Bad Neighbors and The Internet: A Geospatial Analysis of Internet Adoption

Cheng Yee Lim

May 28, 2017

Abstract

The paper investigates the role of location in internet adoption with a spatial econometric methodology. This paper employs country-level data of both proportion of internet users and fixed broadband subscriptions to identify good and bad neighborhoods with spatial clustering techniques. It is argued that location matters because of various kinds of spillovers, including knowledge spillovers and spillovers in terms of attitude, beliefs and social norms. The empirical findings support that internet adoption is affected by adoption rates in countries in close proximity, even when we define countries in close proximity with different types of spatial weights. The paper also addresses potential endogeneity concerns of a spatial lag using an instrumental variables approach and finds that the significance of location remains stable. The findings suggest that there is a growing importance of location in internet adoption globally and imply that in order to promote internet adoption, development aid and international policies may be more effective if they target clusters of countries than single countries.

Keywords: Internet adoption, spatial econometrics, macro-geographic location, neighborhoods

1. Introduction

Digital technologies, such as the internet and mobile phones, have spread at an unprecedented rate. The diffusion process of the Internet has been estimated to reach saturation (1-99%) within a country in 16.1 years, dramatically surpassing the duration taken to reach saturation by other General Purpose Technologies. General Purpose Technologies, such as steam-power and electrification, took a duration of 100 and 60 years to reach saturation respectively. (Ackermann, Angus, and Raschky 2017) In fact, the number of internet users has tripled within a decade, from 1 billion users in 2005 to to 3.2 billion in 2015. (Bank 2016) Besides the unprecedented rate of diffusion, the internet has also provided the means of new possibilities in society. The internet enabled businesses, people and governments to be more connected than ever before. Furthermore, the internet has played a huge role in creating greater ease in communication and access to information.

Previous literature have consistently proven that economic affluence, which can be represented as the GDP per capita and telecommunication network density, is a pertinent determinant of internet diffusion (Beilock and Dimitrova 2003, Hargittai (1999), Crenshaw and Robison (2006), Wunnava and Leiter (2008), Warf (2009), Andres et al. (2010)). Similarly, previous studies have also shown how the demography of a country, such as educational attainment and age, can affect internet adoption. Whilst less validated across studies, institutional qualities, such as political freedom and freedom of speech, were also found to be correlated with greater internet adoption (Wallsten 2005, Andonova and Diaz-Serrano (2009)). However, previous studies have overlooked geographic location and possibilities of spatial spillovers of internet adoption across countries. Countries that are close in spatial proximity may receive spatial spillovers in internet adoption from each other. As illustrated in Figure 1, bordering countries tend to be in similar quintiles of internet adoption worldwide.

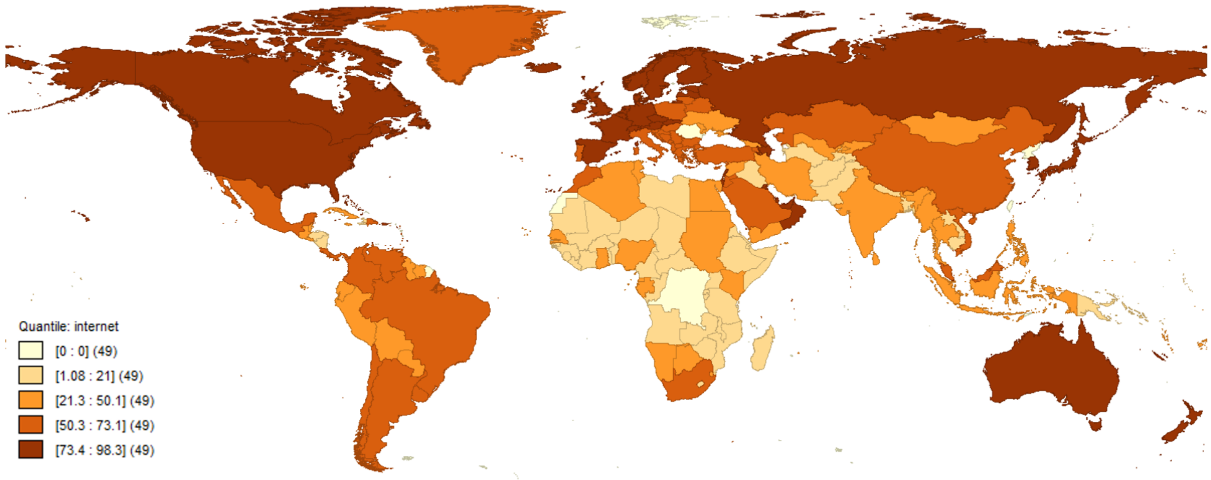


Figure 1: Quintile Map of Internet Adoption in the World in 2015

It is crucial to understand the determinants of internet adoption as the benefits of the digital technologies are multifold. These benefits span across different sectors of the economy consumers, producers, and governments. For consumers and producers, these technologies have the potential to offer new opportunities by creating jobs that leverage human capital, and in turn producing consumer surplus. Furthermore, digital technologies strengthen the government's capability to deliver public services to its citizens and serve as an avenue for citizens to overcome collective action problems. While these benefits are neither assured nor automatic, digital technologies do have the potential of bringing significant gains upon adoption.

Most crucially, digitization is changing the global competition landscape and heralding numerous opportunities for new entrants to penetrate global value chains, contesting against incumbents. This phenomenon is not only applicable to companies, but can also be extended to countries. It is, however, ambiguous and debatable whether the digital revolution is enabling poor countries to smoothly integrate and catch up with the global economy, or whether rich countries will be able to maintain their competitive advantage and accelerate growth by utilizing digital technologies (Bank 2016).

Since the internet is the core technology crucial for reaping the advantages of digitization, the ability to leverage on new technologies, such as the internet, and the state of digital infrastructure are increasingly likely to determine the winners and losers of the technology revolution. This is because such capabilities directly influence which countries, industries, and companies create and lose value. (Hirt and Willmott 2014) It is, therefore, paramount that developing countries promptly close the digital gap between. However, Table 1 shows that the difference in the proportion of internet users in developed and developing countries has increased from 44.05% to 46.97% from 2006 to 2015.

Year	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
Developed Countries	53.45	59.03	61.27	62.90	66.51	67.68	73.84	76.87	79.50	82.25
Developing Countries	9.40	11.92	14.64	17.42	21.07	24.05	26.99	29.51	32.41	35.28
Digital Divide	44.05	47.11	46.63	45.48	45.44	43.63	46.85	47.36	47.09	46.97

Table 1: Difference in Internet Users between Developed and Developing Countries

Source: ITU (2016)

Opposing the optimistic outlook of the Internet potentially improving productivity, communications and enabling new possibilities in developing countries, these benefits may be widening the existing chasm between the haves and the have-nots, due to the lack of skills, resources and infrastructure in developing countries (Norris 2001). This is worrying as the current circumstance of digital inequality seems to favor those with an abundance of resources. This is well-encapsulated with a quote by Rodriguez and Wilson, "when a new technology is introduced into a social setting where scarce resources and opportunities are distributed asymmetrically, there is a greater likelihood that those with more resources will employ them to gain additional ones". (Rodriguez and Wilson, n.d.).

In this paper, we relook at the internet adoption process with the possibilities of spatial spillovers in hopes to devise new strategies to promote internet adoption and abridge the digital gap. On top of previous findings and policy recommendations, the presence of spatial spillovers could provide a new perspective in development aid distribution. If spatial spillovers are shown to be significant, development aid with the goal of promoting digital technologies can be directed to regions or clusters of countries. More fundamentally, the presence of spatial spillovers could point to a misspecification in previous models of internet adoption that have not accounted for spatial possibilities, resulting in biased and inconsistent estimates in the model.

2. Literature Review

Previous literature has established three main channels - economic, demographic and institutional - for internet proliferation in a country. Economic variables, such as telecommunication infrastructure and GDP per capita, were identified as key determinants of Internet diffusion in various studies. Economic variables, such as telecommunication infrastructure and GDP per capita, were identified as key determinants of Internet diffusion in various studies (Beilock and Dimitrova 2003, Hargittai (1999), Crenshaw and Robison (2006), Wunnava and Leiter (2008), Warf (2009), Andres et al. (2010)). Hargittai (1999) is one of the first econometric studies analyzing the spread of the Internet across countries. Using a sample of OECD countries from 1994 to 1997, Hargittai finds that differences in GDP per capita affect diffusion rates, even for OECD countries with comparable social and economic development. Her findings also identify other important determinants such as telecommunications policy and telephone density. Kiiski and Pohjola also found income per capita, telephone access costs, and years of schooling to be significant determinants in their examination of internet penetration in 60 OECD and developing countries. (Kiiski and Pohjola 2002) Beilock and Dimitrova (2003) conducted their empirical study on a sample of 105 countries with a greater variation of socioeconomic levels. Their results affirm previous findings that income per capita is the paramount determinant of aggregate Internet usage in countries. Other important factors from their findings include telephone infrastructure, political freedom and economic openness of a country. Additionally, Chinn and Farlie examined a sample of 161 countries from 1999 to 2001 and attempted to decompose the relative importance of determinants of internet adoption. Their findings asserted that the most influential factor in Internet diffusion is GDP per capita, followed by telephone lines per capita. (Chinn and Fairlie 2006) In their next study, they focused on disparity in internet adoption in developed and developing countries and found that income per capita, telephone density, human capital and legal quality account for poor internet adoption in developing countries.

Other economic variables that have been considered in different comparative studies include openness to trade, regulatory policies of the telecommunication industry and educational attainment. Mixed conclusions have been found for these three variables, depending on the measurements of the data and empirical specification. Low levels of education are expected to impede the accessibility and diffusion of technologies. In Kiiski and Pohjola's (2002) analysis of Internet diffusion in both OECD and developed countries, average years of schooling positively affects Internet diffusion, but competition in the telecom market was found to be insignificant. More recent work also found evidence of school education positively affecting ICT adoption (Cruz-Jesus et al. 2016, Tengtrakul and Peha (2013)). Caselli and Coleman's study found that imports per worker and the attainment of secondary education strongly increases computer diffusion (Caselli and Coleman II 2001). However, Hargittai (1999) and Chinn and Farlie (2007) did not find a significant relationship between education and Internet diffusion, but their results highlighted the significance of telecommunication regulatory policy.

Besides economic factors, previous literature had also taken an interest in investigating the importance of institutions, policies crafted and enforced for ICTs advancement (Wallsten 2005, Andonova and Diaz-Serrano (2009)). In Andonova's analysis of a cross-section of developed and developing countries in 2001, she used different measures of institutional quality, such as civil liberties, political constraints, and political rights, and their resultant impacts on investment climate to posit an explanation of the differences between Internet and mobile phone usage; and found a positive relationship between infrastructural development, institutional environment and Internet usage. (Andonova 2006) Henisz and Zelner also concluded that the risk of the state taking over an investment is a paramount institutional parameter in telecommunications development. (Henisz, Zelner, and others 2001) Additionally, countries with greater political freedom and better property rights are more likely to have higher Internet adoption rates (Crenshaw and Robison 2006).

Many studies of Internet diffusion also considered demographic controls as particular demographic characteristics are expected to advance Internet diffusion. For instance, countries with a younger population and greater urbanization are expected to accept the Internet more readily. Previous studies by Goldfarb and Prince (Goldfarb and Prince 2008) and Chinn and Farlie (2007) hint at a positive relationship between youth and Internet adoption, which aligns with findings from other micro-data (National Telecommunications and Information Administration 2002). On the other hand, findings regarding levels of urbanization in countries are less clear-cut. Andonova (2006) and Crenshaw and Robison (2006) find a significantly positive impact of urbanization, whereas Chinn and Farlie (2007) find opposite results.

3. A Spatial Perspective in Internet Adoption

Despite the extensive literature on Internet adoption, studies have paid little attention to neighborhoods and macro-geographic location thus far (Dohse and Lim 2016). Majority of studies has treated each country as independent units of internet adoption and overlooked the possibilities of cross-country interactions in the adoption process (Comin and Mestieri 2013). This oversight is surprising as macro-geographic neighborhood effects has been well established in the development economics (Collier 2008) and knowledge spillover literature (Keller 2002). In this paper, we explore how internet adoption is affected by neighborhoods and cross-country interactions, a process that can be termed spatial spillover.

Internet adoption in a country could very likely be affected by Internet adoption in countries that are spatially nearer for the following reasons. As previously shown by Keller (2002), international technology diffusion is geographically localized. Keller reports that spillovers of technology diffusion are halved after a distance of 1,200km. Similarly, spatial proximity between countries can also facilitate the transfer of knowledge of new technologies (the internet) by increasing the likelihood of an encounter of an internet adoptee from another country. For instance, assume country A has a growing population of internet users and is in close proximity with country B. As the number of internet adoptees in country A grows, the occurrence of a knowledge transfer between an internet adoptee from country A with an inhabitant from country B becomes more likely. This flow of knowledge is also, otherwise, termed as geographical knowledge spillovers. As geographical knowledge spillovers reduce over distance, technology adoption in a given country tends to benefit neighboring countries more than distant countries.

Furthermore, the theories of technology adoption behavior (Mathieson 1991, Davis Jr (1986), Ajzen (2011)) consistently suggest that attitudes, aptitudes and social norms are important in leading to actual technology adoption. The similar attitudes of people in neighboring countries towards new technologies is well documented by cross-country surveys of attitudes and views towards science and technology in society (Gaskell et al. 2010). A similar argument holds for norms and values. As evidenced by the World Values Survey and related research (Parts 2013, Wach and others (2015), Berggren and Nilsson (2015)), social norms and values tend to be clustered in groups of neighboring countries. Clustering across more distant countries is also possible in countries that share a common history and language, such as Commonwealth countries, but spatial proximity doubtlessly facilitates the spread of norms and values.

The similar social norms and attitudes towards new technologies in neighboring countries are further perpetuated by the reference groups of the inhabitants. The reference group theory states that a reference group may mean a group with which one compares oneself in making a self judgement (Robert K Merton and Kitt 1950). The reference group of inhabitants of a given country can also extend to people of neighboring countries with similar contexts. Peer groups and reference persons play an important role in the establishment of norms and in the forming of attitudes and intentions (Robert King Merton 1968, Venkatesh and Morris (2000), Falck, Heblich, and Luedemann (2012)). Peer or reference groups might also serve as role models for individual behavior, such as Internet adoption. While reference groups need not be restricted to neighboring countries, spatial proximity is clearly helpful as people typically have better knowledge of neighboring countries than distant ones, and identify with people with common cultural and historical roots and common religious beliefs.

