Determinants of Internet use in Spain

Fernando Lera-López[a], Margarita Billon[b*], and María Gil[b]

[a]*Department of Economics, Public University of Navarra, Pamplona, Spain*
[b]*Department of Applied Economics, Autonomous University of Madrid, Madrid, Spain*

(Received 27 May 2009, final version received 29 September 2009)

Abstract

This paper analyzes the impact of a variety of socioeconomic, demographic and regional factors to explain Internet use and the frequency of use by individuals in Spain. We have employed binomial and ordered probit models with a Heckman´s two-stage estimation procedure. This allows us to distinguish between different variables and explain both use and the intensity of use, respectively. Internet use is mainly associated with education, age, occupation, employment in service sector, nationality, urban areas and regional GDP per capita. In contrast, frequency of Internet usage is positively related to broadband connection, education, the ways through which Internet skills are acquired, gender, and population size. Knowledge of differences in the determinants of Internet use and its extent of use may help to specify the most suitable policies for each case your abstract

Keywords: Internet, Digital divide, Spain, Moulton.

JEL Classification: O33, O18, R210

[*] Corresponding author. Applied Economics. Autonomous University of Madrid, Cantoblanco, Ctra. Colmenar Viejo Km. 15 28049, Madrid, Spain. Tel.: +34 914978777; Fax: +34 914974971. Email: margarita.billon@uam.es

1. Introduction

Empirical evidence has shown that the Internet may have a remarkable impact on economic performance, mainly through its effects on productivity, as well as on social well-being (Jorgenson, Ho, and Stiroh 2008). While the potential benefits depend on the diffusion of the Internet and its use, an uneven distribution may result in different digital divides both in terms of access and use.

When studying the digital divide some distinction should be made between Internet access, Internet use and the extent of use of the Internet. Many existing studies focus on the determinants of Internet adoption (Chaudhuri, Flamm, and Horrigan 2005; Flamm and Chaudhuri 2007), which, strictly speaking, refers to the existence of an Internet connection (normally a broadband connection) at the unit of analysis (household, firm, institutions). However, although Internet access is required for Internet use, remarkable differences can also be found in terms of Internet use by individuals. Consequently, other studies focus its attention on the individual use of the Internet, principally at the household level (Mills and Whitacre 2003; Rice and Katz 2003). The research on the extent of use as measured by frequency and diversity of use has been less common (Bucy 2000; Demoussis and Giannakopoulos 2006; OECD 2007).

The literature highlights the fact that the digital divide in Internet adoption is decreasing and the differences in Internet access among different social groups are declining. However, a second level of the digital divide, based on unequal digital usage and related to socio-economic factors may also emerge (OECD 2007). Within this framework, the study of the determinants of Internet use and its intensity becomes an issue of growing relevance. Whereas some researchers have shown that models explaining Internet adoption and use may be rather different (Goldfarb and Prince

2008; Madden and Coble-Neal 2003), there are very few studies analyzing the differences between Internet use and the extent of use (Demoussis and Giannakopoulos 2006; OECD 2007).

In addition, studies on Internet use by individuals at the international level have traditionally explored the impact of several socio-demographic factors. However, the literature has also shown that Internet distribution follows an uneven spatial pattern (Billon, Ezcurra, and Lera-López 2008; Zook 2006) and that positive externalities in relation to the country and region of residence should be considered (Vicente and López 2006). These contributions tell us that it makes sense to account for certain regional variables in the analysis of Internet use.

The present study provides an empirical analysis of the determinants of Internet use and frequency of use, rather than Internet access. The Internet is a technology that registers a very dynamic process of diffusion. Changes in Internet use as well as changes in the determinants explaining it happen at speed. Within this context, it should be noted that the purpose of the paper is to provide a snapshot of the extent of diffusion at a specific point in time using data for 2007.

In comparison with the prior research, the present study extends and expands the results achieved to make several contributions. First, it provides empirical evidence related to Spain, a country which is a good example of how some southern European economies show a low degree of Internet adoption and use, despite their economic development. The study might interest other countries with low digital development, such as other European southern countries (Portugal and Greece) and some developing countries as well as to compare the determinants with the results obtained for other economies.

Second, in contrast with other studies analyzing Internet usage and frequency without developing explanatory models (Bucy 2000; OECD 2007) or using separate

models (Madden and Coble-Neal 2003), our study presents a two-stage model to explain both issues.

In addition, along with economic and socio-demographic factors, we also study the influence of several regional variables employing the Moulton correction (Moulton 1990) to avoid the problem of inefficient estimators when individual and aggregate variables are included in the same model.

Finally, the distinction between Internet usage and the extent of use in the analysis of the determinants of the Internet may be crucial to develop the most appropriate policy actions to avoid a new digital divide within countries in terms of the extent of Internet use.

The reminder of the paper is structured as follows. The next section provides a general overview of Internet use in Spain. Section 3 shows an overview of the conceptual framework and the existing empirical literature. Sections 4 and 5 describe the data and variables, as well as the methodology employed, respectively. Section 6 provides the exploratory analysis on Internet use. The explanatory analysis and models are shown in Section 7. Lastly, Section 8 presents the major conclusions and areas for further research.

2. Household Internet use in Spain

Global Internet usage has seen dramatic growth in the last 20 years. The number of Internet users has gone from 55 million in 1995 to 1,407 million in 2007 (Internet World Stats 2008). Spain has been a part of this remarkable trend: the number of users has grown from 18.7% of the total population in 2002 to 52% in 2007. The number of regular users has evolved from 29% of the total population in 2003 to 44% in 2007 (Spanish National Statistics Institute, INE 2008). However, despite this growth, Spain still lags behind the European average (51%) (EUROSTAT 2008). Figure 1 compares

Internet usage in Spain with that of other EU countries. As shown in Figure 1, Internet usage rates in Spain, as is also the case with Italy, Greece and Portugal, are lower than those of other European countries with less GDP per capita. (Figure 1)

Along with the existing disparities in comparison with its neighbors, Internet adoption is characterized by strong inequality among the Spanish regions. Figure 2 shows Internet usage and GDP per capita in the Spanish regions in 2007. Madrid, Balearic Islands and Catalonia present Internet values above the average, while other regions, such as Extremadura, are clearly below. Also, regions such as Navarra and the Basque Country, having high GDP values, show lower rates than other less developed regions, such as Asturias and Cantabria. In fact, the inequalities in Internet usage at regional levels are higher than those of GDP in the European regions (Billon, Ezcurra, and Lera-López 2008). (Figure 2)

In addition, the socio-demographic evolution of Internet use in Spain exhibits several digital divides. These divides are associated with strong differences in characteristics such as economic status, occupation, education level, age and gender. For example, as shown in Figure 3, there is a permanent divide between those individuals with a secondary education and those without it. This pattern has also been shown to exist in Canada, Korea, the United States and other OECD countries (OECD 2007). The effects of age differences are remarkable, whereas the effects of gender differences are less pronounced due to an increase in Internet use by females (Observatorio Redes 2008) (Figure 3).

Finally, the lack of data for Spain precludes analysis by income groups. However, regional data shows that higher income regions generally have a higher number of Internet users (Fundación Orange 2007).

3. Literature review

3.1 Internet diffusion

Internet adoption is usually explained within the framework of the diffusion theory (Rogers 2003). According to epidemic models, the key to explaining technology diffusion relies on the spread of information about the existence of a new technology as well as learning from experience (Mansfield 1968). The greater the number of people adopting, the higher is the probability that a non-user will become a user. For example, Goolsbee and Klenow (2002) find that people are more likely to buy their first home computer in areas where a high proportion of households already own computers or when a large share of their friends and family own computers.

As these models do not allow for differences in the population, we refer to others, such as those incorporating heterogeneity (Rosenberg 1972), to explain the differences in diffusion rates by the observed economic and demographic characteristics of adopters. In fact, socio-economic characteristics (income, location, employment, education, and family structure), personal and demographic factors (age, gender, disability, ethnicity, and children per family) as well as the attitude towards technology, are often considered preconditions for Internet use. Karshenas and Stoneman (1992) demonstrate that exogenous learning and economic factors are significant determinants of the diffusion process.

In turn, the knowledge spill-over literature (Fujita and Mori 2005) highlights the role of the type of knowledge transmitted in determining Internet diffusion. When knowledge is tacit or non-codified, face-to-face communication facilitates its diffusion. The adoption of technologies such as the Internet may occur more rapidly in populated areas or cities (Gaspar and Glaeser 1998) where the density of sources of knowledge about such technology is higher and where the geographic concentration of knowledge spillovers is founded. In a similar vein, the urban density theory

(Forman 2005; Forman, Goldfarb, and Greenstein 2005) holds that the probability of becoming an Internet user increases as population size and density rise. This is due to the decreasing cost of Internet access as certain resources become more abundant (Karshenas and Stoneman 1995).

