

Serrano-Cinca, C.; Muñoz-Soro, J.F.; Brusca-Alijarde, I.

Conference Paper

A multivariate study of the digital divide and digital exclusion

26th European Regional Conference of the International Telecommunications Society (ITS): "What Next for European Telecommunications?", Madrid, Spain, 24th-27th June, 2015

Provided in Cooperation with:

International Telecommunications Society (ITS)

Suggested Citation: Serrano-Cinca, C.; Muñoz-Soro, J.F.; Brusca-Alijarde, I. (2015) : A multivariate study of the digital divide and digital exclusion, 26th European Regional Conference of the International Telecommunications Society (ITS): "What Next for European Telecommunications?", Madrid, Spain, 24th-27th June, 2015, International Telecommunications Society (ITS), Calgary

This Version is available at:

<https://hdl.handle.net/10419/127178>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

A multivariate study of the digital divide and digital exclusion

C. Serrano-Cinca

*Full Professor at the University of Zaragoza,
Department of Accounting and Finance, Fac. Economía y Empresa, Gran Vía 2, Zaragoza (50.005) Spain,
serrano@unizar.es <http://ciberconta.unizar.es/charles.htm>*

J.F. Muñoz-Soro

*Researcher at the Aragonese Agency for Research and Development (ARAID)
P. T. Walqa, Ctra. Zaragoza N-330a, km. 566 22197 Cuarte (Huesca) Spain.
jfm@unizar.es*

I. Brusca-Alijarde

*Professor at the University of Zaragoza,
Department of Accounting and Finance, Fac. Economía y Empresa, Gran Vía 2, Zaragoza (50.005) Spain,
ibrusca@unizar.es, Phone: 0034976762321*

ABSTRACT

This paper explores the use of Internet, including social networks, e-government services and e-commerce from the perspective of the digital divide. First, it aims to find out what factors explain the digital divide, considered as a multidimensional concept. Second, the article aims to identify the groups that are digitally excluded. The paper is based on survey data (N=2,304) collected in Spain that are analyzed by using multiple regression, principal component analysis and cluster analysis. Two dimensions are identified: the first is the comprehensive use of Internet and the second is the nature of this use, differentiating between a professional use and a recreational and social use of Internet. The paper verifies that factors explaining the digital divide are age, education level, and income. Many unemployed are among the most frequent users of Internet. The most serious problem is encountered in homeworkers, who are mainly woman. NEETs (Not in Education, Employment, or Training) are frequent users of Internet but they only use it for entertainment and to certain extent they are digitally excluded.

KEYWORDS

Digital divide, digital exclusion, e-government, social networks, e-commerce.

INTRODUCTION

The term *digital divide* refers to inequalities in Internet access and the use of Information and Communication Technologies (ICT) that some segments show because of socio-demographic, economic and geographical factors. Although around 50% of European citizens use Internet daily, 30% of the population have never accessed it (European Commission, 2014). Many authors warn that digital exclusion can lead to social exclusion (Klecun, 2008), highlighting that more and more services and goods are delivered electronically and that these are often not available in other ways, an example being e-government online forms. Before designing out public policies, it is necessary to identify the collectives that are digitally excluded, which is the aim of this paper.

The digital divide has been widely studied, especially its drivers (Chinn and Fairlie, 2006; Klecun, 2008; Van Dijk, 2009; Hilbert, 2011b; and Vicente and López, 2011). At first, studies focused on the first-level digital divide between “have” and “have-nots” access to digital information and communications technology (Riggins and Dewan, 2005). Most of these studies reach similar conclusions (Van Dijk, 2009): people with senior age, with lower educational level, positions outside the labor market, and to a lesser extent females and from ethnic minorities, have less access to Internet. But different environments can modify the importance of these factors (Hilbert, 2011a). Rice and Katz (2003) studied Internet access with the mobile phone concluding that the use of Internet and mobile phones was very similar and that a digital divide exists with respect to both Internet and mobile phone usage.

Several researchers propose going beyond access to computers and Internet (Valadez and Durán, 2007) and the second-level digital divide is concerned with differences in people’s online skills, the “digital capability divide” (Attewell, 2001). Furthermore, in Wei et al. (2011) a third-level digital divide has been added to consider the “digital outcome divide”. More specific uses of the Internet include participation in social networks, the use of e-government services and e-commerce. Monsuwé et al. (2004) find that age, gender, education, and income are the factors influencing consumers’ attitude toward online shopping. So, uneven diffusion of e-commerce is creating a further digital divide (Gibbs et al, 2003). E-government is a particularly important case because although Internet is a tool to facilitate greater citizen-participation in government, there are many social groups that cannot benefit from electronic services or participatory activities; hence, the digital divide may affect the ways e-government programs are implemented (Helbig et al., 2009). Taipale (2013) finds that the level of education, income, and city size, explain the use of e-government services. Similar results are obtained by Bélanger and Carter (2009). Ahn (2011) analyzes the digital divide and social

network sites focusing on young people, who are the biggest users of social networks, and concludes that traditional divide indicators are not significant predictors of social network use.

Our study focuses on the digital outcome divide with two objectives. Firstly, we aim to identify socio-demographic and economic variables related to the digital divide in Spain. This is done by analyzing 63 items of Internet use included in a questionnaire (N=2,304) carried out in November 2013. We focus on the type of Internet use, including social networks, e-government services and e-commerce, making up a multidimensional framework. A biplot of principal component analysis (biplot-PCA) allows us to understand the internal structure of Internet use and of the digital divide, which is a contribution of the paper with respect to previous literature. The paper identifies two dimensions of Internet use. The first is related to the comprehensive use of the Internet (including most of the indices commonly used to measure the level of Internet use) and the second dimension differentiates between professional uses (e-commerce and e-government), and recreational uses –whose culmination is the social use of Internet.

The second objective of the paper is to identify digitally excluded segments, which constitutes a serious problem because of the risk of social exclusion (Warren, 2007). Policies to increase Internet use are often aimed at socially disadvantaged groups, but these groups do not always coincide with digitally excluded groups. The use of cluster analysis and a biplot-PCA allow us to obtain different clusters that reflect the levels of use of Internet. Analyzing these clusters and their patterns, we identify the digitally excluded individuals in Spain, which is another contribution of the paper.

The rest of this paper is structured as follows. The next section reviews previous literature and develops the hypotheses. In Section 3, the empirical study is presented, beginning with the sample and methodology and showing the results obtained. In the final section, the conclusions are discussed.

2. LITERATURE REVIEW AND HYPOTHESES DEVELOPMENT

Universal digital inclusion is a basic objective of society and particularly of public administrations, who have initiated plans towards this end. In 2010, the European Commission adopted the Europe 2020 strategy for smart, sustainable and inclusive growth. It includes the Digital Agenda for Europe, a strategy to take advantage of digital technology that contains 7 priority areas, the sixth of which is enhancing digital literacy, skills and inclusion. In Spain the Digital Agenda for Europe was approved in 2013 together with a plan for digital inclusion and employability. