. Henderson and J.F. Thisse (eds.), *Handbook of regional and urban economics*, vol. 4, Elsevier, North-Holland.

Bailey, T. and A.C. Gatrell (1995) *Interactive spatial data analysis*. Longman, Essex.

Bell, D.R. and S. Song (2004) Neighbourhood effects and trial on the Internet: Evidence from online grocery retailing. Wharton School Marketing Department Paper, Philadelphia.

Belsley, D., E. Kuhn and R. Welsh (1980) *Regression diagnostic identifying influential data and source of collinearity*. John Wiley, New York.

Billón, M., R. Ezcurra and F. Lera-López (2007) The spatial distribution of the Internet in the European Union: Does geographical proximity matter? *European Planning Studies* (Forhcoming).

BISER (2004) *ICT infrastructure. The regional dimension*. BISER Domain Report, 10, Rome, http: www.biser-eu.com/results.

Bode E., C. Krieger-Boden and R. Soltwedel (2004) *European integration, regional structural change and cohesion in the EU: Cross-country summary*. EURECO Working Paper, Kiel.

Burridge, P. (1980) On the Cliff-Ord test for spatial autocorrelation. *Journal of the Royal Statistical Society B* 42, 107-8.

Cairncross F. (1997) *The death of distance: how the communications revolution will change our lives*. Harvard Business School Press, Boston.

Cliff, A. and J. Ord (1972) Testing for spatial autocorrelation among regression residuals. *Geographical Analysis* 4, 267-84.

Council of the European Union (2006) *Council decision of 6 October 2006 on community strategic guidelines on cohesion.* Official Journal of the European Union, 2006/702/EC, Brussels.

Demoussis, M., and N. Giannakopoulos (2006) The dynamics of home computer ownership in Greece. *Information Economics and Policy* 18, 73-86.

Del Aguila, A.R. and A. Padilla (2006) Organizational factors affecting Internet technology adoption. *Internet Research* 16.1, 94-110.

Demunter, C. (2005) The digital divide in Europe. *Statistics in Focus* 38, 1-8.

ESPON (2006a) *ESPON Project 1.2.3. Identification of spatially relevant aspects of the Information Society.* Draft Final Report. http://www.espon.eu.

ESPON (2006b) *ESPON project indicators*. *ESPON database 2006*. http://www.espon.eu.

European Commission (1999) *Sixth periodic report on the social and economic situation of the regions in the European Union*. European Commission, Brussels.

European Commission (2005) *The 2005 European e-business readiness index*. European Commission, Brussels.

Fink, C. and C.J. Kenny (2003) W(h)ither the digital divide? *Info, The Journal of Policy, Regulation and Strategy for Telecommunications* 5, 1-19.

Forman, C. (2003) The geographic dispersion of commercial Internet use. In S. Wildman and L. Cranor (eds.) *Rethinking rights and regulations: Institutional responses to new communication technologies*, MIT Press, Cambridge, MA.

Forman, C. (2005) The corporate digital divide: Determinants of Internet adoption. *Management Science* 51.4, 641-54.

Forman, C. and A. Goldfarb (2005) Diffusion of ICT to businesses. In T. Hendershott (ed.), *Handbook of Economics and Information Systems*, Elsevier, forthcoming.

Forman, C., A. Goldfarb and S. Greenstein (2005a) How did location affect adoption of the commercial Internet? Global village vs. urban leadership. *Journal of Urban Economics* 58, 389-420.

Forman, C., A. Goldfarb and S. Greenstein (2005b) Geographic location and the diffusion of Internet technology. *Electronic Commerce Research and Applications* 4.1, 1-13.

Fujita, M. and T. Mori (2005) Frontiers of the new economic geography. *Papers in Regional Science* 84, 377-405.

Fujita, M., P. Krugman and A. Venables (1999) *The spatial economy.* MIT Press, Cambridge, MA.

Galliano, D. (2005) *The evolution of the spatial digital divide. From Internet adoption to internet use by French industrial firms*. Cahiers du GRES 25. University of Montesquieu Bordeaux IV, Bordeaux.