In addition, the literature on economic geography demonstrates that different indicators of Internet activity, such as broadband, are concentrated in urban areas (Gubresic 2008; Malecki 2002). Internet distribution follows an uneven spatial pattern, revealing the emergence of the so-called Internet geography (Zook 2006).

According to these contributions, it is natural to expect that regional features affect Internet use. For example, we may consider the following factors determining the different spatial distribution of Internet diffusion: the level of telecommunications infrastructure (mainly broadband); the location of high-tech firms; the allocation of resources such as human capital and R&D; and the percentage of employment in the services sector, as well as others.

3.2. Socio-demographic factors explaining Internet use

The main socio-demographic factors mentioned by empirical studies attempting to explain Internet use have been education level, income and age. Other variables, such as gender, familiar composition, occupation, ethnicity and location (rural versus urban), have also shown to be significant.

Empirical evidence shows that the higher the education and income levels, the higher the probability of Internet use (Goldfard and Prince 2008; Ono and Zavodny 2007; Rice and Katz 2003). These results are explained by a variety of factors. According the diffusion theories mentioned above (Rogers 2003), more educated individuals tend to achieve higher professional and economic status, being also more prone to adopt innovations such as the Internet. In addition, the perceived benefits and utility from Internet use vary according to economic status and education level

(OECD 2007). Thus, people with higher income levels report higher benefits than people with lower income levels (Mills and Whitacre 2003).

Age is negatively correlated with Internet use, as empirical evidence has profusely shown (Goldfard and Prince 2008; McKeown, Noce, and Czerny 2007; Mills and Whitacre 2003; OECD 2007; Rice and Katz 2003). The impact of age may be attributed to a combination of factors, including the lack of Internet skills as well as differences in attitude among age groups and the perceptions of the benefits associated with its use, which are lower for elderly people (Hargittai 2003; OECD 2007).

The evidence is far from conclusive with respect to gender. Some studies show that the probability of Internet use is lower for women than for men (Bimber 2000; McKeown, Noce, and Czerny 2007). Lower female use may be associated with differences in educational and income levels, as well as with occupation and the distribution of household responsibilities (Bimber 2000; Losh 2003). However, other studies have found no differences between these groups (Goldfard and Prince 2008; Rice and Katz 2003). In any case, it seems that gender is becoming a less significant factor as Internet use increases (OECD 2007).

Family structure influences Internet use as well. In general, there is a positive influence of the presence of children on Internet diffusion in households (Mills and Whitacre 2003; OECD 2007; Singh 2004). However, this variable is not significant in some studies (Bucy 2000; McKeown, Noce, and Czerny 2007), suggesting that although households with children are more likely to have Internet connections, this does not affect the usage by adults in the household.

Related to the effect of occupation on Internet use, its use is to be associated with the employed, rather than with the unemployed and with pensioners in the US (NTIA 2004) and in the EU (Vicente and López 2006).

Internet usage and its extent seem to be conditioned by variables such as ethnicity and nationality. Martin and Robinson (2004) find that Internet usage is lower among citizens not born in the US. NTIA (2004), Hofman and Novak (2000), and Hacker and Steiner (2002) detect racial differences in Internet use. Carveth and Kretchmer (2002) show differences between the US and Western Europe: whereas race or ethnicity is a major predictor of Internet usage in the US, race does not appear to have a real influence on Internet usage rates in Western Europe. Mills and Whitacre (2003) suggest that Internet content may be less attractive for ethnic groups such as black and Hispanics.

Other variables may also help to explain the difference, such as educational and income differences associated with some ethnic groups. However, after controlling income and education, ethnicity is likely to have differential effects on Internet usage and on its frequency of use. The ethnic divides may be related to the lack of Internet skills as well as opportunities for its use (Hacker and Steiner 2002).

The literature also highlights the existence of a digital divide between rural and urban areas. Internet usage by rural residents is usually less frequent than that of the urban population (see Mills and Whitacre (2003) for a literature review). In Europe, Demunter (2005) shows that this digital divide is particularly wide in the Southern countries and in the new member States. In Canada, rural location is one of the three relevant barriers to Internet use (McKeown, Noce, and Czerny 2007). Traditional differences in education, income and skills between areas may partially explain the differences. Additionally, some authors consider whether the digital divide is a consequence of higher infrastructure costs and lower levels of infrastructure investment in less densely populated areas. Other researchers consider that the divide is the result of the lower likelihood of computer and Internet usage in jobs in rural areas (McKeown, Noce, and Czerny 2007; Singh 2004; Vicente and López 2006) and

differences in perceived benefits from Internet use at home (Mills and Whitacre 2003). Also, according to epidemic models, the positive spillovers of Internet usage by other households in the area may increase Internet propensity in areas with high use of Internet.

Along with demand factors, variables associated with supply factors should also be taken into account. The role of Internet access costs is particularly relevant. Kiiski and Pohjola (2002) show that a 50% reduction in Internet access costs would raise the number of computer hosts per capita by 25% over a 5-year period in 23 OECD countries. Other authors have highlighted the influence of Internet costs on its usage in Australia (Madden and Coble-Neal 2003) and across Europe (Demoussis and Giannakopoulos 2006; Vicente and López 2006) although Chaudhuri, Flamm and Horrigan (2005) find that its influence is small in the US. Internet costs are not only determined by the availability of ICT infrastructure but also by the competition among local providers (Grubesic 2008). Urban areas have been characterized by more broadband competition than rural areas.

3.3. Socio-demographic factors explaining the extent of use of the Internet

The literature review of frequency or intensity of Internet use is rather less profuse than the analysis of Internet usage. Nevertheless, the comparison among studies and countries is complicated. Whereas Internet usage is measured by a dichotomous variable (yes or no), frequency can be measured by different indicators. Some studies include the number of online hours per day (Horrigan and Rainie 2002) and per week (Goldfarb and Prince 2008), while others consider the number of access per week and month (Bucy 2000; Demoussis and Giannakopoulos 2006; OECD 2007). Other studies consider the diversity of Internet use as a proxy for intensity of Internet usage (Horrigan and Rainie 2002; OECD 2007). To solve this methodological problem, some researchers have relied on an index including typical measures of frequency,

such as hours per week as well as measures of diversity in Internet usage (Veenhof, Clermont, and Sciadas 2005).

Results on the influence of education and income levels on frequency of Internet usage have been ambiguous. The US case is document in Bucy (2000), while the French case is documented by OECD (2007). Both studies find that the higher the education and income levels, the higher the frequency of Internet usage in terms of access per week. However, Godlfarb and Prince (2008) show that, among US households, online hours per week decrease when both income and education rise. According to the authors, when consuming a fixed cost product takes time (as it happens with the Internet), the value of time must be considered the most relevant cost. As lower-income people have less options to carry out a great variety of different leisure activities due to low wages, their opportunity cost of leisure time is lower. .Consequently, they are much more likely to spend more hours consuming a fixed-cost product such as the Internet. Among Canadian households, Horrigan and Rainie (2002) document that education level explains the number of Internet activities, but not the amount of time spent online. Finally, as the Internet is an interactive technology, specific skills are needed to look for and use online information. These skills are often associated with higher levels of education (Hargittai 2003; Chaudhuri, Flamm, and Horrigan 2005). In the US, individuals having more Internet experience (Horrigan and Rainie 2002) or Internet skills (Hacker and Steiner 2002) are more likely to show a more intensive use of the Internet.

Evidence concerning age and gender is inconclusive as well. Generally speaking, age is negatively correlated to the frequency of use in the US (Bucy 2000) and in Canada (Horrigan and Rainie 2002; OECD 2007) and also in terms of the number of online activities developed in other countries, such as Finland and the Netherlands (OECD 2007). Regarding the influence of gender, men tend to use the

Internet more frequently (Bucy 2000) and in a more varied way (OECD 2007) in the US, Finland, France and the Netherlands. However, an analysis of American users revealed that men spend less time online and engage in fewer Internet activities than do women (Horrigan and Rainie 2002).

Evidence of the influence of family structure on the intensity of Internet use shows that the number of people in a respondent's household is statistically and positively related to the frequency of Internet usage in some cases (Bucy 2000), while in others, the number of children has a negative impact on the number of hours adults spent online (Goldfarb and Prince 2008).