Thus, we argue that there are manifold interactions in the internet adoption process among countries in close proximity. The impact of the neighborhoods might go well beyond mere knowledge spillovers, but also contribute to the formation of attitudes, norms and beliefs that influence internet adoption behavior. Therefore, in addition to the established determinants of internet adoption in the literature, we hypothesize that there are spatial dependence (spatial effects and externalities) in a country's adoption of the internet. In other words, we expect internet adoption in a country to be positively affected by internet adoption of other countries in close proximity.

4. Overview of Research Design

In this paper, we demonstrate that macro-geographic locations matter for cross-country internet adoption with country-level data using a spatial econometric methodology. Firstly, we conduct exploratory spatial data analysis on internet adoption and other key variables and describe the limitations of our data. Secondly, we explain how we overcame the limitations in our data with spatial weights and define countries that are in close proximity. Thirdly, we expound how spatial dependence can be introduced in a regression specification and conduct the necessary diagnostics tests to determine our model specification. Fourthly, we report how the findings of the estimated model reaffirms the importance of spatial spillovers in internet adoption. Lastly, we conclude the paper with methodological and practical implications of our paper.

5. Data and Exploratory Spatial Data Analysis

5.1 Data

Data on telecommunications, institutional characteristics and country demographics have been consolidated for 195 countries in 2004, 2009 and 2014 from the World Bank and International Telecommunication Union. 2014 was selected as the most recent year of study for the completeness of the relevant variables in our study. We then model the diffusion process with an evenly spaced five year interval from 2014, to 2009 and to 2004. Among the 195 countries, 119 of them were categorized as upper middle and high income countries according to the World Bank Country Classifications. Using the World Bank Country Classifications, we defined upper middle and high income countries as developed countries and the low and lower middle income countries as developing countries. A limitation of our data set is that the World Bank reflects data on only 195 countries when they are 245 countries in the world. These missing countries[1] will lead to other countries having "missing neighboring countries" in the data set.

The main metric we will be using for internet adoption is the percentage of the population as internet users in a country, which is also the most common measure of internet adoption in the literature. Since technology adoption is made up a series of individual adoption decisions, the most relevant measure of technology adoption is the ratio of actual to potential users (Andres et al. 2010). Furthermore, the percentage of population as internet users is preferred over other metrics, such as computer penetration rates and internet subscribers, as it includes not only internet access from the household, but also public places, such as workplaces, universities and internet cafes (Andres et al. 2010). We will also conduct a robustness check on the metric for internet adoption with the proportion of the population with fixed broadband subscribers. According to the International Telecommunication Union (ITU), fixed broadband subscriptions are fixed subscriptions with high-speed access to public internet.[2] Economic variables such as GDP per capita, level of

[1] The treatment of missing countries will be explained in Section 6.1.

[2] Fixed broadband subscriptions refers to fixed subscriptions to high-speed access to the public Internet (a TCP/IP connection),

urbanization, telephone fixed line subscribers were included as a proxy of a country's ability in providing the relevant telecommunication infrastructure in the country. Demographic characteristics, such as proportion of population aged 65 and primary education enrollment rate, were also included to reflect the likelihood of inhabitants adopting new technologies. Lastly, the number of days taken to set up a business is used as a proxy for institutional quality and press freedom in a country. Technology use metrics for each country are obtained from the International Telecommunication Union. Socioeconomic data, such as economic indicators, institution quality, and country demographics, are largely obtained from the World Bank Database of World Development Indicators. The summary statistics and spatial distributions of these variables for each time period can also be found in Appendix 1 and 2.

5.2 Exploratory Spatial Data Analysis

We use a series of descriptive maps to describe the spatial patterns of internet users and fixed broadband subscriptions by country in 2004, 2009 and 2014. Figure 1 and 2 illustrate the spatial distribution of internet users (per 100 people) and spatial clusters of internet adoption in the world. Figure 1 shows a series of quintile maps of internet adoption for the different time periods. From Figure 1, we can visualize that countries in closer proximity tend to be in the similar quintiles (similar shades of colors). Over time, we also observe that countries with neighboring countries in high quintiles of internet adoption tend to enter higher quintiles in the next time period. For example, the majority of countries in Western Europe and Northern Europe are part of the highest quintile in internet adoption in 2004. In the next time period (2009), we observe an change in Eastern Europe countries such as Ukraine and Belarus from 20-40% percentile to the 40-60% percentile in internet adoption. Similar trends can also be found in North Asian countries bordering Russia.

at downstream speeds equal to, or greater than, 256 kbit/s. This includes cable modem, DSL, fiber-to-the-home/building, other fixed (wired)-broadband subscriptions, satellite broadband and terrestrial fixed wireless broadband. This total is measured irrespective of the method of payment. It excludes subscriptions that have access to data communications (including the Internet) via mobile-cellular networks. It should include fixed WiMAX and any other fixed wireless technologies. It includes both residential subscriptions and subscriptions for organizations. (ITU, 2016)

Figure 3 and 4 illustrate the spatial distribution of fixed broadband subscriptions (per 100 people) and spatial clusters of fixed broadband subscriptions in the world. While the adoption rates of internet usage and fixed broadband are vastly differently, the overall pattern in 2009 and 2014 is similar. The quintile map in 2004 has to be interpreted with caution as more than 50% of the countries have a reported fixed broadband subscriptions (per 100 people) of 0. This is unsurprising as fixed broadband was at its premature stages of adoption. There is also higher internet adoption than broadband adoption worldwide, this suggests that users do not only access the internet through fixed broadband, but also with more obsolete methods such as dial-up internet access or new-age mobile phones.

We further illustrate these spatial patterns by generating Local Moran's I cluster maps with 999 permutations. The Local Moran's I cluster maps allow us to categorize countries into statistically significant spatial clusters. Low-low areas or cold spots are countries with a below-average internet adoption rates surrounded by similar neighbors with a below-average internet adoption rates. High-high neighborhoods or hot spots are countries with above-average internet adoption rates surrounded by neighbors with an above-average internet adoption rates. Permutations allow us to determine how likely it is to find the actual spatial distribution of a set of values by comparing them with a set of randomly generated values. The more permutations conducted, the more precise the findings. Figure 2 shows the Local Moran's I cluster maps with queen contiguity weights[3] generated from country centroids[4] with 999 permutations. From Figure 2, we can observe a growing high-high cluster in Europe and a growing low-low cluster in Central and South Africa from 2004 to 2014, with the exception of the country South Africa. This exception could be attributed to consistently strong South African economy in the Sub-Saharan Africa region, thus this is unsurprising that South Africa is at the forefront of internet adoption in Africa. In the bottom map of Figure 2, South Africa is labelled as high-low country, which means it has above-average internet adoption rates and are surrounded by countries with

[3]The Local Moran's I cluster maps were also generated for the following spatial weights: three-nearest neighbors, four-nearest neighbors, five-nearest neighbors weights, and queen contiguity weights generated from country polygons in the world map. The clusters generated with other weights were similar to the clusters presented in Figure 2. The rest of the Local Moran's I cluster maps can be found in Appendix 3.

[4]Queen contiguity weights generated from the country centroids was selected as the most preferable spatial weights in our study. The justification for the selected spatial weights will be discussed later in Section 6.1.

below-average internet adoption rate. The spatial clusters found for broadband adoption are rather similar to internet adoption spatial clusters. In figure 4, a growing high-high cluster and low-low cluster can also be observed in Europe and Africa respectively.

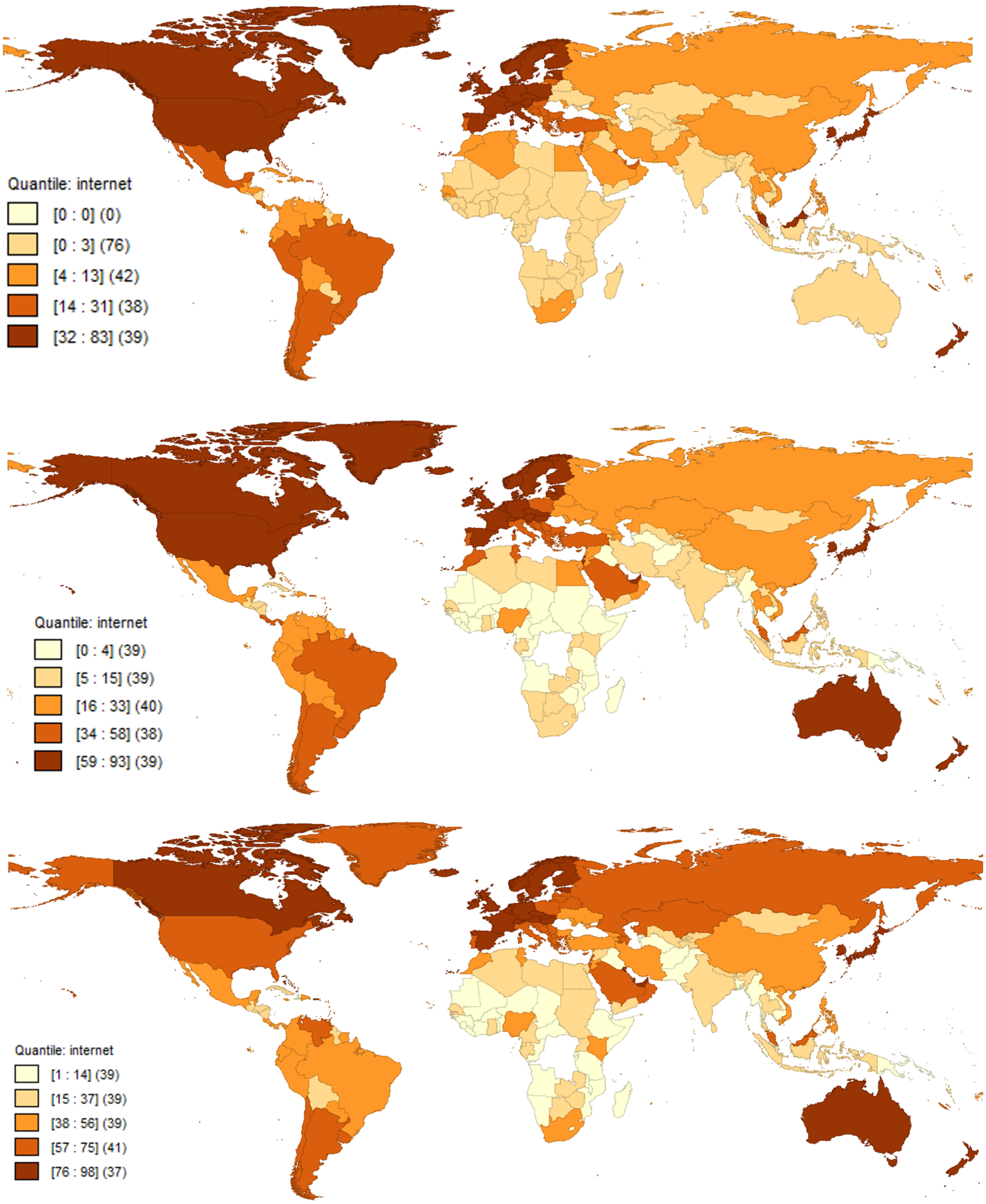


Figure 2: Quintile World Map of Internet Adoption for 2004 (Top) for 2009 (Middle) for 2014 (Bottom)

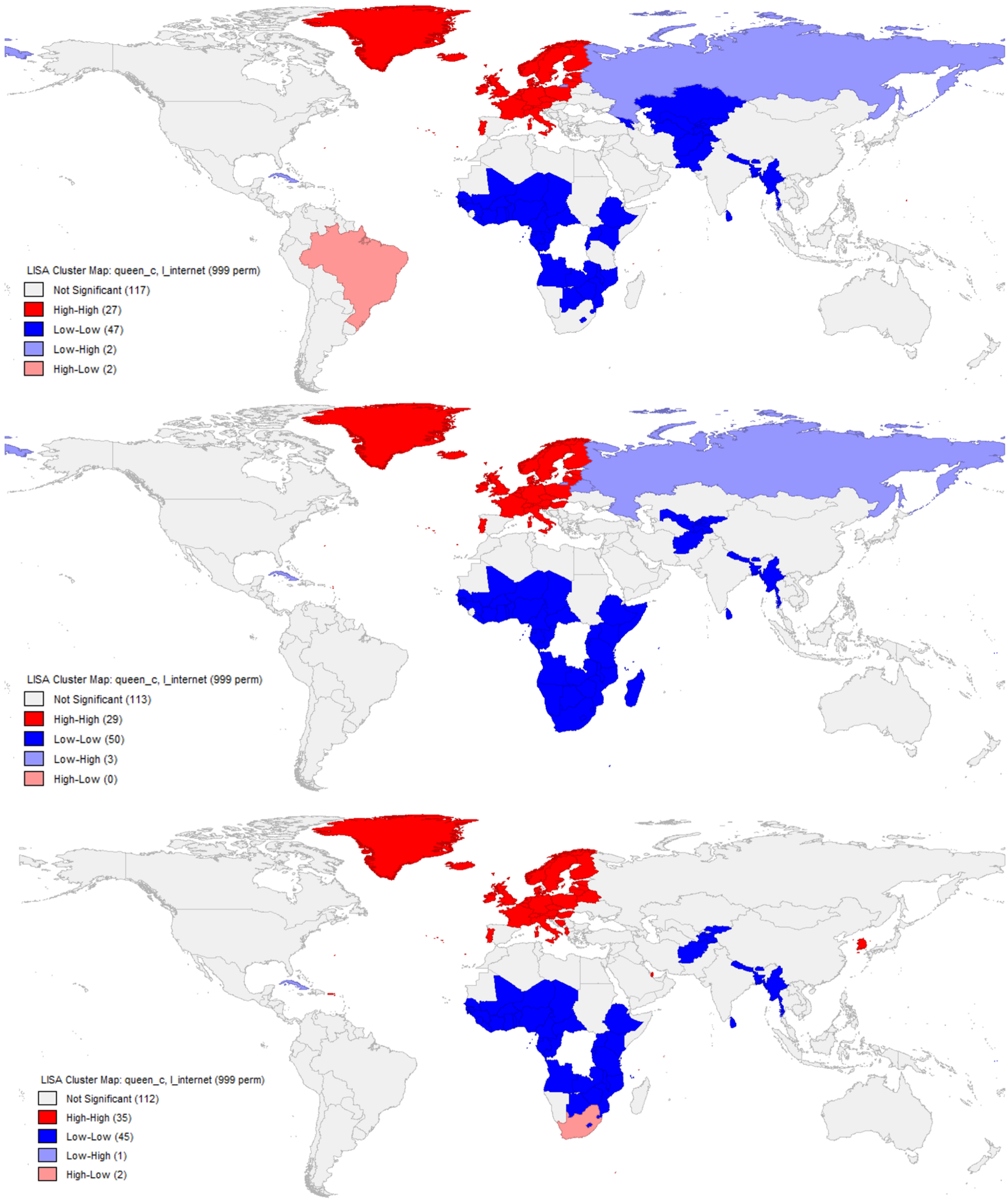


Figure 3: LISA Cluster Map of Internet Adoption in 2004 (top), 2009 (middle), and 2014 (bottom)