Geroski, P.A. (2000) Models of technology diffusion. *Research Policy* 29, 603-25.

Giovannetti, E., K. Neuhoff and G. Spagnolo (2003) Agglomeration in the Internet: Does space still matter? The MIX-IXP Case. In M. Kagami and M. Tsuji (eds.), *Economic development through Industrial agglomeration: Factors which attract firms*, IDE-JETRO, Tokio.

Greenstein, S. and J. Prince (2006) The diffusion of the Internet and the geography of the digital divide in the United States. In C. Avgerou, D. Quah and R. Silverstone (eds.) *Oxford Handbook of ICTS*, Oxford University Press, forthcoming.

Haining, R. (1990) *Spatial data analysis in the social and environmental sciences*. Cambridge University Press, Cambridge.

Hallet, M. (2002) Regional specialisation and concentration in the EU. In Cuadrado-Roura (ed.) *Regional convergence in the European Union: Facts, prospects and policies,* Springer, Berlin.

Karshenas, M. and P. Stoneman (1995) Technological diffusion. In P. Stoneman (ed.), *Handbook of the Economics of Innovation and Technological Change*, Blackwell, Oxford.

Kaufmann, A., P. Lehner and F. Todtling (2003) Effects of the Internet on the spatial structure of innovation networks. *Information Economics and Policy* 15, 402-24.

Kolko, J. (2000) The death of cities? The death of distance? Evidence from the geography of commercial internet usage. In I. Vogelsang and B.M. Compaine (eds.), *The Internet Upheaval*, MIT Press, Cambridge, MA.

Kolko, J. (2002) Silicon mountains, silicon molehills: Geographic concentration and convergence of Internet industries in the US. *Information Economics and Policy* 14, 211-32.

Krugman, P. (1991) Increasing returns and economic geography, *Journal of Political Economy* 99, 482-99.

Le Gallo, J. and C. Ertur (2003) Exploratory spatial data analysis of the distribution of regional per capita GDP in Europe, 1980-95. *Papers in Regional Science* 82, 175-201.

Leamer, E.E. and M. Storper (2001) The economic geography of the Internet age. *Journal of International Business Studies* 32, 641-665.

Malecki, E. (2002) The economic geography of the Internet´s infrastructure. *Economic Geography* 78.4, 399-424.

Martin, S. and J. Robinson (2004) The income digital divide. *IT & Society* 1, 1-20.

Milicevic, I. and K. Gareis (2003) *Disparities in ICT take-up and usage between EU regions*. Workshop The regional effects of the New Information Society, Milano.

Neven, D. and C. Gouyette (1995) Regional convergence in the European Community. *Journal of Common Market Studies* 33, 47-65.

Ottaviano, G. and D. Puga (1998) Agglomeration in the global economy: A survey of the new economic geography. *The World Economy* 21, 707-731.

Pinkse, J. and E. Slade (1998) Contracting in space: an application of spatial statistics to discrete-choice models. *Journal of Econometrics* 85, 125-54.

Quah, D. (1996a) Empirics for economic growth and convergence. *European Economic Review* 40, 1353-75.

Quah, D. (1996b) Twin peaks growth and convergence in models of distribution dynamics. *The Economic Journal* 106, 1045-1055.

Rodríguez-Pose, A. and M. Vilalta-Bufí (2005) Education, migration, and job satisfaction: the regional returns of human capital in the EU. *Journal of Economic Geography* 5, 545-566.

Rogers, E.M. (1995) *Diffusion of innovations.* Free Press, New York.

Rosenberg, N. (1972) Factors affecting the diffusion of technology. *Explorations in Economic History* 10, 3-33.

Silverman, B. (1986) *Density estimation for statistics and data analysis*. Monographs on Statistics and Applied Probability, 26. Chapman and Hall, London.

Sternberg, R. and M. Krymalowski (2002) Internet domains and the innovativeness of cities/regions. Evidence from Germany and Munich. *European Planning Studies* 10.2, 252-273.