The digital divide between rural and urban areas also shows up in terms of the frequency of Internet use. According to OECD (2007), Finnish residents of urban areas show a more varied use of the Internet than those living in rural areas. Also, the frequency of Internet use in Australia is higher in non-remote areas than in remote ones (ABS 2002). As mentioned in the previous section, this result may stem from the lower level of ICT infrastructure in rural areas, such as broadband connection. In fact, the type of connection has a clear influence on the frequency of Internet use and on the variety of online activities. Having a broadband connection at home clearly increases the frequency of Internet use in the US (Horrigan and Rainie 2002; Horrigan and Smith 2007) and in many European countries (OECD 2007).

3.4. Regional factors influencing Internet usage and its extent

Together with personal and socio-demographic factors, some studies use regional variables to explain Internet usage by individuals and households. The inclusion of these types of variables usually seeks to explain geographical disparities in Internet adoption. Some studies, such as those by Vicente and López (2006) and Demoussis and Giannakopoulos (2006) for the European countries and by Schleife (2006) for German regions, aim to explain differences in Internet use by highlighting the role of

different aggregate variables. These studies consider variables related to the level of economic development, such as GDP per capita, the rate of unemployment, trade openness, R&D intensity (R&D expenditure, patents and employment in high-technology sectors) as well as variables related to sectoral composition (percentage of employment in the services sector, agriculture and industry). Other variables analyzed include the population composition (percentage of foreign population, percentage of monoparental households) as well as several indicators related to Internet costs, infrastructure (Internet hosts) and spill-over (percentage of Internet users in the region).

Results for the European Union show that the influence of GDP per capita and R&D intensity is positive and significant in explaining Internet usage (Demoussis and Giannakopoulos 2006; Vicente and López 2006). Also, Internet host and trade openness are positively related to Internet usage, whereas sectoral composition is positively related to Internet intensity (Demoussis and Giannakopoulos 2006). This research also documents a negative effect between the cost of Internet use and the probability being an Internet user, although the effect is lower than that of the rest of variables (Vicente and López 2006). In Demoussis and Giannakopoulos (2006), the cost of Internet access is only relevant in explaining Internet intensity, but not Internet usage.

For the German regions, Schleife (2006) finds that individual characteristics are more important than regional factors in explaining individual decisions to use the Internet. In fact, only the percentage of monoparental families, as well as the rate of unemployment, influences the probability of becoming a user. Schleife also finds that a larger number of regional Internet users has a positive impact on the probability of becoming an Internet user.

Although these studies obtain very interesting and significant results when regional variables are included in the Internet usage models, they are subject to some methodological limitations. These problems are due to the simultaneous inclusion of both individual and regional (aggregate) variables, which generates inconsistent estimators. To overcome this problem, this study employs the Moulton correction, proposed in 1990 by Moulton (1990).

In summary, theoretical contributions and empirical evidence reveal the role of socio-demographic factors and certain regional features in explaining Internet use and the extent of use at the international level. We test these findings for the Spanish case in the following sections.

4. Research model and methodology

The empirical evidence about the economic analysis of Internet demand by households and individuals is based on a standard neoclassical utility maximization framework (Demoussis and Giannakopoulos 2006; Fairlie 2004; Mills and Whitacre 2003; Madden and Simpson 1997; Vicente and López 2008). Within this context and considering the prior literature we present the following research model.

This study's objective is twofold. First, we seek to analyze the determinants of individual Internet use. Second, we wish to analyze the determinants of the intensity of Internet use. Logically as the intensity of Internet use is related to the Internet usage decision, both decision models should be jointly modeled (Demoussis and Giannakopoulos 2006).

In a first step, the decision of using the Internet may be considered as a discrete choice based on a simple linear random utility model. Demand models are usually formulated by comparing utilities (U_{ij}) that individuals (i) obtain from every alternative (j). In this case, the choice modeled is to use the Internet (j =1) or not (j =

0). The utility assigned by individual *i* to every alternative *j* is a linear function of a vector of individual characteristics (X_i), consisting of observable heterogeneity. This way, individuals with different characteristics assign different values to every choice. On the other side, the term ε_{ij} represents the unobservable heterogeneity or random perturbation.

$$U_{ij} = X'_{ij}\,\beta_{ij}+\varepsilon_{ij} \qquad j = 0, 1;\ i = 1, 2, \ldots, n \qquad (1)$$

Every individual will choose alternative j = 1 if this utility is higher than that under alternative j = 0. As utility is not observable, we will denote Yi = 1 as the event that the "individual uses the Internet," while Yi = 0 represents the complementary event. So, in probabilistic terms, the probability of choosing alternative "he/she uses the Internet" turns out to be:

$$\begin{gathered} P(Y_i = 1) = P(U_{i1} > U_{i0}) = P(X'_{i1}\beta_1 + \varepsilon_{i1} > X'_{i0}\beta_0 + \varepsilon_{i0}) = \\ P[X'_i \cdot (\beta_1 - \beta_0) + (\varepsilon_{i1} - \varepsilon_{i0}) > 0] = P(X'_i\beta + \varepsilon_i > 0) = \\ P(\varepsilon_i > -X'_i\beta) = F(X'_i\beta) \end{gathered} \qquad (2)$$

This decision is usually modeled using probit or logit discrete choice models. We use a probit model in this study, assuming that the regressors' random perturbations are normally distributed.

Once the main factors explaining Internet use are determined, in a second stage we study the factors affecting the frequency of use, which obviously requires previous use. However, restricting our analysis to the individuals who had used the Internet before would introduce a sample selection bias. As Demoussis and Giannakopoulos (2006) point out, if there are systematic reasons for why people with an Internet connection do not use it, and the analysis is only restricted to Internet users, then the parameter estimates in the model will be inconsistent (Greene 2008; Heckman 1979). To avoid this problem, we control sample selection bias using the standard sample selection model considered by Heckman (1979)[1]. However, it should

be noted that the dependent variable in the second equation (frequency of use) is not continuous, but ordinal with the three categories shown in Figure 4. For this reason, the Heckman standard sample selection model has to be extended to accommodate ordinality (Dubin and Rivers 1990).

To account for the ordinal character of Internet frequency[2], we use an ordered probit model based on the following specification:

$$T^{*}_{ik} = \gamma'_{ik} Z_{ik} + u_{ik} \qquad k = 0, 1, 2;\ i = 1, 2, \ldots, n \qquad (3)$$

where T^{*}_{ik} is a latent variable with the following structure (Greene 2008):

$$T_i = \begin{cases} 0 \text{ if } T^{*}_i \leq \mu_0 = 0 \\ 1 \text{ if } \mu_0 \leq T^{*}_i \leq \mu_1 \\ 2 \text{ if } \mu_1 \leq T^{*}_i \leq \mu_2 \\ ::::: \\ :::: \\ K \text{ if } T^{*}_i > \mu_{k-1} \text{ and } 0 < \mu_1 < \mu_2 < L < \mu_{k-1} \end{cases} \qquad \ldots..(4)$$

The variable of interest, T*ik, is a continuous and unobserved variable capturing the extent of Internet usage. The observed categories, Ti, are assumed to represent an ordered partitioning of this continuous scale, where Ti is the observed category for the ith individual, being γ the vector of coefficients, Zi is the vector of explanatory variables for the ith individual, ui is the standard normal random error and the μk are threshold parameters. Consequently, the higher the γ values, the higher the probability of the frequent Internet use. There are three response levels in our analysis. Consequently, the probabilities according to, an accumulated normal distribution, $\Phi(\cdot)$, are given by:

$$\begin{aligned} P(Ti=0) &= \Phi(-\gamma' iZi) \\ P(T_i=1) &= \Phi(\mu_1 - \gamma'_i Z_i) - \Phi(-\gamma'_i Z_i) \\ P(T_i=2) &= \Phi(\mu_2 - \gamma'_i Z_i) - \Phi(\mu_1 - \gamma'_i Z_i) \\ P(T_i=3) &= \Phi(\mu_3 - \gamma'_i Z_i) - \Phi(\mu_2 - \gamma'_i Z_i) \end{aligned} \qquad (5)$$

To sum up, the Internet demand model is estimated in two stages. In the first stage, we use a binomial probit model. This statistical model allows us to identify the factors explaining Internet use (Yi) for the whole sample. The second model seeks to determine the factors explaining the frequency of Internet use, (Ti) (Figure 4).We use an ordered probit model and consider the Heckman correction. We introduce the parameter lambda (λi, the inverse Mills ratio of the first stage) into the model, which is the correction term of the selection bias, adopting this expression:

$$\lambda_i = \frac{\phi(X_i'\beta)}{1-\Phi(X_i'\beta)} \text{ if } Y_i=1 \qquad (6)$$

Finally, we wish to mention that a particular case arises in our probit model, related to some variables which are not provided by the survey but that can be affecting the decision of using the Internet (GDP per capita and the percentage of employment in service activities, as we will see in following sections). As we have information from other sources (Spanish National Statistics Institute, INE 2004), we have proceeded to input these variables from this aggregated statistic at a regional level to our survey. This causes a statistical problem identified by Moulton in 1990, as we have already mentioned. This problem occurs when combining individual cross-section data with aggregated data (in our case, at a regional level). An intra-group correlation in the random perturbation arises, causing estimators to be inefficient. For this reason, we have used a probit estimation introducing the Moulton[5] correction, which controls for this effect. Thus, standard errors (and consequently individual significance of every variable) are provided with this correction.