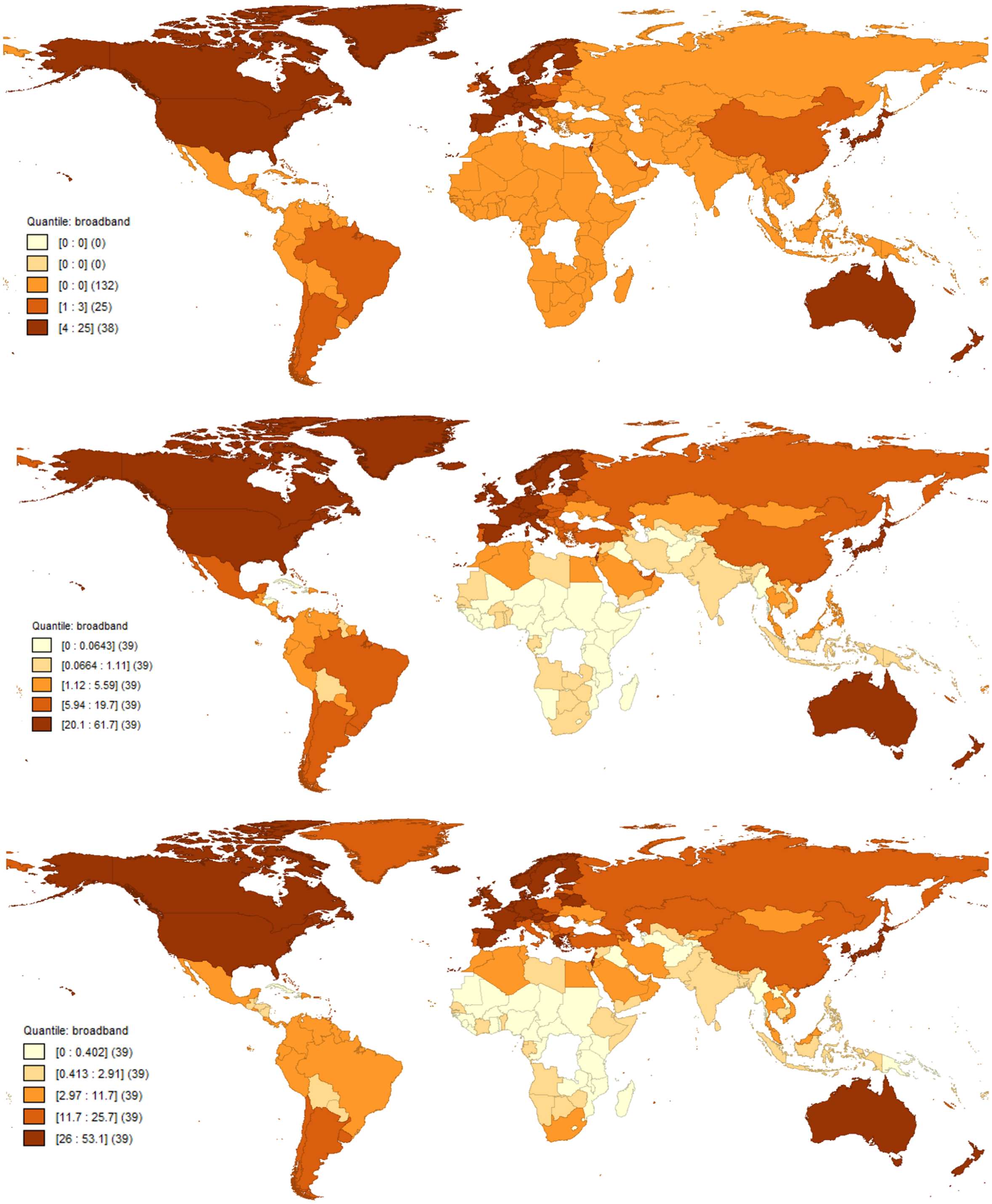


Figure 4: Quintile World Map of Broadband Subscription for 2004 (Top) for 2009 (Middle) for 2014 (Bottom)

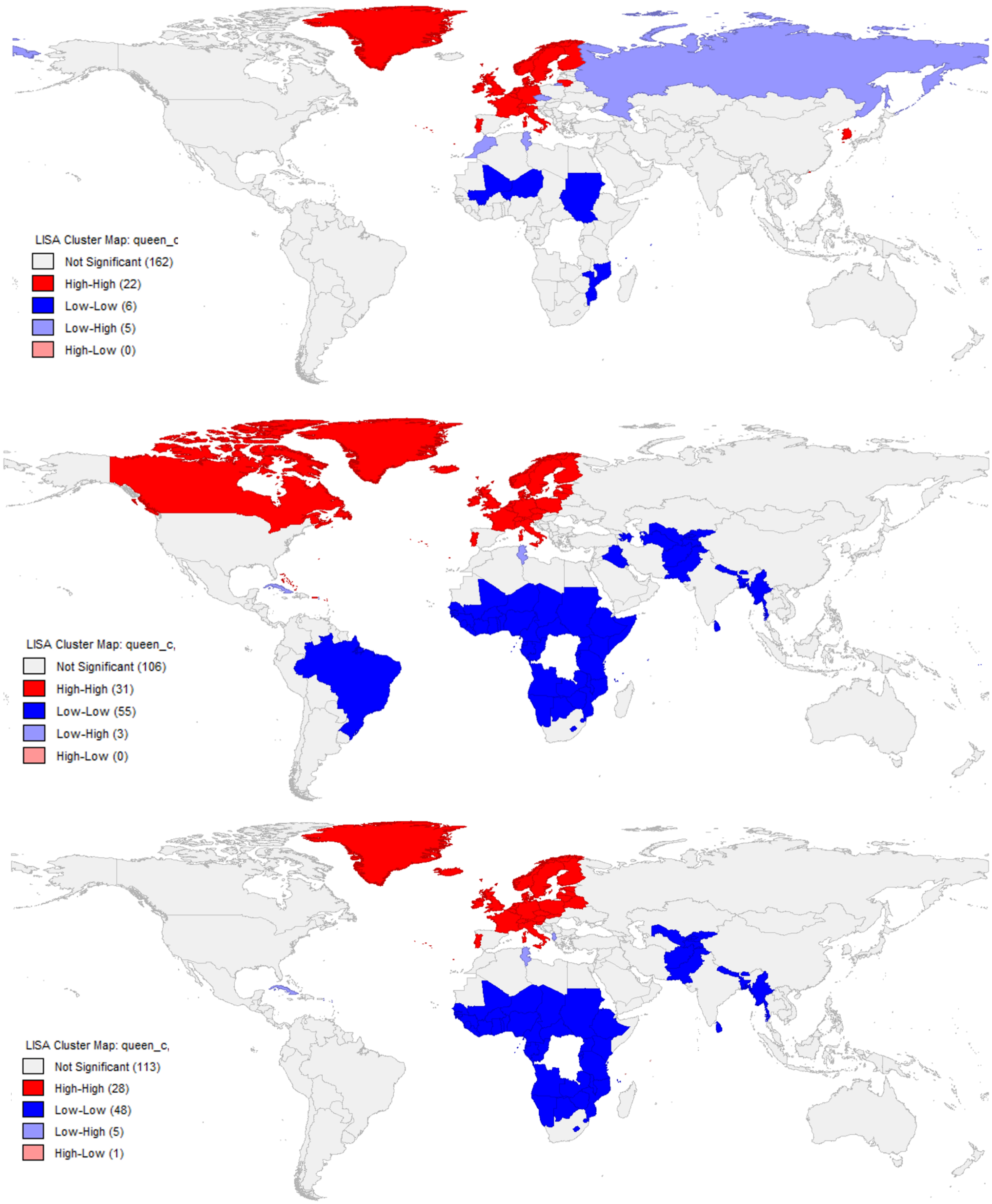


Figure 5: LISA Cluster Map of Broadband Subscription in 2004 (top), 2009 (middle), and 2014 (bottom)

6. Methodology

To account for spatial dependence, the traditional econometric methodology can be augmented with a spatial lag or a spatial error term according to the type of spatial externalities present. In the first specification, spatial dependence is conceptualized as an interaction between observational units. Observational units interact with each other and may unconsciously copy and experience peer effects in their decision-making process. Econometrically, this form of spatial dependence can be modelled as the dependent variable defined as a function of its value at neighboring locations. More explicitly, the spatial lag of the dependent variable encapsulates the “neighbor” or spatial effect, which is represented as Wy. (Anselin, A. Murray, and Rey 2013)

The spatial lag at location i is thus defined as:

$$Wy_i = \sum_{j=1}^{n} w_{ij} y_j$$

where W_{ij} is the spatial weights. Spatial weights are row standardized by convention, where $\sum_j W_{ij} = 1$. With row standardized weights, the spatial lag, Wy, is obtained as a weighted average of neighboring values. This regression specification, which includes a spatially lagged dependent variable, is known as the spatial autoregressive model or spatial lag model.

The second way spatial dependence can enter the regression specification is through the error term. The unobserved effects due to neighboring locations results in non-zero off-diagonal elements in the covariance structure of random error terms. Thus, spatial error autocorrelation is a case of non-spherical error variance-covariance, where $E[\epsilon_i \epsilon_j] \neq 0$ for $i \neq j$. When identified with a spatial error term, the spatial error dependence can be specified with a spatial stochastic process model. Examples of such models are a spatial moving average form (SMA), a spatial autoregressive form (SAR) and the conditional autoregressive model (CAR).

A typical spatial regression follows the specific form:

$$y = \rho W y + X\beta + u - (1)$$

where Wy is the spatial lag of the dependent variable, X is a vector of independent variables and u, the error term. Rather than having a normal error term, the second specification contains a spatially dependent error term. The normal error term is replaced by spatially dependent errors which are most commonly specified as the SAR form (2) or the SMA form (3)

$$u = \lambda W u + u - (2)$$

as the SAR form, or

$$u = \lambda W v + v$$

as the SMA form,

where Wu and Wv are the spatial lags of error term.

With the presence of spatial dependence, models are misspecified when they omit spatial lags or spatial error terms. (Anselin 1988) Similar to an omission of relevant variable, the missing spatial lags or error terms will result in biased estimates and misleading inferences (Anselin and Arribas-Bel 2011). Thus, the methodology of spatial econometrics involves testing for the potential presence of these misspecifications and using appropriate models to account for the spatial dependence.

6.1 Diagnostic Tests and Model Specification

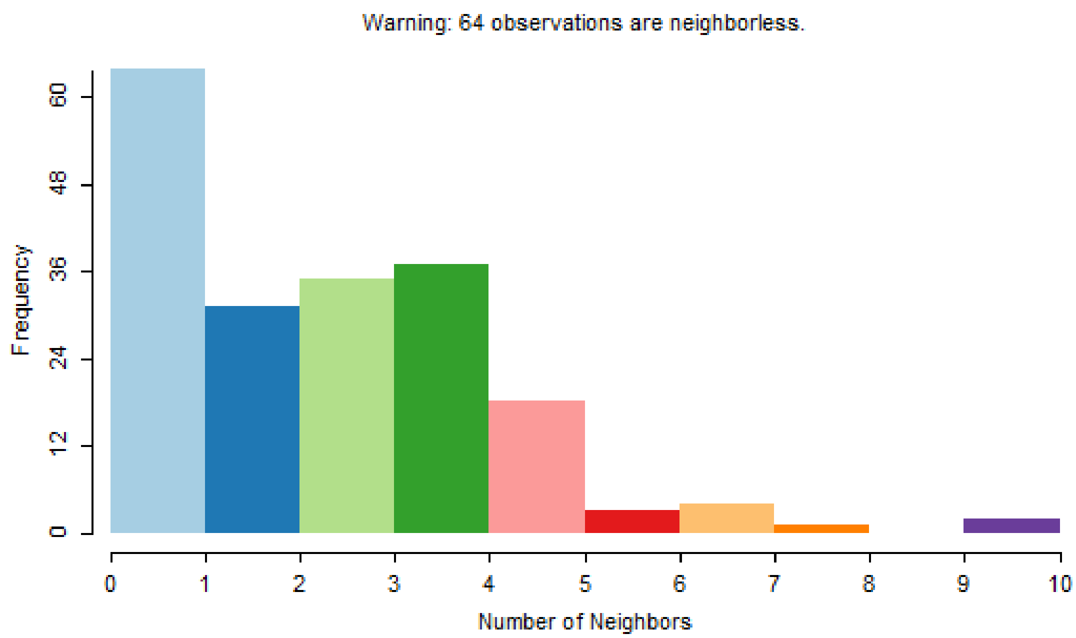
We will begin with performing an ordinary least squares (OLS) regression of the internet adoption on the independent variables. An appropriate spatial model is then selected, according to the OLS residuals and a series of diagnostics of spatial effects. The diagnostic tests consist of a series of Lagrange Multiplier tests against spatial lag dependence (LM-lag) or spatial error dependence (LM-Error) and the robust forms of those tests. The spatial models will use maximum likelihood (ML) estimation or 2SLS estimation, depending on the outcome of normality diagnostic test. If there is a strong evidence of non-normality, 2SLS estimation will be used.