UNDERSTAND (2006) *Understand. Result synopsis.* UNDERSTAND Conference. Brussels.

Van Oort, F. and O. Atzema (2004) On the conceptualization of agglomeration economies: The case of new firm formation in the Dutch ICT sector. *Annals of Regional Science* 38, 263-90.

Weltevreden, W., O. Atzema, K. Frenken, K. Kruif and F. Van Oort (2005) *The geography of Internet adoption by retailers*. Papers in Evolutionary Economic Geography 05.10. Utrecht University, Utrecht.

Zook, M.A. (2000) The web of production: the economic geography of commercial Internet content production in the United States. *Environment and Planning A* 32.3, 411-26**.**

Zook, M.A. (2002) Grounded capital: Venture financing and the geography of the Internet industry, 1994-2000. *Journal of Economic Geography* 2, 151-77.

Zook, M.A. (2003) Underground globalization: mapping the space of flows of the Internet adult industry**.** *Environment and Planning A* 35.7, 1261-86.

Zook, M.A. (2006) The geographies of the Internet**.** *Annual Review of Information Science and Technology* 40, 53-78**.**

Figures and Tables

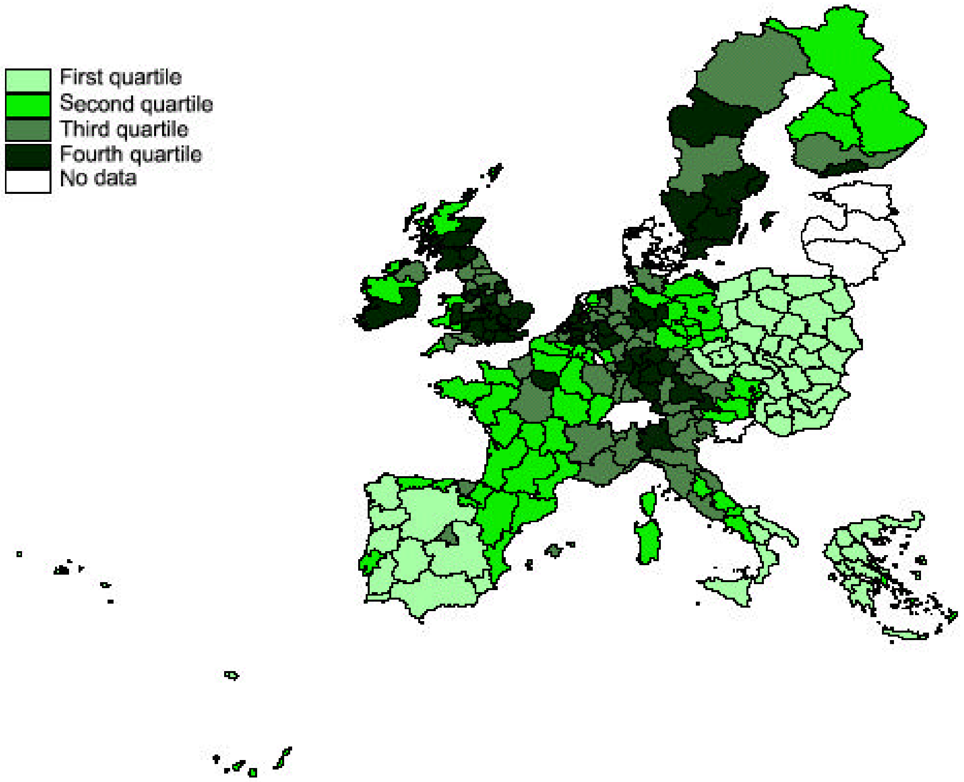

Figure 1: The spatial distribution of the proportion of firms with their own website.

Figure 2: Moran's scatterplot.