5. Data and variables

The data come from the survey of Equipment and Usage of ICT at the household level (in Spanish "*Encuesta sobre Equipamiento y Uso de Tecnologías de Información y*

Comunicación en los Hogares," TICH) published by the Spanish National Statistics Institute, INE 2008) in each semester since 2005. The survey is the only one covering ICT equipment in Spanish households, as well as their use at the national level. It allows us to understand the level of ICT development of Spanish individuals and households. Data are from the first six months of 2007, which provide updated information about Internet use for 61,834 Spanish individuals over 16 years, conforming 22,198 households. In addition, note that the survey is launched throughout Spain using a three-stage sampling process stratified by regions. One important particularity of this database refers to the way of filling in the questionnaire. The general questionnaire has two parts: one related to ICT equipment and Internet access in the household and one devoted to asking about personal ICT use as and some individual socioeconomic characteristics. These latter questions are answered by just one member of the household, who is randomly selected.

The dependent variable in the first stage is the variable indicating whether the individual uses the Internet or not. In line with the conceptual framework and empirical evidence mentioned above, some of the explanatory variables aim to capture the impact of the socio-demographic factors explaining Internet use. The variables selected are the following: age, gender, education level, occupation and nationality. In addition, we also include variables such as household structure, people under 10-15 and the degree of urbanization. The degree of urbanization measures the urbanization level where the individuals surveyed live.

Unfortunately, the survey does not provide information such as the cost of Internet per person, ethnicity and household income. However, empirical evidence has demonstrated a close relationship between income level, occupation and education level. For this reason, we consider using one of those variables to serve as a

proxy for income level. Ethnicity may be indirectly analyzed using the nationality variable. Unfortunately, there is no proxy for the impact of the cost of Internet.

Along with socio-demographic variables capturing individual characteristics, we have included other regional variables associated with certain economic development features, which may affect the individual decision to use the Internet. These variables are GDP per capita and the percentage of employment in the services sector. These variables are available at the county level. As they were not included in the survey, we imputed them according to individuals´ region of residence. The regional variables were obtained through different official statistics developed by the Spanish National Statistics Institute. Other significant variables included in other studies at national level are not available at regional level in Spain: ICT infrastructure (hosts, websites, etc.), R&D intensity and trade openness. The inclusion of other variables to measure possible spill-over effects and to test the influence of epidemic models, such as the percentage of Internet users in the region, was rejected due to endogeneity concerns.

In the second stage, the dependent variable is frequency of use. As previously explained, frequency of use is measured as an ordered variable with three different categories: daily usage, weekly usage and monthly usage. Unfortunately, the survey offers no information on individuals' number of online hours and on the number of online activities. We consider gender, age, education level combined with Internet competencies and people under 15 living in the household, together with urbanization level, as independent variables. We have also included two other variables associated with Internet use: ADSL Internet access at home and a variable capturing the means through which individuals acquire their skills to use the Internet. The first allows us to analyze the impact of infrastructure on use, while the second provides information about how Internet skills are acquired.

6. Exploratory analysis

Table 1 shows the interval frequencies for the categorical variables as well as the mean and standard deviation of the only continuous variable. Over the period analyzed, we see that the percentages of Internet users and non-users are quite similar. However, the number of those using the Internet is slightly higher. With regard to the individual socio-demographic features, the average age of Internet users is 35, while that of non-users is 58. Then, age differences are remarkable: non-Internet users are older and show greater age dispersion. (Table 1)

Regarding gender, Internet use is slightly higher for men than for women (by 8%). The difference becomes greater at the highest level of education. Among highly educated individuals, Internet use increases by 87.5%, while 90% of those with no education or only a primary education do not use the Internet. Results by occupation are conclusive for the three categories. Most students are Internet users (over 98%). Likewise, 70% of workers use the Internet. On the opposite side, we find that most pensioners and housewives do not tend to use the Internet.

Other variables considered include the presence of children between 10-15 years in the household. Among households with children of this age, the percentage of Internet users is 18%; in other households, this percentage decreases to 12%.

With regard to the degree of urbanization, we find that the bigger the municipality, the higher the percentage of Internet users. However, we wish to highlight the fact that there is a slightly higher percentage of users if a municipality is the capital city of the region l, compared to non-capital municipalities with more inhabitants.

Similar to our previous analysis, it is also possible to characterize the frequency of use for those individuals who use the Internet. The methodology employed by the Spanish National Statistics Institute for this survey distinguishes

between daily, weekly and monthly frequencies of use. Daily frequency refers to the use of the Internet at least 5 days/per week. Weekly frequency refers to the use of the Internet between one and four days every week. Finally, monthly frequency includes use at least once a month, but less than once a week.

Out of the 11,568 individuals who claim to use the Internet, only 8,837 use it at least once a month. This means that 2,731 individuals use the Internet less than once a month. Those users have been excluded from the sample because less than monthly use is not considered frequent.

The descriptive statistics of Internet users' characteristics according to the frequency of use are shown in Table 2. According to the empirical evidence, the factors determining the frequency of Internet use should not be the same as those explaining the decision to use the Internet. Consequently, some variables included to explain Internet usage are not the same as those included to explain the intensity of use. (Table 2)

There are no remarkable gender differences, although men show slightly more frequent daily use. There are small differences in mean and standard deviation with respect to age. The average age decreases as the frequency of use increases.

With respect to the education level and the Internet competencies that individuals have, it seems that daily use is strongly associated with having advanced Internet competencies, though education is also important (the highest percentage is observed for the category of individuals with Higher education and advanced competencies, 76.5%). Concerning the degree of urbanization, we find that the percentage of daily users decreases in smaller municipalities.

The type of Internet connection becomes a key variable in studying the frequency of Internet use. There is an increasing relationship between frequency of

use and the existence of broadband in households, such as ADSL. Other types of connections lower the frequency of use.

Finally, regarding the channel through which individuals acquire Internet skills, we are interested in identifying the most common channel for acquiring Internet knowledge[3] by frequency of use. Thus, it seems that the category "other means" is the most frequent when Internet is used daily. However, we note that this is a rare option (only 3% of daily internet users). The most common avenue through which daily users acquire Internet skills is own learning, followed by regular and company training. On the other hand, those who use the Internet the least tend to learn through courses for adults, from other users or on their own.

7. Results

7.1. Determinants of Internet use

As mentioned above, we have estimated a two-stage model, using the Statistical software Stata, version 10. In the first stage, we estimate the determinants of the decision to use the Internet (yes or no). In the second stage, we study the frequency of use (daily, weekly and monthly).

The first stage is estimated with a probit model, using a robust to heteroscedasticity Huber-White estimator of variance. The results are shown in Table 3. The model adjustment is good regarding all adjustment measures provided. The model is significant at a global level, as shown by the likelihood test shows (Wald Chi2). The value of the pseudo R2 is relatively high (47.44%) within the context of micro econometric models such as this one. Finally, the model's predictive capacity is high; taking the value 0.5211[4] as the cutting point, the accuracy of the model (when the dependant variable is 1) is 85.87% (79.75% when the dependant variable is 0). We study the variables independently, interpreting their significance, sign and value. The interpretation of the coefficients can be done in probabilistic terms (marginal effects),

that is, analyzing how the probability of Internet usage varies with changes in every independent variable, ceteris paribus the rest of variables. (Table 3)

Looking at the results, we see that both regional variables included are significant and show a positive impact on the probability of Internet use. Thus, as GDP or the percentage of employment in services increases, the probability of demanding the Internet rises. Interestingly, the percentage of employment in service activities is the most powerful regional determinant of Internet use (value of 0.2704, while the marginal effect on GDP is only 0.0087). The role of GDP confirms previous findings showing that economic development is positively associated with Internet usage (Demoussis and Giannakopoulos 2006; Vicente and López 2006).

We highlight the following results when looking at socio-demographic variables. First, education level has a great impact on the probability of Internet use. People with secondary education have a 38% higher likelihood of becoming an Internet user than people with no education or qualifications. Thus, higher education increases the probability of Internet use by 61%. The positive influence of education is corroborated by other empirical evidence. This influence may be explained by higher benefits and utility associated with more educated groups (Mills and Whitacre 2003) as well as the fact that this group possesses the required Internet skills. These skills are often associated with higher education (Chaudhuri, Flamm, and Horrigan 2005; Hargittai 2003); their absence may be a relevant barrier to Internet use.