Spatial Weights

Prior to conducting the diagnostics to determine the model specification, we formalize the notion of neighbors or countries that are spatially close with spatial weights, w_{ij}, where two spatial units i and j are a priori likely to interact. In our case, the spatial weights matrix is defined as a 195 by 195 positive matrix W with elements w_{ij}, where w_{ij} is non-zero for neighbors (likely to interact) and is zero if i and j are not neighbors (unlikely to interact). Spatial weights matrix can be constructed based on geography, distance and k-nearest neighbors. Geography-based weights is the most suitable for cross-country studies as spatial units i and j are considered as neighbors if they share a common border. In general, distance-based weights follow an inverse distance function, constructed based on distance between points or polygon centroids. However, there is a lack of literature that details the explicit distance cut-off suitable for internet adoption cross-country studies. Since the choice of spatial weights are crucial to the identification of neighbors, distance-based weights may not be suitable for our particular study. There are also similar issues with the selection of k in the construction of k-nearest neighbors weights. Furthermore, k-nearest neighbors enforces the same number of neighbors for all observations and may result in countries which are very far apart in distance to be clustered together. Thus, we will use row-standardized geography-based weights for our main model and conduct robustness checks with k-nearest neighbors weights (see Appendix 4).

There are three types of geography-based weights: rook contiguity weights, bishop contiguity weights and queen contiguity weights. Rook contiguity weights treat spatial units that touch on edges as neighbors. Bishop contiguity weights treat spatial units that touch on corners as neighbors. On the other hand, queen contiguity weights treat spatial units that touch on both edges and corners as neighbors. Queen contiguity weights is more suitable to deal with the idiosyncrasies of maps than the rook and bishop contiguity weights. Initially, we used queen contiguity weights to define spatially close countries (neighbors) as countries who share a common border. However, 32% of the countries were neighborless (isolates) due to missing countries in the World Bank World Development Index data set. This results in an uneven distribution of neighbor cardinality, which is showed by the frequency histogram of the neighbors each country has in Figure 6.

Figure 6: Frequency histogram of number of neighbors for queen contiguity weights



from	0.000	1.000	2.000	3.000	4.000	5.000	6.000	7.000	8.000	9.000
to	1.000	2.000	3.000	4.000	5.000	6.000	7.000	8.000	9.000	10.000
#obs	64	31	35	37	18	3	4	1	0	2
% of total	32.821	15.897	17.949	18.974	9.231	1.538	2.051	0.513	0.000	1.026
sd from mean	-0.446	0.000	0.122	0.691	1.259	1.827	2.395	2.963	3.531	4.100

min: 0, max: 9, median: 2, mean: 1.78462, s.d.: 1.76003, #obs: 195

To solve the problem of isolates due to missing data, we created our spatial weights with queen contiguity weights from the polygon centroids of each country. As shown in Figure 7, the result is a Thiessan polygon map laid over the centroids and queen contiguity weights generated from these polygons. In Figure 8, the histogram presents a much more symmetric and compact distribution of neighbor cardinalities. The median number of neighbors is 6 and the mean is 5.70.

Figure 7: Thiessan polygon map laid over centroids

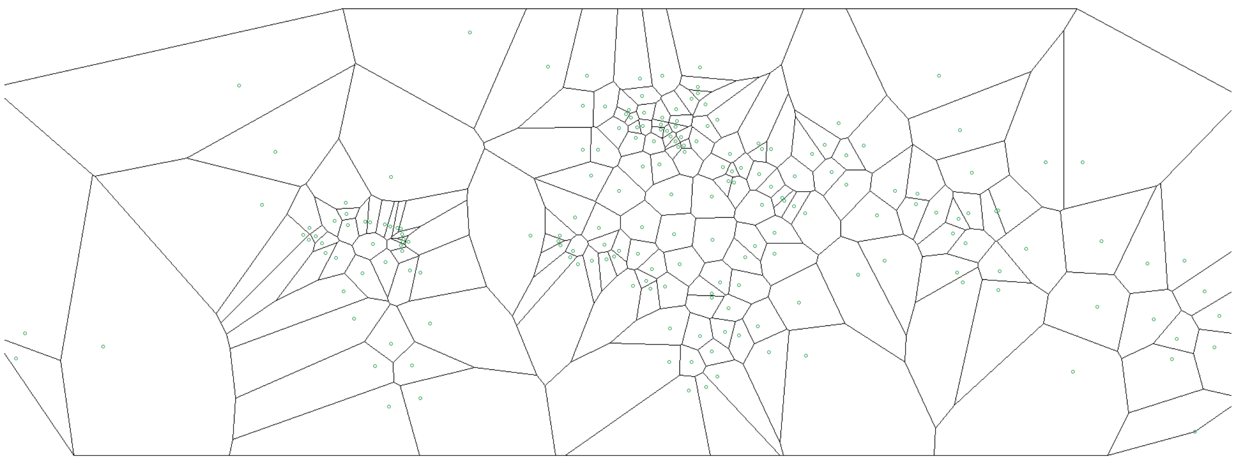
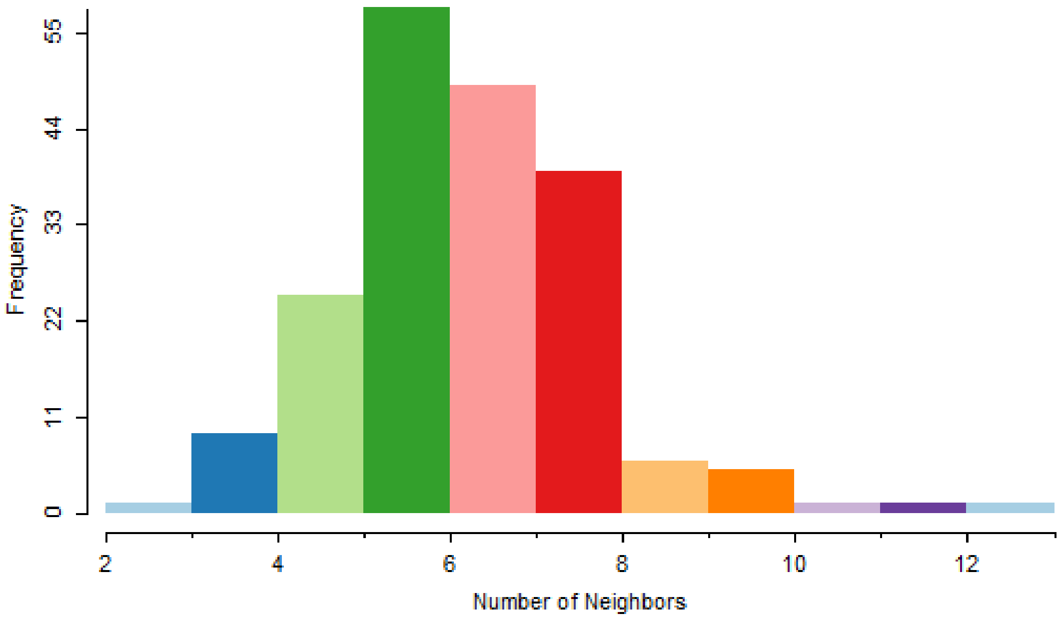


Figure 8: Frequency histogram of number of neighbors for queen contiguity weights based on centroids



from	2.000	3.000	4.000	5.000	6.000	7.000	8.000	9.000	10.000	11.000	12.000
to	3.000	4.000	5.000	6.000	7.000	8.000	9.000	10.000	11.000	12.000	13.000
#obs	1	9	25	58	49	39	6	5	1	1	1
% of total	0.513	4.615	12.821	29.744	25.128	20.000	3.077	2.564	0.513	0.513	0.513
sd from mean	-1.831	-1.153	-0.476	0.000	0.201	0.879	1.556	2.234	2.911	3.589	4.266

min: 2, max: 12, median: 6, mean: 5.70256, s.d.: 1.47618, #obs: 195

Non-Spatial Diagnostics

Table 3, 4, and 5 presents the non-spatial diagnostics results of the OLS regression for internet users in 2004, 2009, and 2014. As displayed in Table 2, the multicollinearity condition number for the three OLS regressions are 9.75, 9.53 and 9.98. Since the rule of thumb is to be concerned with multicollinearity condition numbers above 30, it suggests that there will not be any potential problems with multicollinearity with our specification. The Jarque-Bera test statistic for normality of errors is rejected at 1% significance level with a value of 142.3 and 37.4 for the regression for year 2004 and 2009. The Jarque-Bera test statistic for normality of errors is 3.487 for year 2014, thus the null hypothesis of normality of errors is not rejected at the 10% significance level.Thus, the Generalized Methods of Moments (GMM) estimation will be used for year 2004 and 2009, whereas the OLS regression for year 2014 will be estimated with the Maximum Likelihood Estimator. Two tests, Breusch-Pagan test and Koenker-Bassett test, were conducted to detect heteroskedasticity. Out of the three tests, only the Koenker-Bassett test is robust to non-normality of errors. The Breusch-Pagan test has a reported value of 184.4, 142.2, and 15.2 and the null hypothesis of homoskedasticity are rejected at a 5% significance level in 2004, 2009 and 2014. This is further corroborated with the reported values of Koenker-Bassett test for 2004, 2009 and 2014 OLS regressions of 59.7, 71.0, 12.4 which also rejects the null hypothesis of homoskedasticity at 10% significance level. Based on the corroboration of the previous two tests, there is strong evidence suggesting the assumption of homoskedasticity is violated. Thus, we will estimate the OLS regression with heteroskedasticity-consistent standard errors such as White standard errors.

Table 2: Multicollinearity tests for 2004, 2009 and 2014 OLS regressions

Year	Multicollinearity Condition Number
2004	9.75
2009	9.53
2014	9.98

Table 3: Non-spatial diagnostics for 2004 OLS regression

Test	DF	Value	Probability
Jarque-Bera	2	142.253	0.0000
Breusch-Pagan test	7	184.418	0.0000
Koenker-Bassett test	7	59.680	0.0000

Table 4: Non-spatial diagnostics for 2009 OLS regression

Test	DF	Value	Probability
Jarque-Bera	2	37.3541	0.00000

Test	DF	Value	Probability
Breusch-Pagan test	7	142.2652	0.00000
Koenker-Bassett test	7	71.0495	0.00000

Table 5: Non-spatial diagnostics for 2014 OLS regression

Test	DF	Value	Probability
Jarque-Bera	2	3.487	0.1749
Breusch-Pagan test	7	15.231	0.0332
Koenker-Bassett test	7	12.404	0.0880

Spatial Diagnostics

With the specification of a queen contiguity weights matrix, the Lagrange Multiplier statistics is calculated for the spatial lag, spatial error and higher orders of spatial autocorrelations with the OLS regression for three time periods. The diagnostics for spatial dependence in 2014, 2009, and 2004 are reported in Table 3, 4, and 5 respectively. The Lagrange Multiplier (lag) in the 2004 is highly significant at a 0.01% significance level, with a value of 17.4. This is further corroborated with the robust Lagrange Multiplier (lag) with a value of 8.14, which is also significant at a 0.01% significance level. Similarly, the Lagrange Multiplier (lag) and robust Lagrange Multiplier (lag) for the cross section 2009 are also significant at 0.01 significance level with a value of 23.8 and 18.2 respectively. The Lagrange Multiplier (lag) and robust Lagrange Multiplier (lag) in the 2014 are also highly significant at a 0.01% significance level, with a value of 23.8 and with a value of 24.1 respectively. The results of the spatial autocorrelation diagnostics for all three time periods support the presence of spatial autocorrelation.

The Lagrange Multiplier (error) test for the presence of spatial error autocorrelation. However, the rejection of the null hypothesis of LM (error) test might be due to not only the presence of spatial error autocorrelation, but also the presence of spatial autocorrelation. Thus, the robust Lagrange Multiplier (error) test attempts to disentangle this effect and only identify the presence of spatial error autocorrelation. For the cross-section 2004, the Lagrange Multiplier (error) test has a reported value of 9.95 and is significant at a 5% significance level, but the robust Lagrange Multiplier (error) test has a value of 0.719 and is highly insignificant. Thus, the statistic has moved from significant to insignificant. This suggests that the Lagrange Multiplier (error) test was most plausibly picking up the error spatial autocorrelation due to the spatial lag, rather than actually having spatial error autocorrelation in its model. Similarly in 2009, the Lagrange Multiplier (error) test has a reported value of 5.58 and is significant at a 5% significance level, but the robust Lagrange Multiplier (error) test has a value of 0.059 and is highly insignificant. Thus, the diagnostics report that spatial error autocorrelation is absent for the OLS regression for year 2009 as well. The diagnostics for spatial error autocorrelation in 2014 were insignificant for both Lagrange Multiplier (error) and robust Lagrange Multiplier at a 10% significance, thus spatial error autocorrelation was absent as well.

Finally, the last diagnostic, Lagrange Multiplier (SARMA), is a test for higher order spatial autocorrelation. The value of the final diagnostic test is highly significant for all three OLS regressions. The results suggest that it is probable that a higher order model specification is the correct alternative model. We have to interpret the results of the Lagrange Multiplier (SARMA) with extra caution as there is two degrees of freedom in the test. The presence of two degrees of freedom could increases the likelihood of rejecting the joint null hypothesis when one marginal test has a very high value. Therefore, we would only consider higher order models if residual spatial autocorrelation is still present in the model with a single spatial parameter.

Table 6: Diagnostics for spatial dependence for 2004 cross section with queen contiguity weights

Test	MI/DF	Value	Probability
Lagrange Multiplier (lag)	1	17.368	0.0000
Robust LM (lag)	1	8.138	0.0043
Lagrange Multiplier (error)	1	9.950	0.0016
Robust LM (error)	1	0.719	0.3964
Lagrange Multiplier (SARMA)	2	18.087	0.0001

Table 7: Diagnostics for spatial dependence for 2009 cross section with queen contiguity weights

Test	MI/DF	Value	Probability
Lagrange Multiplier (lag)	1	23.7663	0.00000
Robust LM (lag)	1	18.2413	0.00002
Lagrange Multiplier (error)	1	5.5841	0.01812
Robust LM (error)	1	0.0591	0.80799
Lagrange Multiplier (SARMA)	2	23.8254	0.00001

Table 8: Diagnostics for spatial dependence for 2014 cross section with queen contiguity weights

Test	MI/DF	Value	Probability
Lagrange Multiplier (lag)	1	23.8078	0.00000
Robust LM (lag)	1	24.0991	0.00000
Lagrange Multiplier (error)	1	1.9404	0.16363
Robust LM (error)	1	2.2316	0.13521
Lagrange Multiplier (SARMA)	2	26.0395	0.00000

6.2 Model Specification

Considering the presence of spatial autocorrelation, our model is specified with a spatial lag with its independent variables from the OLS regression:

$$internet_i = \alpha + X\beta + Winternet_i + \epsilon_i - (3)$$

where α is a constant for all observations, X is a vector of control variables and β is the estimated parameter, $Winternet$ is the spatial lag of dependent variable (percentage of internet users in a country), and ϵ is the error term. The control variables included in our OLS regression are GDP per capita, proportion of telephone fixed lines subscribers, number of days to set up a business, primary school enrollment rates, proportion of population aged above 65, percentage of urbanization and a dummy variable for developed countries. Since the weights used in our model is row-standardized queen contiguity weights, the spatial lag of the model is average of internet adoption of the neighbors of a given country. However, the presence of the spatial lag of the dependent variable could induce endogeneity or simultaneity, and violate the assumption of exogenous regressors in OLS regression. In order to correct for potential problems of endogeneity, we perform two stage least squares estimation to the model and instrument the spatial lag with spatially lagged explanatory variables, which are also known by the literature to be the most appropriate instruments.