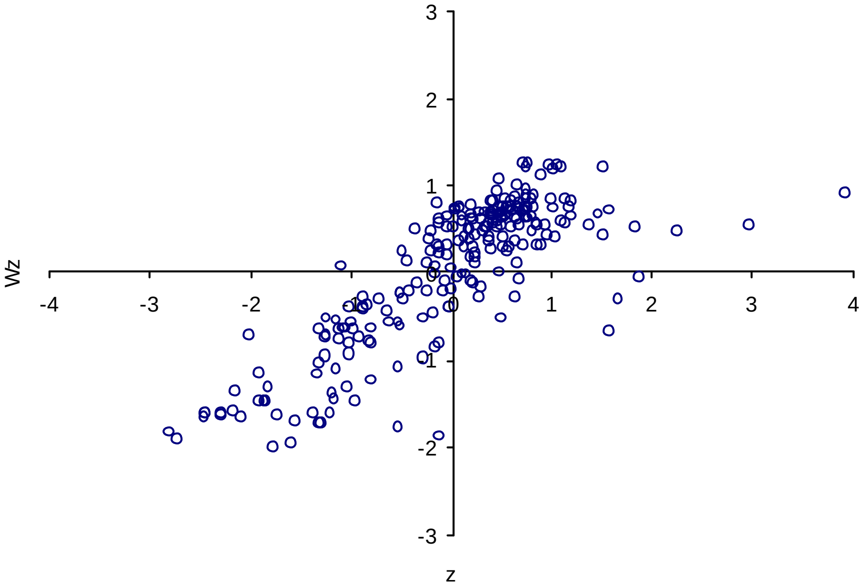

Figure 3: Spatial distribution of local Moran's I.

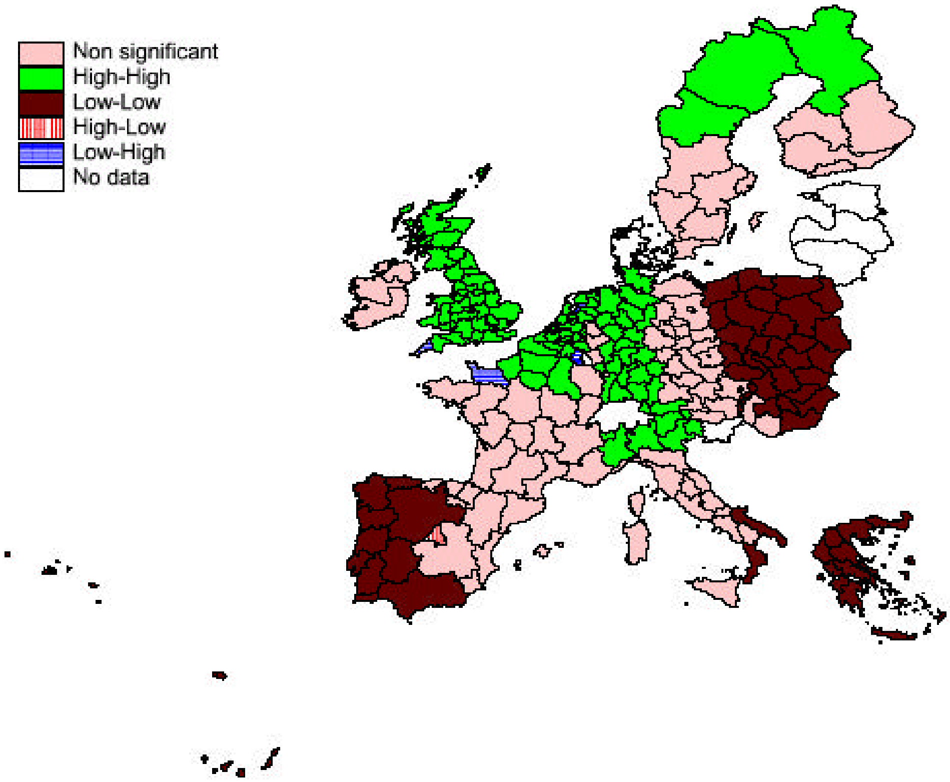

Figure 4: Density functions of the spatial distribution of the proportion of firms with their own website and GDP per capita (sample average equal to 100).