Second, all age and gender groups are statistically significant. In comparison with the reference category (woman over 64 years), belonging to other categories increases the probability of using the Internet. The young (men and women) between 16 and 24 have the highest probability, followed by men and women between 25 and 34 years. In the third place we have the individuals between 35-44 years, followed by the group between 45 and 54 years. Finally, individuals over 64 years are the least

likely to use the Internet. There are no significant gender differences in Internet usage, whereas Internet usage clearly decreases as age increases. The negative influence of age is corroborated by widespread empirical evidence. It may reflect differences in attitudes and lifestyles by different groups of age as well as the fact that the needs and benefits from its use decrease with age. It seems that the digital divide is decreasing with respect to gender, confirming the empirical evidence mentioned above.

Third, the presence of children between 10 and 15 years in the household is not significant in explaining Internet usage in Spain, confirming the available empirical evidence for other countries (Bucy 2000; McKeonn, Noce, and Czerny 2007). Although households with children may be more likely to adopt Internet, this fact does not affect Internet use by adults.

Fourth, there are no significant differences in Internet usage between the retired people and housewives. Being a student increases the probability of becoming a new Internet user by 48%, while being worker, unemployed and being in another situation, respectively, increases the probability of being an Internet user by 26% and 12%, compared with the reference category (retired people). Therefore, it seems that Internet usage may be partially explained by the usefulness of the Internet for many professional activities. In addition, Internet usage seems to be particularly useful for students and for the unemployed in searching for a new job. Both groups of people may have a lot of time to spend using the Internet.

Fifth, nationality is statistically significant at the 0.05 level. Being Spanish increases the likelihood of being an Internet user. Even controlling for education, nationality seems to have differential effects on Internet usage, corroborating the empirical work by Hofman and Novak (2000), Hacker and Steiner (2002), and Martin and Robinson (2004) for the US. Considering that non-Spanish nationality mostly includes people from South America, North of Africa and Western Europe (Romania,

Bulgaria), this distinction may be related to the lack of Internet skills as well as to the existence of less opportunities for and less benefits from f its use.

Finally, the population size has an important influence on Internet usage. Living in municipalities of up to 50,000 inhabitants has no impact on Internet usage. The bigger the urban area, the higher the probability of being an Internet user. Consequently, we detect a digital divide between urban areas and rural areas (less than 50,000 inhabitants) in Spain, corroborating the available empirical evidence. This may be the consequence of lower infrastructure investment in rural areas as well as of the differences in the perceived benefits from Internet use between citizens from urban and rural areas.

However, the value of the marginal effect when the individual lives in a city with more than 50,000 inhabitants is very similar to that registered when he lives in a smaller regional capital. Thus, there is a regional capital effect. When controlling for education and other socio-demographic variables living in the main city has a positive effect on the likelihood of being an Internet user regardless of the number of inhabitants.

7.2. Determinants of the frequency of Internet use

Now that we have analyzed the determinants of Internet usage, we study the relevant factors explaining the frequency of Internet use. In this second stage, we consider Internet users with a frequency of use over “once in the last three months” exclusively. To avoid inconsistencies in the parameter estimations, we use the standard sample selection model by Heckman (1979). The significant value of the Lambda parameter allows us to confirm the existence of sample selection in the use of the Internet. The results are shown in Table 4, provided for each frequency category: daily, weekly, monthly. As with the preceding model, we estimate this one using a

robust to heteroscedasticity procedure. The model adjustment is good. Variables will be interpreted as mentioned before, using their marginal effects. (Table 4)

The first result to highlight is the fact that, although the significance levels are similar for the three categories (daily, weekly and monthly), the estimated values of the coefficients show clear and important differences in values and signs. The marginal effects are very similar for the weekly and monthly frequencies of Internet use, but totally different for the daily frequency. Consequently, we can assume that the determinants of daily Internet use are the opposite of the relevant factors explaining less frequent Internet use.

The second interesting result is the fact that the variable capturing the inverse Mills ratio (Lambda parameter) in the three models is significant, implying that there is a selection bias in the frequency of Internet use.

In this second model, the positive influence of Internet competencies and education is confirmed (Bucy 2000; Hacker and Steiner 2002; OECD 2007). Combining these two variables, we may study the interaction effect of formal education and Internet skills. Results show that all the categories are significant, except the one considering people with secondary studies and basic Internet skills, regarding to the reference category. More educated individuals are more prone to do a more frequent use of the Internet, but it is remarkable the non negligible impact of personal skills on it. According to the results, having advanced Internet competencies increases the daily frequency of use and this effect is even more remarkable when the individual is more educated. For example, higher education and advanced competencies in Internet increase the probability of daily frequency use by 35%, ten points more than secondary education and advanced competencies. A possible explanation may be associated with differences in the benefits and utility of Internet usage, depending on education level as well as on the level of Internet skills.

The results show that the influence of age on the frequency of Internet usage is statistically significant, but the impact differs according to the category of use frequency. Daily use decreases with age. In contrast, age has a positive impact on weekly and monthly Internet use. To sum up, the use of the youngest is associated with a higher daily Internet use, while aged people are more likely to use the Internet at a weekly or monthly frequency. Besides, the variable square of age is also significant, despite the imperceptible value of the coefficient, showing a slightly non-linear effect. It may lead us to remark the importance of youth on the frequency of use of the Internet. The role of gender confirms the previous findings: men tend to use the Internet more frequently than women (Bucy 2000; OECD 2007). When controlling for education and Internet infrastructure (ADSL), age and gender are relevant. This may suggest significant differences in the purpose as well as in the benefits of Internet usage, depending on the age groups and on gender. Having children between 10 and 15 years old seems to have a negative effect on frequency of use, confirming the results shown by Goldfarb and Prince (2008).

In analyzing the role of population size, it seems that living in dense urban areas has a positive and statistically significant influence on the frequency of Internet use. The bigger the size of population, the higher likelihood of daily use of the Internet. This positive impact corroborates prior literature (ABS 2002; OECD 2007). As controlling for Internet infrastructure with this variable is relevant, this result seems unrelated to the lower level of ICT infrastructure in rural areas.

With respect to the type of Internet connection, our results corroborate previous findings: having a broadband connection at home is the largest single factor in explaining the frequency of Internet use (Horrigan and Rainie 2002). Having a broadband connection at home makes a respondent about 19% more likely to use the Internet daily, decreasing the weekly and monthly use of Internet.

Finally, the way through which Internet skills are acquired has a clear impact on Internet frequency. Thus, learning by practice and self-teaching explain a higher frequency of use, followed by company and regular training. These four ways of learning have a negative impact on the weekly and monthly use of Internet. Learning from other adults increases non-daily frequency of use. Finally, less frequent users of the Internet tended to learn mainly through courses for adults. However, this variable is not statistically significant.

8. Conclusions and implications

This paper analyses the determinants of Internet usage and the extent of use at individual level for a southern European country. As the bulk of the literature has been focused on the United States and on the European Union as a whole, the study might interest other some EU countries, such as Portugal and Greece as well as developing countries that despite high rates of economic growth register low rates of digital development.

We extend the available research by studying not only the use, but also the extent of Internet use for Spain and analyzing not only the influence of economic and socio-demographic variables at the individual level, but also the influence of some regional variables. In order to distinguish between usage and frequency of use, we employed binomial and ordered probit models with a Heckman two-stage estimation procedure. In addition, we estimated probit models including the Moulton correction in order to avoid inefficient estimators, given that the models include both individual and aggregate (regional) variables.

The results confirm the findings of the scarce existing research on Internet adoption and usage for other European countries (Demoussis and Giannakopoulos 2006; OECD 2007), showing the need for the distinction between Internet use and the extent of use. Internet usage is mainly associated with high education levels, certain

age groups, some student and worker occupations, an elevated percentage of employment in service sector, Spanish nationality, living in an urban area and high GDP per capita in the county. In contrast, frequency of Internet usage is positively and mainly related to broadband connection, education and having advanced Internet competencies, differences in the ways through which Internet skills are acquired, gender, and population size. Frequency of use is negatively related to the presence of children in the household and to age. Moreover, if we compare the two models, only age has the same influence on both Internet usage and the extent of usage. Factors such as gender and family structure have a significant influence on the frequency of Internet usage, but not on its use. The education level shows a clearly linear relationship with Internet usage while education and Internet skills have a positive impact on the frequency of use. One of the main findings is that education and age are the largest determinants of Internet usage in Spain, whereas education and Internet skills, and broadband adoption are the main factors in explaining the frequency of Internet usage.