7. Results

Table 9 displays the results of the instrumental variable regression with control variables. We find that internet adoption is positively affected by adoption rates in countries in close proximity. This result is robust over different years and also holds when a rich set of other potential determinants of internet adoption are included as control variables. We consider the consistent and high significance of the spatial lag as strong evidence pointing to the significance of spatial proximity (neighborhood) in determining internet adoption rates. As previously discussed in Section 3, there are good theoretical basis for internet adoption in countries in close proximity to be an important determinant of internet adoption. Good neighbors create positive spillovers, whereas 'bad neighbors' create less or none. These spillovers take not only the form of knowledge spillovers, but also extends to attitudes, beliefs, and social norms that may be facilitated by neighboring countries being a reference group. For instance, risk-taking and technical progress can be characterized as positive values with a spread of social norms from neighboring countries. Furthermore, the coefficient of $W_{internet}$ is increasing over the years, suggesting the growing importance of neighboring countries' internet adoption patterns on a given country. The coefficient of the spatial lag increased from 0.248 in 2004, to 0.337 in 2009 and 0.405 in 2014. To put this into perspective, an unit percentage increase in the average internet adoption rate in neighboring countries results in 0.248 increase in internet adoption rates in the given country in 2004. In 2014, a unit increase in the average internet adoption rate in neighboring countries now result in 0.405 increase in internet adoption rates in a given country. This is unsurprising as the internet is a technology with positive network externalities. The presence of network externalities suggests that the greater number of internet users, the greater the incentive to adopt the internet. Thus, the impact of spatial spillovers are likely to grow larger and larger as more people adopt the internet.

An example of a good neighborhood is Europe. Countries in Western Europe already had above average internet adoption rates and had neighboring countries with above average internet adoption rates in 2004, forming a hot-hot cluster. In 2009, more countries from Central Europe joined the hot-hot cluster such as Croatia, Bosnia and Herzegovina and Poland. In 2014, the hot-hot cluster included Eastern European countries such as Greece, Turkey, Ukraine and Serbia. Besides obvious examples of good neighbors, a group of neighboring post-Soviet states[5] also shows impact of a good neighborhood among developing countries. As the World Bank reports, there has been a series of national infrastructure plans released in Central Asian countries to leverage on digital technologies and promote ICT adoption. On the other hand, a bad neighborhood is characterized by

[5]The group of post-Soviet states include Kazakhstan, Kyrgyzstan Republic, Tajikistan and Uzbekistan.

the lack of positive spillovers. Sub-Saharan Africa is an example of a bad neighborhood, with underdeveloped infrastructure, low income per capita, and a lack of human capital investments (Bank 2016). For instance, the bad neighborhood of Sub-Saharan Africa has hindered internet adoption in Egypt and Swaziland, despite them having higher income per capita than Guyana and Bolivia in South America from 2004 to 2014.

Table 9: 2SLS regression for internet adoption in 2014, 2009 and 2004

internet	2014	2009	2004
Constant	1.37	-6.686***	-4.579***
	(2.94)	(1.877)	(1.383)
W_internet	0.405***	0.337***	0.248*
	(0.073)	(0.0766)	(0.116)
gdp	0.000211***	0.000295**	0.000410**
	(0.0000444)	(0.0000965)	(0.000129)
tele	0.318***	0.286*	0.295**
	(0.0812)	(0.136)	(0.123)
urban	0.129**	0.121**	0.0296
	(0.0489)	(0.0438)	(0.0362)
old	0.272	0.775**	0.771**
	(0.239)	(0.301)	(0.317)
pri_rate	0.180***	0.100	-0.0483
	(0.0382)	(0.0558)	(0.0757)
bus_days	-0.0910**	-0.0134	-0.0128
	(0.0356)	(0.00908)	(0.0139)
developed	14.714***	7.73**	1.374
	(2.787)	(2.44)	(1.748)
N	195	195	195
R^2	0.8239	0.8345	0.7833

Standard errors in parentheses $*$ p $<$ 0.05, $**$ p $<$ 0.01, $***$ p $<$ 0.001

8. Robustness Checks

We also validated the robustness of the presence of spatial dependence in internet adoption by testing for the presence of spatial dependence in fixed broadband subscriptions. Table 10, 11, 12 shows the results of the spatial dependence diagnostic tests in 2004, 2009, and 2014 for fixed broadband subscriptions. With the exception of the year 2004, the diagnostic test results show that there is a presence of a spatial lag in 2009 and 2014 at a 5% significance level. This corroborates the previous findings presented in Section 6.1. Validating the presence of spatial dependence with more than one metric for internet adoption reaffirms our hypothesis that internet adoption in a country to be positively affected by internet adoption of other countries in close proximity, even after controlling for other usual determinants of internet adoption.

Table 10: Diagnostics for spatial dependence in broadband adoption for 2004 cross section with queen contiguity weights

Test	MI/DF	Value	Probability
Moran's I (error)	0.0986	2.6433	0.00821
Lagrange Multiplier (lag)	1	2.8607	0.09077
Robust LM (lag)	1	0.0564	0.81230
Lagrange Multiplier (error)	1	5.0036	0.02529
Robust LM (error)	1	2.1993	0.13807
Lagrange Multiplier (SARMA)	2	5.0600	0.07966

Table 11: Diagnostics for spatial dependence in broadband adoption for 2009 cross section with queen contiguity weights

Test	MI/DF	Value	Probability
Moran's I (error)	0.0659	1.9082	0.05637
Lagrange Multiplier (lag)	1	10.0735	0.00150
Robust LM (lag)	1	7.7388	0.00540
Lagrange Multiplier (error)	1	2.3354	0.12646
Robust LM (error)	1	0.0007	0.97916
Lagrange Multiplier (SARMA)	2	10.0742	0.00649

Table 12: Diagnostics for spatial dependence in broadband adoption for 2014 cross section with queen contiguity weights

Test	MI/DF	Value	Probability
Moran's I (error)	-0.0004	0.3056	0.75988
Lagrange Multiplier (lag)	1	11.6964	0.00063
Robust LM (lag)	1	15.4939	0.00008
Lagrange Multiplier (error)	1	0.0001	0.99282
Robust LM (error)	1	3.7976	0.05133
Lagrange Multiplier (SARMA)	2	15.4940	0.00043

9. Conclusion

Our findings of our study is two fold. Firstly, internet adoption is positively affected by internet adoption rates of countries in close proximity. The significance of countries in close proximity stems from the presence of spatial spillovers, which we define more broadly as both knowledge spillovers and the formation of attitudes, beliefs and social norms. Our results suggest that countries, especially developing countries, could be caught in a "information trap" where there is a lack of positive spillovers and a dominance of negative spillovers. Thus, it will be arduous for these countries to escape this trap on their own without sudden improvements in economics, infrastructure and institutions. Our results illuminates an alternative perspective to be considered in development aid distribution for ICT adoption. On top of usual considerations of aid distribution, international policies supporting internet adoption can be more effective if they target clusters of countries in close proximity rather than single countries. Furthermore, since internet adoption varies geographically within countries as well, a more precise analysis can be conducted with more granular data specifying internet adoption by agglomerations or sub-regions in the future. This will also allow better definitions of neighborhoods within and across countries.

Secondly, another finding of our paper is methodological. The results of our study point to the relevance of spatial econometrics in internet adoption, and possibly technology adoption. As explained in Section 3, the concept of knowledge spillovers in technology adoption is not novel, but the more intangible definitions of spillovers - in the forms of attitudes, beliefs and social norms - have yet to be explored. These spillovers, both tangible and intangible, can be accounted with suitable metrics with a spatial econometric methodology. For instance, an individual's likelihood to accept technologies (Gaskell et al. 2010) can be modeled with a spatial econometric regression model to account for possible spatial spillovers, the presence of a spatial lag or spatial error could then suggest the importance of other countries in the formation of attitudes towards technology in one country. Thus, spatial econometrics can augment present econometric models for technology adoption to model for spillovers and facilitate a more thorough analysis of spillovers in technology adoption.

Appendix

Appendix A

Internet	GDP per capita	Telephone lines	Old
Min. : 1.38	Min. : 286	Min. : 0.000	Min. : 1.039
1st Qu.:17.73	1st Qu.: 1870	1st Qu.: 3.101	1st Qu.: 3.502
Median :46.20	Median : 5484	Median : 13.356	Median : 5.884
Mean :46.08	Mean : 15829	Mean : 18.604	Mean : 8.237
3rd Qu.:70.11	3rd Qu.: 18595	3rd Qu.: 29.116	3rd Qu.:12.702
Max. :98.16	Max. :178713	Max. :132.953	Max. :25.705
NA	NA's :14	NA	NA's :10

Urbanization	Days to Business
Min. : 8.55	Min. : 0.50
1st Qu.: 39.31	1st Qu.: 8.10
Median : 58.66	Median : 14.00
Mean : 58.18	Mean : 22.68
3rd Qu.: 77.22	3rd Qu.: 28.00
Max. :100.00	Max. :144.00
NA	NA's :16

Table A.1: Summary Statistics of Variables in 2014

Internet	GDP per capita	Telephone lines	Old
Min. : 0.22	Min. : 190.4	Min. : 0.01851	Min. : 0.6987
1st Qu.: 7.30	1st Qu.: 1227.6	1st Qu.: 3.61227	1st Qu.: 3.4148
Median :26.00	Median : 4576.3	Median : 16.31757	Median : 5.6907
Mean :31.24	Mean : 14211.5	Mean : 21.12441	Mean : 7.6307
3rd Qu.:50.80	3rd Qu.: 16772.2	3rd Qu.: 32.58993	3rd Qu.:11.3422
Max. :93.00	Max. :152877.4	Max. :118.90455	Max. :22.2150
NA's :2	NA's :7	NA's :2	NA's :10

Urbanization	Days to Business
Min. : 9.249	Min. : 0.50
1st Qu.: 36.941	1st Qu.: 12.00
Median : 56.608	Median : 20.00
Mean : 56.718	Mean : 35.64
3rd Qu.: 75.533	3rd Qu.: 39.00
Max. :100.000	Max. :690.50
NA	NA's :20

Table A.2: Summary Statistics of Variables in 2009

Internet	GDP per capita	Telephone lines	Old
Min. : 0.00	Min. : 0.0	Min. : 0.00	Min. : 0.000
1st Qu.: 1.00	1st Qu.: 699.9	1st Qu.: 3.20	1st Qu.: 3.173
Median : 7.00	Median : 2649.9	Median : 13.57	Median : 4.875
Mean :17.59	Mean : 10427.5	Mean : 20.98	Mean : 6.886
3rd Qu.:27.00	3rd Qu.: 10502.0	3rd Qu.: 33.09	3rd Qu.:10.158
Max. :83.00	Max. :123289.5	Max. :101.50	Max. :19.294

Urbanization	Days to Business
Min. : 9.00	Min. : 0.0
1st Qu.: 35.00	1st Qu.: 4.0
Median : 55.00	Median : 31.0
Mean : 54.76	Mean : 35.9
3rd Qu.: 73.50	3rd Qu.: 47.5
Max. :100.00	Max. :202.0

Table A.3: Summary Statistics of Variables in 2004

Appendix B

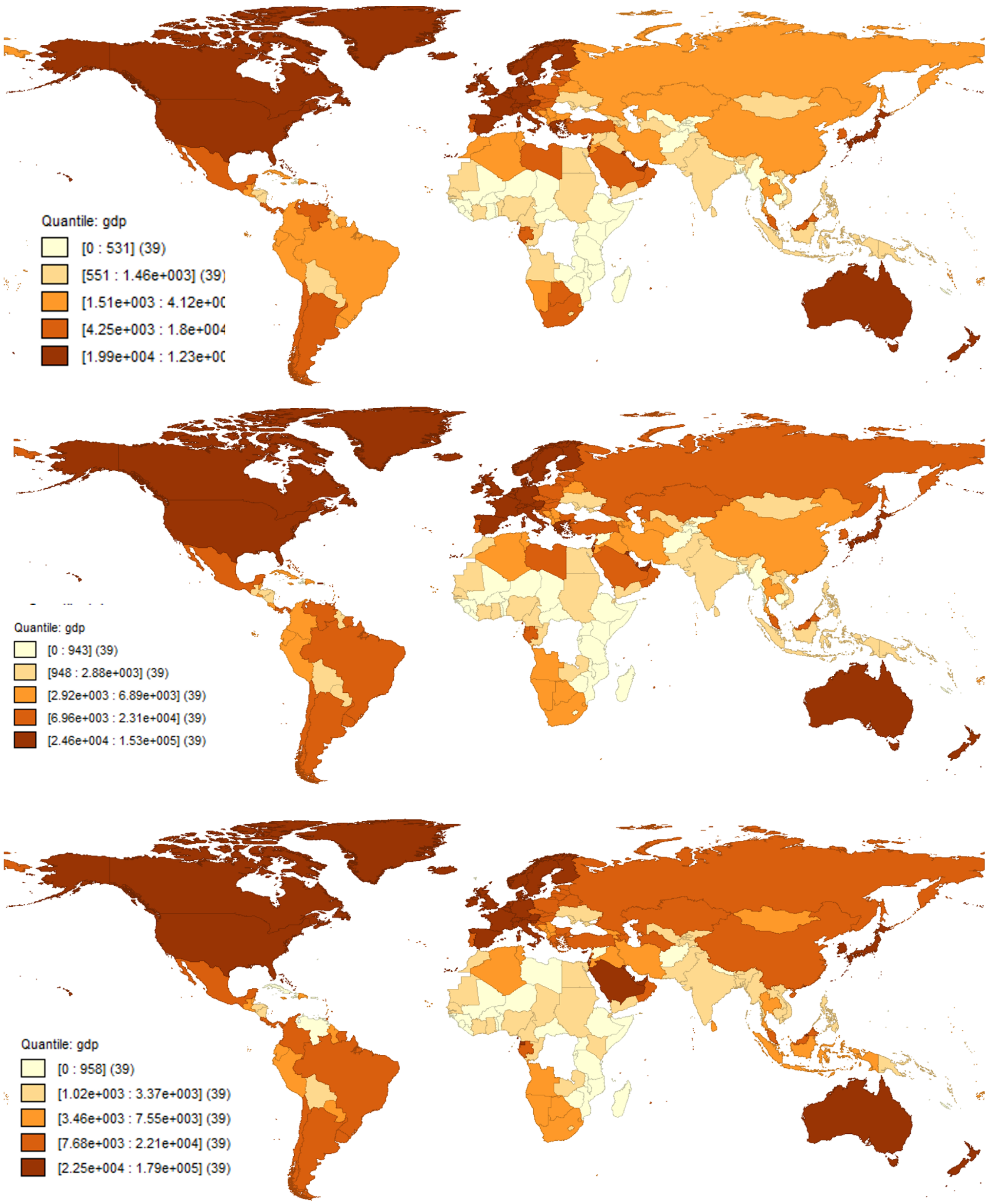


Figure B.1: Quintile World Map of GDP per capita in 2004 (Top), 2009 (Middle), and 2014 (Bottom)