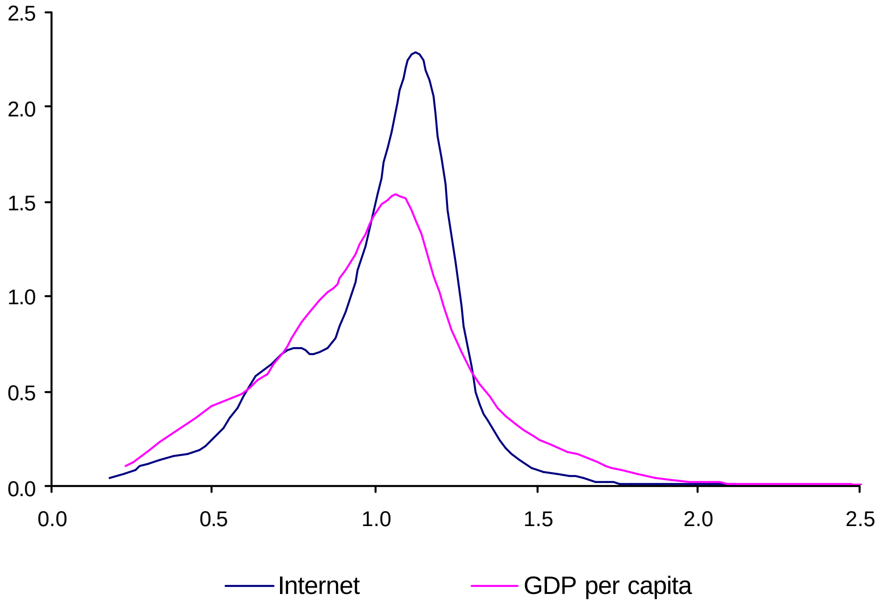

Table 1: Regression analysis.

Variables	**OLS**	**ML**
Constant	0.546	0.648
	(0.000)	(0.000)
GDPpc	0.395	0.374
	(0.000)	(0.000)
EDU	0.072	0.041
	(0.000)	(0.009)
DEN	0.000	0.003
	(0.849)	(0.090)
AGR	-0.094	-0.087
	(0.000)	(0.000)
IND	-0.003	-0.008
	(0.841)	(0.620)
SER	0.084	0.031
	(0.028)	(0.420)
λ		0.782
		(0.000)
LIK	293.806	354.992
AIC	-573.611	-695.985
SC	-549.276	-671.649
Moran's I	15.761	
	(0.000)	
LMERR	191.258	
	(0.000)	
R-LMERR	146.183	
	(0.000)	
LMLAG	71.452	
	(0.000)	
R-LMLAG	26.376	
	(0.000)	
Likelihood ratio test for $\lambda=0$		122.373
		(0.000)
LMLAG*		2.614
		(0.106)

Notes: The dependent variable is the proportion of firms with their own website. p-values are in parentheses. LIK is the value of the maximum likelihood function, while AIC and SC are the Akaike and Schwarz information criteria, respectively. LMERR is the Lagrange multiplier test for residual spatial autocorrelation and R-LMERR is its robust version. LMLAG is the Lagrange multiplier test for spatially lagged dependent variable and R-LMLAG is its robust version. LMLAG* is the Lagrange multiplier test on spatial lag dependence in the spatial error model.

Table 2: Spatial dependence within countries and between countries.

Spatial weights matrix	**Within-countries**	**Between-countries**
Moran's I	18.018	1.432
	(0.000)	(0.152)
LMERR	242.538	1.801
	(0.000)	(0.180)
R-LMERR	185.802	2.324
	(0.000)	(0.127)
LMLAG	93.460	3.185
	(0.000)	(0.074)
R-LMLAG	36.725	3.707
	(0.000)	(0.054)

Notes: p-values are in parentheses. LMERR is the Lagrange multiplier test for residual spatial autocorrelation and R-LMERR is its robust version. LMLAG is the Lagrange multiplier test for spatially lagged dependent variable and R-LMLAG is its robust version.