On the other hand, according to the second model, it is possible to assume that the determinants of the daily use of Internet are different from the relevant factors explaining less frequent use. For example, daily use of the Internet is positively determined by the broadband connection at home, higher education and Internet skills levels, and certain ways of acquiring Internet skills. For instance, having advanced Internet skills increases the daily use of the Internet and this effect is even more significant when the individual is more educated. Self-learning through trial-and-error, the use of manuals and books and other more formal methods, such as courses sponsored by employers, are the most common ways to acquire these skills. Being male and living in areas with high population density is positively related to daily use. In contrast, weekly and monthly use of Internet is statistically significant and

negatively influenced by the previous variables and positively influenced by acquiring Internet skills from colleagues, by the presence of children between 10 and 15 years in the household and by age.

In addition, the results allow us to compare Spanish Internet usage to that of other European countries. The empirical evidence is far from conclusive about the determinants of Internet usage in the EU. Generally speaking, our results coincide with the empirical evidence of the influence of age, education levels and occupations on Internet usage (Demoussis and Giannakopoulos 2006; Demunter 2005; Vicente and López 2006). Also, our analysis confirms the results obtained by Demunter (2005) on the lack of influence of gender on Internet usage, although for Southern European countries, Demoussis and Giannakopoulos (2006) found that men show a higher use than women. In addition, we can corroborate the evidence shown by Demunter (2005) and Arend and Steiner (2005) regarding the positive role of urban areas in explaining Internet usage. Looking at the frequency of Internet usage, our results are very similar to those obtained by Demoussis and Giannakopoulos (2006) for age, gender, education and size of population variables.

According to our findings, the distinction between Internet usage and the extent of usage in Spain within the EU framework may have interesting implications for helping to reduce the digital divide. First, if the policy purpose is to increase the number of people accessing the Internet, it may be convenient to develop specific strategies for low-educated, medium-age and old people, particularly in rural areas and for citizens not born in Spain. These specific factors may spread the perception of benefits of Internet use for these social groups and may influence their acquisition of some basic Internet skills. These actions are expected to develop outside of schools, such as in firms, at cultural associations, and at clubs for retired and immigrant associations. Also, general strategies oriented towards improving the educational level

of the population as well as towards fostering economic development and the services sector may have a positive impact on Internet adoption.

However, if the policy purpose is to increase the frequency of Internet usage, the first step would be to encourage broadband connection in homes by subsiding access to some specific social groups: medium-age and old people, females and people without higher education living in municipalities with less than 50,000 inhabitants. Also, the positive effect of developing Internet skills, particularly through self-teaching, the use of manuals and books and company training, should encourage government to support programs fostering lifelong learning attitudes among workers and students. Also, the implementation of basic Internet courses to foster Internet skills among specific groups, such as the low educational level groups, women and the elderly should be promoted. Some other initiatives should be oriented to integrate the Internet in the school as well as in Higher Education. It is necessary to promote ICT-based educational resources and teachers training as well as to convey the opportunities and utilities associated with the Internet throughout the Spanish society. At the government level, the development of e-government may provide additional utility and reasons for a greater frequency of Internet usage by many social groups.

The main limitation of the paper is the lack of data. At the individual level, information about variables such as income, Internet skills or Internet costs would enrich the analysis and would allow comparisons with other countries. From the supply side, information about the level of competency among Internet providers, the level of Internet infrastructure and the extent of development of broadband connection in comparison with other European countries would allow us to compare the demand and the supply sides of the usage patterns. At the county and national levels, it would be very useful to have information on regional variables, such as R&D intensity, level of Internet adoption by firms, and level of e-government development. Another

interesting variable would be the percentage of users in the region or in the county. This data would have allowed us to test epidemic models.

Further research may account for the existence of unobserved factors that may have power to explain the low degree of household Internet use in Spain. Cultural and attitudinal variables may influence the perceived benefits and utility. Other facts, such as the low level of e-government in Spain relative to that of the EU[6] may explain the low use of some Internet utilities. In addition, the deficiencies in the Spanish education system and its relationship with Internet skills deserve more attention in the academic literature on Internet adoption and use, considering the disappointing Spanish results in terms of education outcomes[7]. Further research is clearly needed on this topic, as well as on the differences registered in the level of Internet skills among countries.

Finally, further research might consider to study changes in Internet use and extent of use over time. To extend the scope of analysis provided in the present paper, we might use panel data to study the diffusion process of the Internet in Spain.

Acknowledgments

The authors would like to thank Professor Jaume Garcia, Eva Medina, Salvador Ortiz and Javier Hualde for their helpful comments. The authors acknowledge the helpful comments from two anonymous referees.

Notes

1. The model still relies on a normality assumption for the error term and functional form assumptions.
2. The frequency of use (T_i) refers to: daily frequency when the individual uses Internet at least five days per week; weekly when Internet is used at least once a week, but not every day; monthly when Internet is used at least once a month but not every week.
3. The categories shown are not exclusive: an individual may acquire Internet skills through different channels.
4. For further information regarding the convenience of choosing the share of individuals taking the value 1 as cutting point (Internet users), see Greene (2008).
5. For a general analysis of this problem, see Moulton (1990).
6. According to the UN-E government readiness, Arevalo, Fernández and Messía de la Cerda (2006) show that in 2004 within the EU-25 ranking, Spain was ranked 23.
7. In PISA reports (OECD 2000, 2003, 2006), Spain is usually placed in the last positions regarding students' acquisition of different skills.

References

Arend, M., and P. Steiner. 2005. Eurobarometer data on digital divides and ICT-disparities in EU15. *International Journal of Technology, Policy and Management* 5: 48-60.

Arevalo , G., E. Fernández, and J. Messía de la Cerda. 2006. The state of e-government and e-participation in European regions. Paper presented at the International Multi-Conference on Computing in the Global Information Technology, Bucharest, Romania. August 1-3.

Australian Bureau of Statistics, ABS (2002) Australia´s digital divide. Report. ABS, Sidney. http://www.abs.gov.au (accessed June 8, 2008).

Billon, M., R. Ezcurra, and F. Lera-López. 2008. The spatial distribution of the Internet in the European Union: Does geographical proximity matter?. *European Planning Studies* 16: 119-142.

Bimber, B. 2000. Measuring the gender gap on the Internet. *Social Science Quarterly* 81: 1-10.

Bucy, E.P. 2000. Social access to the Internet. *Harvard International Journal of Press and Politics* 5: 50-61.

Carveth, R., and S.B. Kretchmer. 2002. Policy options to combat the digital divide in Western Europe. *Informing Science* 5: 115-123.

Chaudhuri, A., K.S. Flamm, and J. Horrigan. 2005. An analysis of the determinants of Internet access. *Telecommunications Policy* 29: 731-755.

Demoussis, M., and N. Giannakopoulos. 2006. Facets of the digital divide in Europe: Determination and extent of Internet use. *Economics of Innovation and New Technology* 15: 235-246.

Demunter, C. 2005. The digital divide in Europe. *Statistics in Focus 38/2005*.

Dubin, J.A, and D. Rivers. 1990. Selection bias in linear regression, logit and probit models. In *Modern methods of data analysis*, ed. J. Fox and J.S. Long, 410-442. Newbury Park, CA: Sage Publications.

EUROSTAT. 2008. EUROSTAT Information Society Statistics. Office for Official Publications of the European Communities, Luxembourg. http://epp.eurostat.ec.europa.eu/portal/page/portal/information_society/introduction (accessed June 2, 2008).

Fairlie, R.W. 2004. Race and the Digital Divide. *Contributions to Economic Analysis and Policy* 3: 155-194.

Flamm, K., and A. Chaudhuri. 2007. An analysis of the determinants of broadband access. *Telecommunications Policy* 31: 312-326.

Forman, C. 2005. The corporate digital divide: Determinants of Internet adoption. *Management Science* 51: 641-654.

Forman, C., A. Goldfarb, and S. Greenstein. 2005. Geographic location and the diffusion of Internet technology. *Electronic Commerce Research and Applications* 4: 1-13.

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

Fundación Orange. 2007. *e-España 2007*. Madrid: Fundación Orange.

Gaspar, J., and E. Glaeser. 1998. Information Technology and the future of cities. *Journal of Urban Economics* 43: 136-156.

Goldfarb, A., and J. Prince. 2008. Internet adoption and usage patterns are different: Implications for the digital divide. *Information Economics and Policy* 20: 2-15.