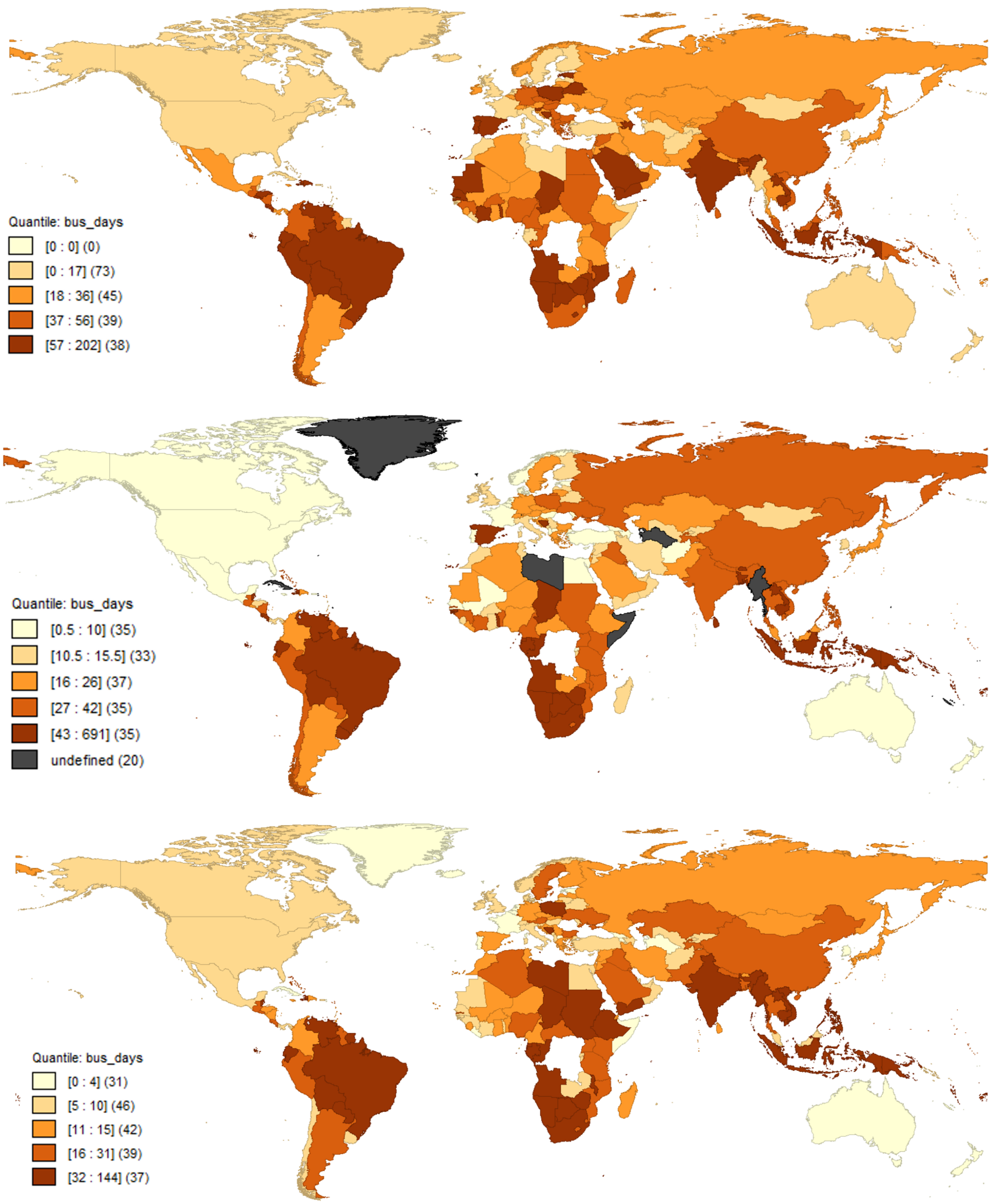


Figure B.2: Quintile World Map of Number of Days to Set Up a Business in 2004 (Top), 2009 (Middle), and 2014 (Bottom)

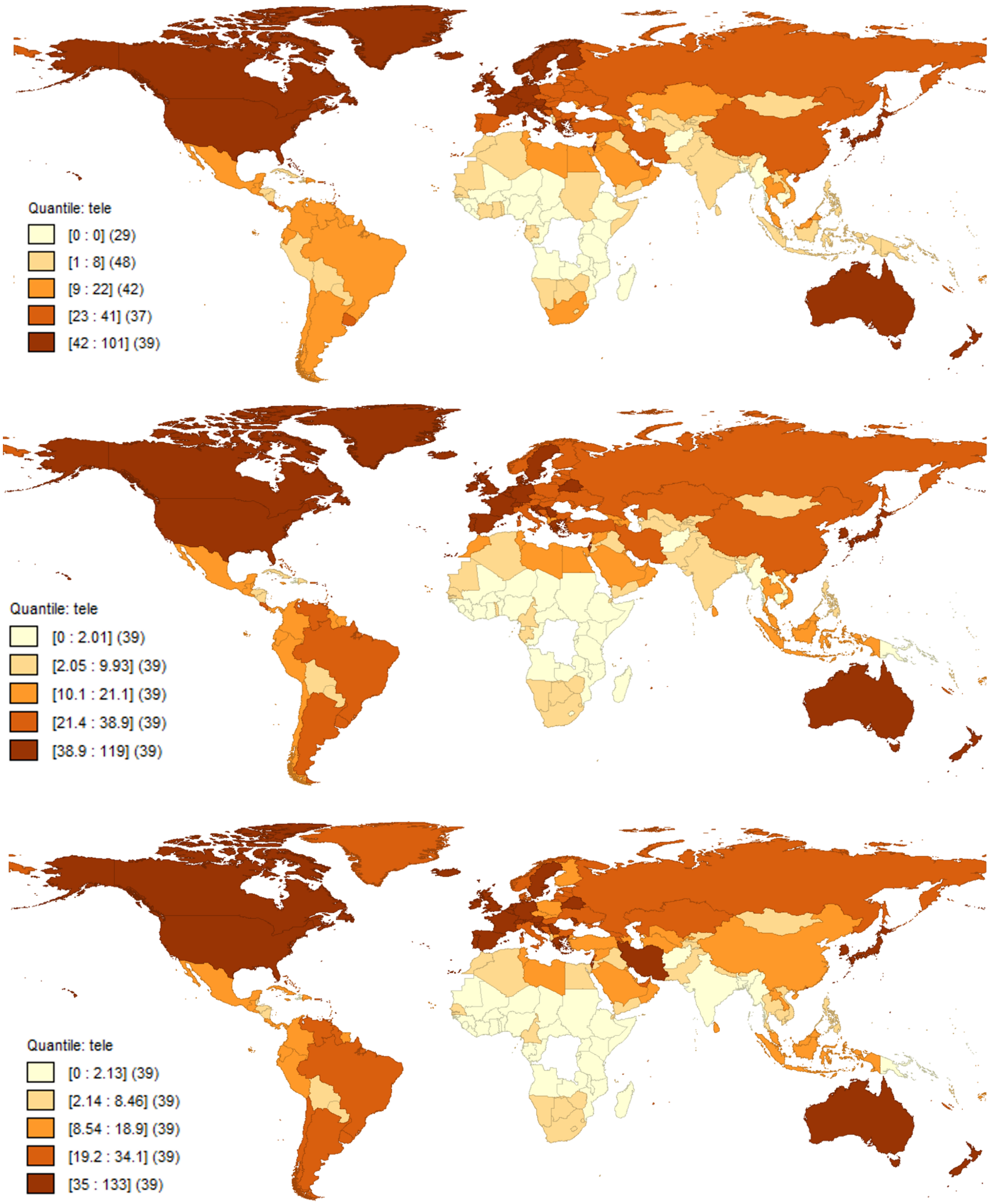
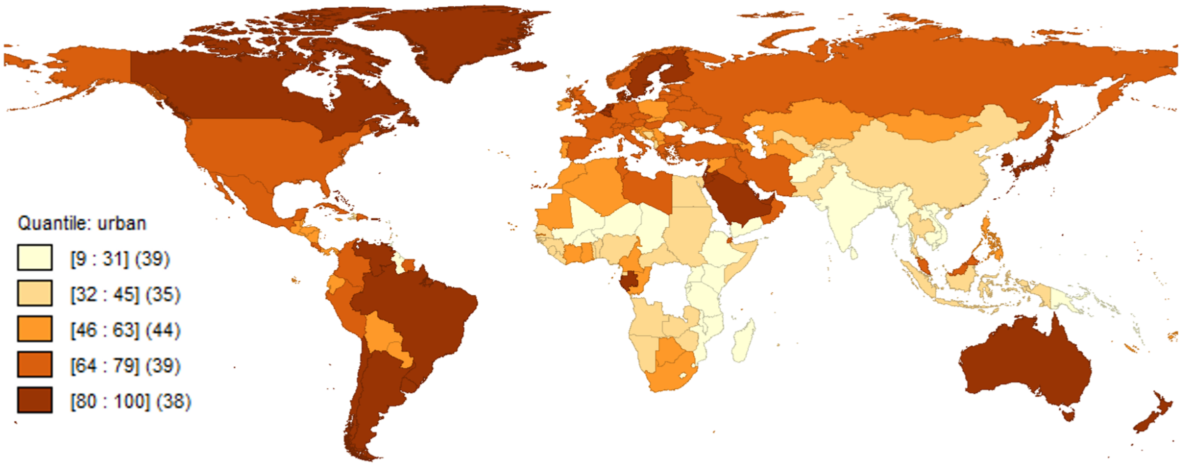
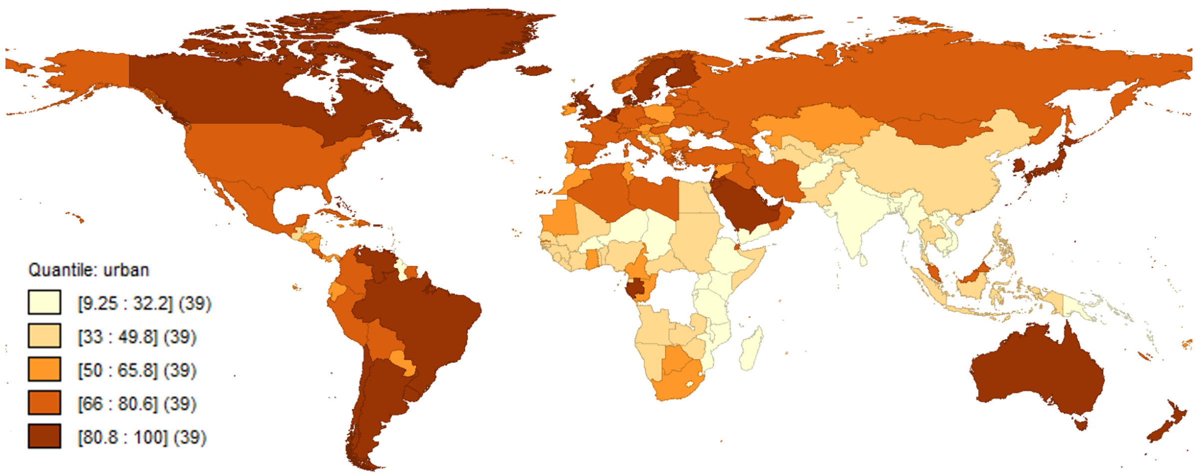


Figure B.3: Quintile World Map of Proportion of Telephone Fixed Line Subscribers in 2004 (Top), 2009 (Middle), and 2014 (Bottom)





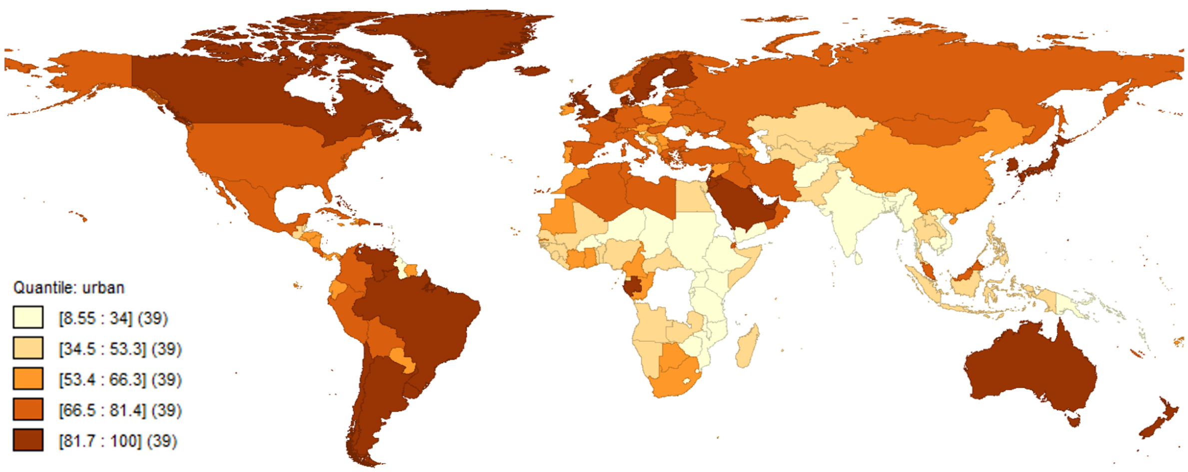


Figure B.4: Quintile World Map of Proportion of Urbanized Areas in 2004 (Top), 2009 (Middle), and 2014 (Bottom)

Appendix C

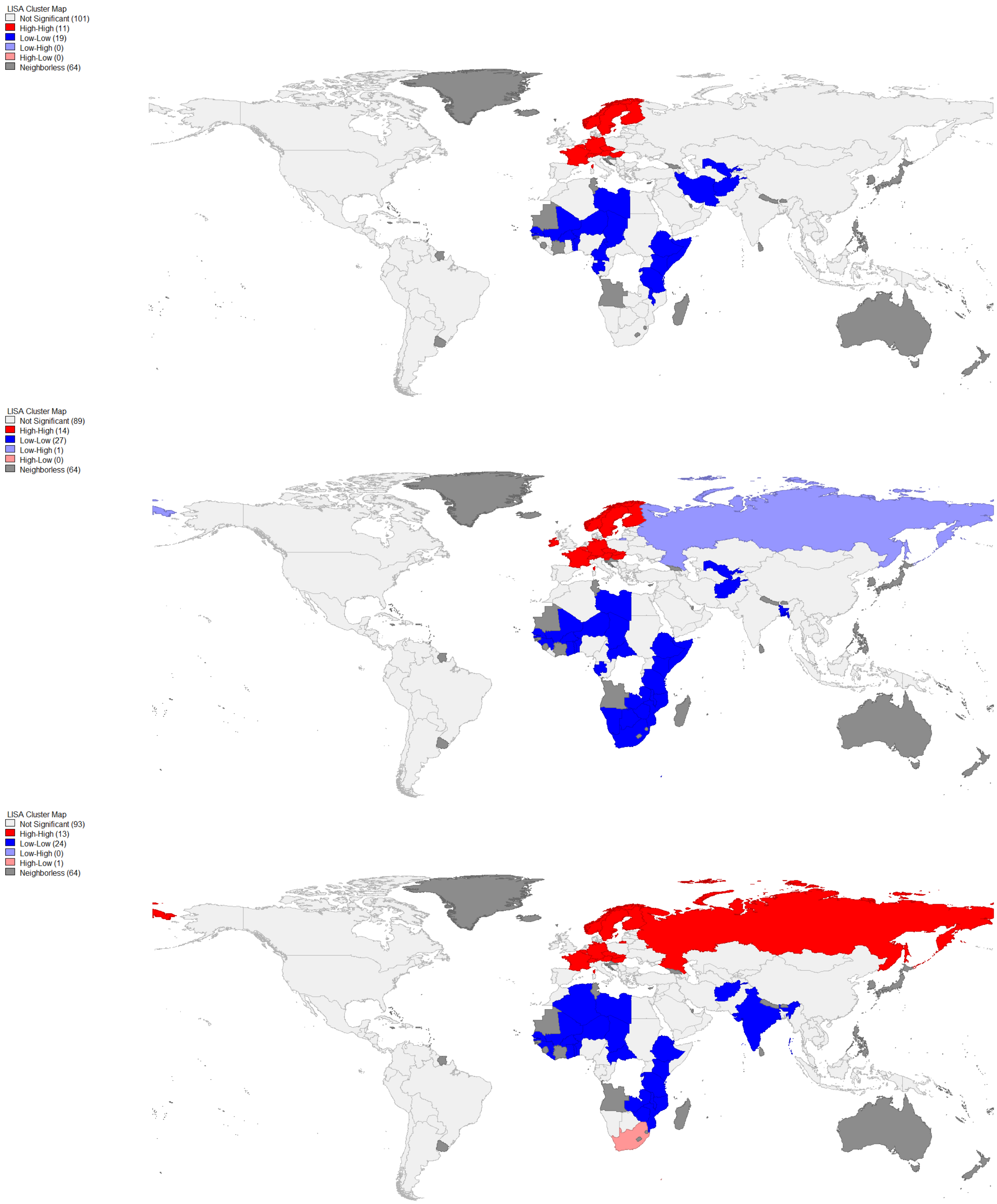


Figure C.1: LISA Cluster Map of Internet Adoption with queen continguity weights in 2004 (top), 2009 (middle), and 2014 (bottom)

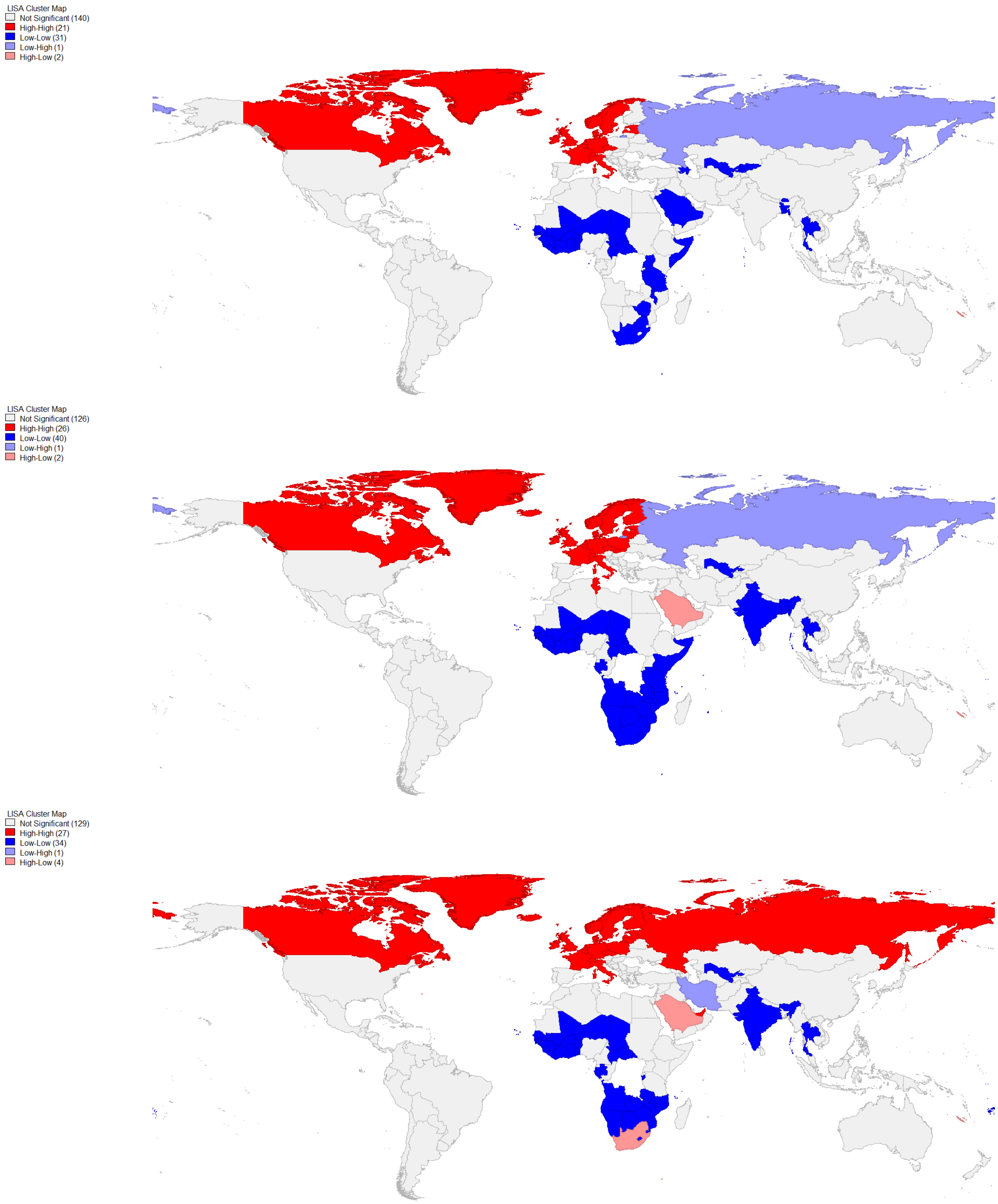


Figure C.2: LISA Cluster Map of Internet Adoption with 3-nearest neighbors spatial weights in 2004 (top), 2009 (middle), and 2014 (bottom)

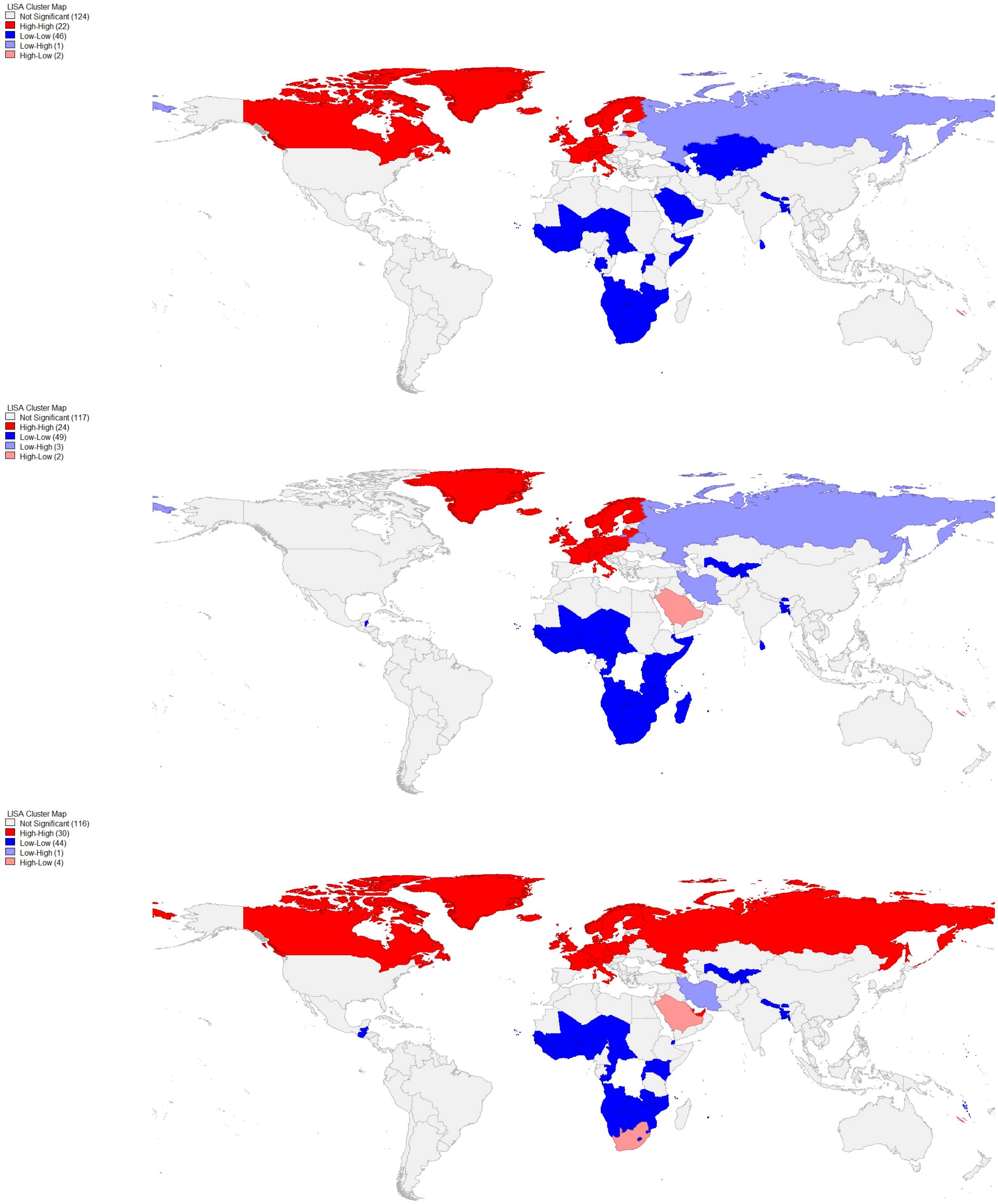


Figure C.3: LISA Cluster Map of Internet Adoption with 4-nearest neighbors spatial weights in 2004 (top), 2009 (middle), and 2014 (bottom)