Goolsbee, A., and P.J. Klenow. 2002. Evidence on learning and network externalities in the diffusion of home computers. *Journal of Law and Economic* 45: 317-343.
Greene, W. 2008. *Econometric Analysis*. Upper Saddle River, NJ.: Prentice Hall.
Grubesic, T.H. 2008. The spatial distribution of broadband providers in the United States: 1999-2004. *Telecommunications Policy* 32: 212-233.
Hacker, K.L., and R. Steiner. 2002. The Digital Divide for Hispanic Americans. *Howard Journal of Communications* 13: 267-283.
Hargittai, E. 2003. The digital divide and what to do about it. In *New Economy Handbook*, ed. C.J. Derek. 822-841.San Diego: Academic Press.
Heckman, J. 1979. Sample selection bias as a specification error. *Econometrica* 47: 153-161.
Hoffman, D.L., and T.P. Novak. 2000. The growing digital divide: Implications for an open research agenda. In *Understanding digital divide*, eds. E. Brynjolfsson and B. Kahin, 245-260. Cambridge: The MIT Press.
Horrigan, J.B., and L. Rainie. 2002. The broadband difference. How online Americans´ behavior changes with high-speed Internet connections at home. Pew Internet & American Life Project, Washington. http://www.pewinternet.org/ (accessed July 20, 2008).
Horrigan, J.B., and A. Smith. 2007. Home broadband adoption 2007. Pew Internet & American Life Project, Washington. http://www.pewinternet.org/ (accessed July 15, 2008).
Internet World Stats. 2008. Internet users in the World. March 2008. http://www.Internetworldstasts.com (accessed June 10, 2008).
Jorgenson, D.W., M.S. Ho, and K.J. Stiroh. 2008. A retrospective look at the U.S. productivity growth resurgence. *Journal of Economic Perspectives* 22: 3-24.
Karshenas, M., and P. Stoneman. 1992. The role of exogenous and endogenous learning and economic factors in the diffusion of new technology. *The Journal of Forecasting* 11: 577-601.
Karshenas, M., and P. Stoneman. 1995. Technological diffusion. In *Handbook of the economics of innovation and technological change*, ed. P. Stoneman, 265-297. Oxford: Blackwell.
Kiiski, S., and M. Pohjola. 2002. Cross-country diffusion of the Internet. *Information Economics and Policy* 14: 297-310.
Losh, S.C. 2003. Gender and educational digital chasms in computer and Internet access and use over time: 1983-2000. *IT & Society* 1: 73-86.
Madden, G., and G. Coble-Neal. 2003. Internet use in rural and remote Western Australia. *Telecommunications Policy* 27: 253-266.
Madden, G., and M. Simpson 1997. Residential broadband subscription demand: an econometric analysis of Australian choice experiment data. *Applied Economics* 29: 1073-1078.
Malecki, E. 2002. The economic geography of the Internet´s infrastructure. *Economic Geography* 78: 399-424.
Mansfield, E. 1968. *Industrial research and technological innovation*. New York: Norton.
Martin, S.P., and J.P. Robinson. 2004. The income digital divide: An international perspective. *IT&Society* 1: 1-20.
McKeown, L., A. Noce, and P. Czerny. 2007. Factors associated with Internet use: Does rurality matter?. *Rural and Small Town Canada Analysis Bulletin* 7(3), Ottawa. http://www.statcan.ca (accessed September 5, 2008).

Mills, B.F., and B.E. Whitacre. 2003. Understanging the non-metropolitan-metropolitan digital divide. *Growth and Change* 34: 219-243.
Moulton, B.R. 1990. An illustration of a pitfall in estimating the effects of aggregate variables on micro units. *Review of Economics and Statistics* 72: 334-338.
National Telecommunications and Information Administration, NTIA. 2004. A nation online: Entering the broadband age. US. Department of Commerce, NTIA, Washington, DC. http://ntia.doc.gov/reports.html (accessed June 10, 2008).
N-Economía. 2007. *Penetración regional de la nueva economía*. Madrid: CEPREDE.
Observatorio Redes. 2008. *Estudio sobre actividades realizadas en Internet*. Madrid: Observatorio Redes. http://observatorio.red.es/hogares-ciudadanos/articles/129 (accessed June 10, 2008).
OECD (2000, 2003, 2006) Programme for International Student Assessment (PISA). htpp:// www.oecd.org (accessed October 1, 2008).
OECD. 2007. Broadband and ICT access and use by households and individuals. Working Party on the Information Society. DSTI/ICCP/IE (2007)4/FINAL. Paris: OECD.
Ono, H., and M. Zavodny. 2007. Digital inequality: A five country comparison using microdata. *Social Science Research* 36: 1135-1155.
Rice, R. E., and J.E. Katz. 2003. Comparing Internet and mobile phone usage: Digital divides of usage, adoption, and dropouts. *Telecommunications Policy* 27: 597-623.
Rogers, E. 2003. *Diffusion of innovations*. New York: Free Press.
Rosenberg, N. 1972. Factors affecting the diffusion of technology. *Explorations in Economic History* 10: 3-33.
Schleife, K. 2006. Regional versus individual aspects of the digital divide in Germany. Discussion Paper 06-085, Centre for European Economic Research, ZEW. ftp://ftp.zew.de/pub/zew-docs/dp/dp06085.pdf (accessed June 16, 2008).
Singh, V. 2004. Factors associated with household Internet use in Canada, 1998-2000. Research Paper. Statistics Canada, Agriculture and Rural Working Paper Series, 66. Ottawa. http://www.statcan.ca (accessed September 4, 2008).
Spanish National Statistics Institute, INE. 2004. Encuesta de Población Activa. INE, Madrid. http://www.ine.es/inebmenu/indice/htm (accessed May 25, 2008).
Spanish National Statistics Institute, INE. 2008. *Encuesta sobre Equipamiento y Uso de Tecnologías de Información y Comunicación en los Hogares (TICH).* Madrid: INE.
Veenhof, B., Y. Clermont, and G. Sciadas. 2005. Literacy and digital technologies: Linkages and outcomes. Statistics Canada, Connectedness Series, 12. Ottawa. http://www.statcan.ca (accessed September 4, 2008).
Vicente, M.R., and A.J. López. 2006. Patterns of ICT diffusion across the European Union. *Economics Letter* 93: 45-51.
Vicente, M.R., and A.J. López. 2008. Some empirical evidence on Internet diffusion in the New Member States and Candidate Countries of the European Union. *Applied Economics Letters* 15: 1015-1018.
Zook, M.A. 2006. The geographies of the Internet. *Annual Review of Information Science and Technology* 40: 53-78.

Figure 1. Percentage of Internet users in the European countries. 2007

Source: Data adapted from EUROSTAT (2008).
Note: This variable represents the percentage of individuals between 16 and 74 years who have access to the Internet at least once per day in the last three months.

Figure 2. Internet users and GDP per capita in Spanish regions. 2007. Index Numbers

Source: Data adapted from INE (2008) and N-Economía (2007). National average = 100.

Figure 3. Trends in Internet users according to the Spanish educational levels

Source: Data adapted from INE (2008).

Figure 4. Stages in the decision making process of Internet use and Intensity of Internet use by individuals

Source: Author´s work.

Table I. Descriptive statistics* on sociodemographic characteristics according to the Internet usage in Spain, 2007

Independent variables			
Variable	*Categories*	*Non-Internet users* ($Y_i = 0$)	*Internet users* ($Y_i = 1$)
Age	(18-89 years)	35.31 (12.60)	57.79 (17.22)
Nationality	Spanish	48.27%	51.73%
	Non-Spanish	41.44%	58.56%
GenderAge (Gender and age groups)	M16-24	9.08%	90.92%
	M25-34	19.25%	80.75%
	M35-44	33.57%	66.43%
	M45-64	56.34%	43.66%
	M+65	92.36%	7.64%
	W16-24	7.69%	92.31%
	W25-34	20.46%	79.54%
	W35-44	39.73%	60.27%
	W45-64	65.29%	34.71%
	W+65	96.35%	3.65%
Educa (Education level)	No studies/qualifications and Primary	89.93%	10.07%
	Secondary	36.74%	63.26%
	Higher Education	12.43%	87.57%
Sitlab (Occupation)	Unemployed	43.03%	56.97%
	Student	1.59%	98.41%
	Housewife	79.73%	20.27%
	Retired	91.84%	8.16%
	Others	52.63%	47.37%
	Workers	29.95%	70.05%
Men10-15 (Children between 10 and 15 years in the household)	No	38.06%	61.94%
	Yes	49.67%	50.33%
Urban (Degree of urbanization)	<10,000 inhabitants	58.61%	41.39%
	10,000-20,000 inhabitants	52.04%	47.96%
	20,000-50,000 inhabitants	47.69%	52.31%
	50,000-100,000 inhabitants, no capital	43.72%	56.28%
	>100,000 inhabitants, no capital	44.78%	55.22%
	Other capitals	43.64%	56.36%
	>500,000 inhabitants, capital	39.29%	60.71%
National average		**47.89%**	**52.11%**
Number of observations		**10,629**	**11,568**

*Note: frequencies for the categorical variables, and mean and standard deviation (in brackets) for the only continuous variable (age).
Source: Data adapted from INE (2008).