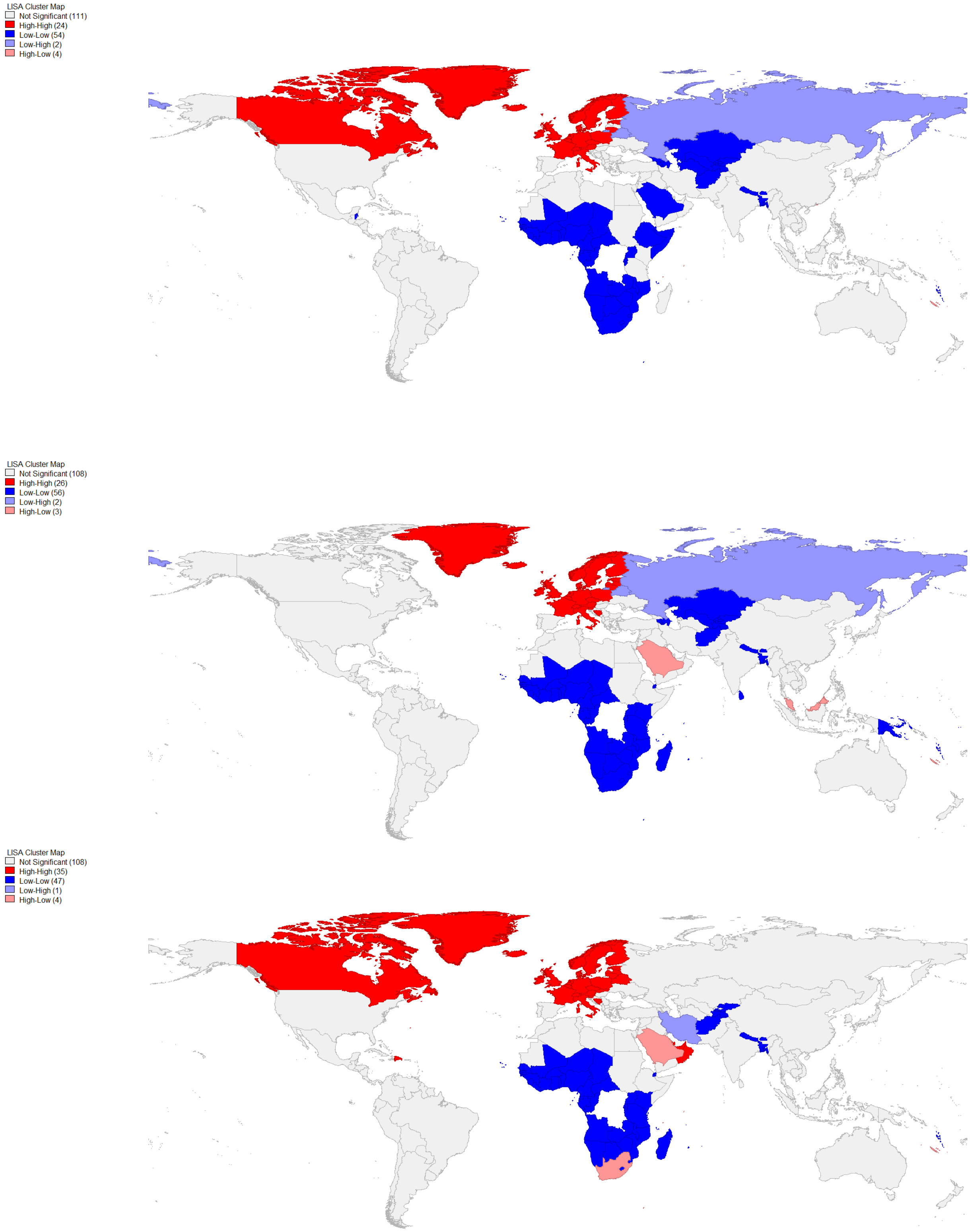


Figure C.4: LISA Cluster Map of Internet Adoption with 5-nearest neighbors spatial weights in 2004 (top), 2009 (middle), and 2014 (bottom)

Appendix D

Test	MI/DF	Value	Probability
Moran's I (error)	0.1234	3.4519	0.00056
Lagrange Multiplier (lag)	1	8.0569	0.00453
Robust LM (lag)	1	2.4314	0.11893
Lagrange Multiplier (error)	1	8.3026	0.00396
Robust LM (error)	1	2.6770	0.10181
Lagrange Multiplier (SARMA)	2	10.7340	0.00467

Table XX: Spatial Diagnostics Test for 2014 OLS regression with 5-nearest neighbors weights

Test	MI/DF	Value	Probability
Moran's I (error) 0.0123	0.5922	0.55373	
Lagrange Multiplier (lag) 1	16.8850	0.00004	
Robust LM (lag) 1	20.5635	0.00001	
Lagrange Multiplier (error) 1	0.0677	0.79475	
Robust LM (error) 1	3.7462	0.05293	
Lagrange Multiplier (SARMA) 2	20.6312	0.00003	

Table XX: Spatial Diagnostics Test for 2014 OLS regression with 4-nearest neighbors weights

Test	MI/DF	Value	Probability
Moran's I (error)	0.0492	1.2178	0.22332
Lagrange Multiplier (lag)	1	14.0924	0.00017
Robust LM (lag)	1	14.8107	0.00012
Lagrange Multiplier (error)	1	0.8233	0.36423

Robust LM (error)	1	1.5416	0.21438
Lagrange Multiplier (SARMA)	2	15.6340	0.00040

Table XX: Spatial Diagnostics Test for 2014 OLS regression with 3-nearest neighbors weights

Test	MI/DF	Value	Probability
Moran's I (error)	0.0942	2.6980	0.00698
Lagrange Multiplier (lag)	1	12.7998	0.00035
Robust LM (lag)	1	8.3313	0.00390
Lagrange Multiplier (error)	1	4.8309	0.02795
Robust LM (error)	1	0.3624	0.54717
Lagrange Multiplier (SARMA)	2	13.1622	0.00139

Table XX: Spatial Diagnostics Test for 2009 OLS regression with 5-nearest neighbors weights

Test	MI/DF	Value	Probability
Lagrange Multiplier (lag)	1	11.8206	0.00059
Robust LM (lag)	1	9.1548	0.00248
Lagrange Multiplier (error)	1	2.6693	0.10230
Robust LM (error)	1	0.0035	0.95302
Lagrange Multiplier (SARMA)	2	11.8241	0.00271

Table XX: Spatial Diagnostics Test for 2009 OLS regression with 4-nearest neighbors weights

Test	MI/DF	Value	Probability
Lagrange Multiplier (lag)	1	15.7399	0.00007
Robust LM (lag)	1	10.4242	0.00124

Lagrange Multiplier (error)	1	5.3757	0.02042
Robust LM (error)	1	0.0600	0.80644
Lagrange Multiplier (SARMA)	2	15.7999	0.00037

Table XX: Spatial Diagnostics Test for 2009 OLS regression with 3-nearest neighbors weights

Test	MI/DF	Value	Probability
Moran's I (error)	0.1234	3.4519	0.00056
Lagrange Multiplier (lag)	1	8.0569	0.00453
Robust LM (lag)	1	2.4314	0.11893
Lagrange Multiplier (error)	1	8.3026	0.00396
Robust LM (error)	1	2.6770	0.10181
Lagrange Multiplier (SARMA)	2	10.7340	0.00467

Table XX: Spatial Diagnostics Test for 2004 OLS regression with 5-nearest neighbors weights

Test	MI/DF	Value	Probability
Moran's I (error)	0.1161	2.9165	0.00354
Lagrange Multiplier (lag)	1	6.7051	0.00961
Robust LM (lag)	1	2.2242	0.13586
Lagrange Multiplier (error)	1	6.0035	0.01428
Robust LM (error)	1	1.5226	0.21723
Lagrange Multiplier (SARMA)	2	8.2277	0.01634

Table XX: Spatial Diagnostics Test for 2004 OLS regression with 4-nearest neighbors weights

Test	MI/DF	Value	Probability

Moran's I (error)	0.1531	3.2270	0.00125
Lagrange Multiplier (lag)	1	7.9098	0.00492
Robust LM (lag)	1	2.0066	0.09661
Lagrange Multiplier (error)	1	7.9730	0.00475
Robust LM (error)	1	2.0699	0.15024
Lagrange Multiplier (SARMA)	2	9.9797	0.00681

Table XX: Spatial Diagnostics Test for 2004 OLS regression with 3-nearest neighbors weights

Bibliography

Ackermann, Klaus, Simon D Angus, and Paul A Raschky. 2017. "The Internet as Quantitative Social Science Platform: Insights from a Trillion Observations." *ArXiv Preprint ArXiv:1701.05632*.

Ajzen, Icek. 2011. "Theory of Planned Behavior." *Handb Theor Soc Psychol Vol One* 1 (2011): 438.

Andonova, Veneta. 2006. "Mobile Phones, the Internet and the Institutional Environment." *Telecommunications Policy* 30 (1). Elsevier: 29–45.

Andonova, Veneta, and Luis Diaz-Serrano. 2009. "Political Institutions and Telecommunications." *Journal of Development Economics* 89 (1). Elsevier: 77–83.

Andres, Luis, David Cuberes, Mame Diouf, and Tomas Serebrisky. 2010. "The Diffusion of the Internet: A Cross-Country Analysis." *Telecommunications Policy* 34 (5). Elsevier: 323–40.

Anselin, Luc. 1988. "Lagrange Multiplier Test Diagnostics for Spatial Dependence and Spatial Heterogeneity." *Geographical Analysis* 20 (1). Wiley Online Library: 1–17.

Anselin, Luc, A. Murray, and S. Rey. 2013. *The Oxford Handbook of Quantitative Methods, Volume 1: Foundations, Chapter 8: Spatial Analysis.* Oxford University Press.

Anselin, Luc, and Daniel Arribas-Bel. 2011. "Spatial Fixed Effects and Spatial Dependence." Citeseer.

Bank, World. 2016. *Digital Dividends World Bank Report 2016.*

Beilock, Richard, and Daniela V Dimitrova. 2003. "An Exploratory Model of Inter-Country Internet Diffusion." *Telecommunications Policy* 27 (3). Elsevier: 237–52.

Berggren, Niclas, and Therese Nilsson. 2015. "Globalization and the Transmission of Social Values: The

Case of Tolerance." *Journal of Comparative Economics* 43 (2). Elsevier: 371–89.

Caselli, Francesco, and Wilbur John Coleman II. 2001. "The Us Structural Transformation and Regional Convergence: A Reinterpretation." *Journal of Political Economy* 109 (3). The University of Chicago Press: 584–616.

Chinn, Menzie D, and Robert W Fairlie. 2006. "The Determinants of the Global Digital Divide: A Cross-Country Analysis of Computer and Internet Penetration." *Oxford Economic Papers*. Oxford Univ Press.

Collier, Paul. 2008. *The Bottom Billion: Why the Poorest Countries Are Failing and What Can Be Done About It*. Oxford University Press, USA.

Comin, Diego A, and Marti Mestieri. 2013. "Technology Diffusion: Measurement, Causes and Consequences." National Bureau of Economic Research.

Crenshaw, Edward M, and Kristopher K Robison. 2006. "Globalization and the Digital Divide: The Roles of Structural Conduciveness and Global Connection in Internet Diffusion." *Social Science Quarterly* 87 (1). Wiley Online Library: 190–207.

Cruz-Jesus, Frederico, Maria Rosalia Vicente, Fernando Bacao, and Tiago Oliveira. 2016. "The Education-Related Digital Divide: An Analysis for the Eu-28." *Computers in Human Behavior* 56. Elsevier: 72–82.

Davis Jr, Fred D. 1986. "A Technology Acceptance Model for Empirically Testing New End-User Information Systems: Theory and Results." PhD thesis, Massachusetts Institute of Technology.

Dohse, Dirk, and Cheng Yee Lim. 2016. "Macro-Geographic Location and Internet Adoption in Poor

Countries: What Is Behind the Persistent Digital Gap?" Kiel Working Paper.

Falck, Oliver, Stephan Heblich, and Elke Luedemann. 2012. "Identity and Entrepreneurship: Do School Peers Shape Entrepreneurial Intentions?" *Small Business Economics* 39 (1). Springer: 39–59.

Gaskell, George, Sally Stares, Agnes Allansdottir, Nick Allum, Paula Castro, Yilmaz Esmer, Claude Fischler, et al. 2010. "Europeans and Biotechnology in 2010 Winds of Change?" European Commission.

Goldfarb, Avi, and Jeff Prince. 2008. "Internet Adoption and Usage Patterns Are Different: Implications for the Digital Divide." *Information Economics and Policy* 20 (1). Elsevier: 2–15.

Hargittai, Eszter. 1999. "Weaving the Western Web: Explaining Differences in Internet Connectivity Among Oecd Countries." *Telecommunications Policy* 23 (10). Elsevier: 701–18.

Henisz, Witold J, Bennet A Zelner, and others. 2001. "The Institutional Environment for Telecommunications Investment." *Journal of Economics & Management Strategy* 10 (1). Citeseer: 123–47.

Hirt, Martin, and Paul Willmott. 2014. "Strategic Principles for Competing in the Digital Age." *McKinsey Quarterly* 2: 93–108.

Keller, Wolfgang. 2002. "Geographic Localization of International Technology Diffusion." *The American Economic Review* 92 (1). American Economic Association: 120–42.

Kiiski, Sampsa, and Matti Pohjola. 2002. "Cross-Country Diffusion of the Internet." *Information Economics and Policy* 14 (2). Elsevier: 297–310.

Mathieson, Kieran. 1991. "Predicting User Intentions: Comparing the Technology Acceptance Model with the Theory of Planned Behavior." *Information Systems Research* 2 (3). INFORMS: 173–91.

Merton, Robert K, and Alice S Kitt. 1950. "Contributions to the Theory of Reference Group Behavior."

Continuities in Social Research: Studies in the Scope and Method of The American Soldier. The Free Press Glencoe, IL, 40–105.

Merton, Robert King. 1968. *Social Theory and Social Structure*. Simon; Schuster.

National Telecommunications and Information Administration. 2002. “A NATION ONLINE : How Americans Are Expanding Their Use of the Internet.” *Population (English Edition)*, no. February: 92.

Norris, Pippa. 2001. *Digital Divide: Civic Engagement, Information Poverty, and the Internet Worldwide*. Cambridge University Press.

Parts, E. 2013. “Social Capital, National Values and Attitudes Towards Immigrants: Empirical Evidence from the European Union and Neighbouring Countries. Wp 3.” SEARCH Working Paper 3.19.

Rodriguez, Francisco, and Ernest J Wilson. n.d. “Are Poor Countries Losing the Information Revolution?” *InfoDev Working Paper* 26651.

Tengtrakul, Pitikorn, and Jon M Peha. 2013. “Does Ict in Schools Affect Residential Adoption and Adult Utilization Outside Schools?” *Telecommunications Policy* 37 (6). Elsevier: 540–62.

Venkatesh, Viswanath, and Michael G Morris. 2000. “Why Don’t Men Ever Stop to Ask for Directions? Gender, Social Influence, and Their Role in Technology Acceptance and Usage Behavior.” *MIS Quarterly*. JSTOR, 115–39.

Wach, Krzysztof, and others. 2015. “Impact of Cultural and Social Norms on Entrepreneurship in the Eu: Cross-Country Evidence Based on Gem Survey Results.” *Zarzkadzanie W Kulturze*, no. 16 (1). Wydawnictwo Uniwersytetu Jagiellonskiego: 15–29.

Wallsten, Scott. 2005. “Regulation and Internet Use in Developing Countries.” *Economic Development and*

Cultural Change 53 (2). The University of Chicago Press: 501–23.

Warf, Barney. 2009. "The Rapidly Evolving Geographies of the Eurasian Internet." *Eurasian Geography and Economics* 50 (5). Taylor & Francis Group: 564–80.

Wunnava, P, and DB Leiter. 2008. "Determinants of Inter-Country." *Zentrum Fur Europaische Wirtschafts-forshung Gmbh (ZEW), Discussion Paper*, no. 3666.