Table II. Descriptive statistics* on sociodemographic characteristics according to the Internet frequency of use in Spain, 2007

Independent variables					
Variable	*Categories*		*At least once a week*	*Every week*	*Daily*
Gender	Woman		16.96%	29.28%	53.76%
	Man		12.98%	26.57%	60.46%
Age	(18-89 years)		36.57 (12.49)	35.79 (12.89)	34.31 (12.34)
Children between 10 and 15 years in the household	No		14.25%	27.14%	58.61%
	Yes		17.34%	30.86%	51.80%
Education and Internet competencies level	Primary education-Basic competencies		38.37%	41.34%	20.29%
	Primary education-Advanced competencies		16.33%	36.80%	46.87%
	Secondary education-Basic competencies		34.34%	36.97%	28.69%
	Secondary education-Advanced competencies		10.05%	27.61%	62.34%
	Higher education-Basic competencies		21.82%	34.06%	44.13%
	Higher education-Advanced competencies		4.24%	19.26%	76.50%
Degree of urbanization	High		12.70%	25.55%	61.75%
	Medium		17.56%	29.06%	53.38%
	Low		18.06%	33.02%	48.92%
Access to the Internet with broadband connection	From home		8.25%	24.78%	66.97%
	From others		22.76%	31.31%	45.93%
	With other devices		26.25%	34.10%	39.65%
Means for Internet skills acquisition	Regular training	No	16.97%	29.50%	53.54%
		Yes	9.89%	24.92%	65.19%
	Courses for adults	No	13.94%	27.64%	58.42%
		Yes	15.71%	28.41%	55.87%
	Company training	No	16.21%	29.32%	54.47%
		Yes	9.26%	23.59%	67.15%
	On his own (books, manuals)	No	18.05%	31.29%	50.65%
		Yes	8.90%	22.59%	68.51%
	By practicing	No	27.31%	31.47%	41.22%
		Yes	12.95%	27.44%	59.61%
	Learning from others	No	14.44%	26.96%	58.60%
		Yes	14.45%	28.13%	57.42%
	By other means**	No	14.59%	28.03%	57.38%
		Yes	8.69%	21.14%	70.17%
National average			**14.82%**	**27.82%**	**57.36%**
Number of observations			**1,309**	**2,458**	**5,069**

*Note: frequencies for the categorical variables, and mean and standard deviation (in brackets) for the only continuous variable (age).

**Note: Although the survey provides this option, this category will not be finally used in the later analysis due to its small weight (only 3% of users) compared to the rest of means for Internet skills acquisition.

Source: Data adapted from INE (2008).

Table III. First stage: Probit model estimation results of Internet usage in Spain, 2007

Variable	Categories	Marginal Effect	Robust standard error	t-ratio	p-value
GDP_pc_1000		0.0087	0.0023	3.76	0.000
% Employment in service activities		0.2704	0.0985	2.74	0.006
Gender (Men/women) and age groups (in years)	M16-24	0.5182	0.0189	13.88	0.000
	M25-34	0.4804	0.0174	20.06	0.000
	M35-44	0.3826	0.0256	12.40	0.000
	M45-64	0.3073	0.0289	9.63	0.000
	M+65	0.0725	0.0278	2.59	0.010
	W16-24	0.5218	0.0147	17.77	0.000
	W25-34	0.4636	0.0218	15.14	0.000
	W35-44	0.3654	0.0294	10.33	0.000
	W45-64	0.2871	0.0309	8.62	0.000
Nationality	Spanish	0.0839	0.0438	1.89	0.059
Education	Secondary	0.3892	0.0181	20.15	0.000
	Higher Education	0.6179	0.0162	30.76	0.000
Occupation	Unemployed	0.1673	0.0310	5.07	0.000
	Student	0.4898	0.0151	13.38	0.000
	Housewife	0.0150	0.0317	0.47	0.636
	Others	0.1272	0.0380	3.22	0.001
	Workers	0.2663	0.0199	12.87	0.000
Children between 10-15 years	Yes	0.0383	0.0243	1.57	0.116
Degree of urbanization	10,000-20,000 hab.	0.0024	0.0228	0.10	0.917
	20,000-50,000 hab.	0.0392	0.0290	1.35	0.178
	50,000-100,000 hab, no capital	0.0661	0.0192	3.41	0.001
	>100,000hab., no capital	0.0899	0.0165	5.37	0.000
	Other capitals	0.0847	0.0193	4.36	0.000
	>500,000 hab., capital	0.0973	0.0144	6.71	0.000
Number of observations	22,187				
Log pseudo-likelihood	-8.073				
Wald-Chi2 (26)	9341.44				
Prob>Chi2	0.000				
Pseudo R-Square	0.4744				
% corrected predictions (Yi=1)	85.87%				
% corrected predictions (Yi=0)	79.75%				
Reference category: Spanish retired woman over 64 years with no studies or with primary studies, without children in the household between 10-years, and living in cities with less than 10,000 inhabitants.					

Source: Author´s work adapted from INE (2008).

Table IV. Second stage: Ordered probit model estimation results of Internet frequency, corrected for selectivity, in Spain, 2007

		Monthly frequency			Weekly frequency			Daily frequency		
Variable	**Categories**	**Marginal effect**	**t-ratio**	**p-value**	**Marginal effect**	**t-ratio**	**p-value**	**Marginal effect**	**t-ratio**	**p-value**
Gender	Man	-0.0180	-3.6	0.000	-0.02124	-3.63	0.000	0.03920	3.63	0.000
Age		0.0042	3.85	0.000	0.0050	3.84	0.000	-0.00915	-3.86	0.000
Age squared		0.0000	-3.79	0.000	-0.0001	-3.79	0.000	0.0001	3.8	0.000
Children between 10-15 years	Yes	0.0288	4.19	0.000	0.0311	4.61	0.000	-0.0599	-4.42	0.000
Education & Internet competencies	Primary education-Advanced competencies	-0.0698	-8.06	0.000	-0.1272	-5.34	0.000	0.1971	6.12	0.000
	Secondary education-Basic competencies	0.0026	0.15	0.880	0.0031	0.15	0.879	-0.0057	-0.15	0.880
	Secondary education-Advanced competencies	-0.1046	-7.89	0.000	-0.1463	-7.16	0.000	0.2509	7.54	0.000
	Higher education-Basic competencies	-0.0385	-2.3	0.021	-0.0541	-1.97	0.049	0.0926	2.1	0.036
	Higher education-Advanced competencies	-0.1466	-9.7	0.000	-0.2064	-9.39	0.000	0.3530	9.69	0.000
Degree of urbanization	High density	-0.0313	-4.87	0.000	-0.0361	-5.02	0.000	0.0673	4.98	0.000
	Medium density	-0.0063	-0.91	0.364	-0.0077	-0.89	0.373	0.0140	0.9	0.369
ADSL access to Internet	At home	-0.0898	-8.97	0.000	-0.0976	-9.89	0.000	0.1874	9.62	0.000
	From other places	0.0132	1.39	0.164	0.0154	1.42	0.156	-0.0286	-1.41	0.159
Means for Internet skills acquisition	Regular training	-0.0162	-2.89	0.004	-0.0198	-2.82	0.005	0.0360	2.86	0.004
	Courses for adults	0.0055	1.02	0.310	0.0064	1.03	0.303	-0.0119	-1.02	0.306
	Company training	-0.0421	-8.19	0.000	-0.0561	-7.36	0.000	0.0982	7.8	0.000
	Learning on his own (books and manuals)	-0.0440	-9.02	0.000	-0.0547	-8.61	0.000	0.0987	8.94	0.000
	Learning by practicing	-0.0574	-5.63	0.000	-0.0540	-7.18	0.000	0.1114	6.35	0.000
	Learning from other people	0.0302	5.87	0.000	0.0394	5.38	0.000	-0.0696	-5.62	0.000
Lambda parameter		0.0410	2.85	0.004	0.0488	2.85	0.004	-0.0898	-2.86	0.004
Number of observations	8,706.00									
Log pseudo-likelihood	-7,175.04									
Wald-Chi2 (17)	2,184.34									
Prob>Chi2	0.0000									
Pseudo R-Square	0.1321									

Reference category: woman, no studies or with primary studies and basic Internet competencies, without ADSL access and without children between 10-15 years, living in rural areas (low-population density areas) and acquiring Internet skills by other means.

Source: Author´s work adapted from INE (2008).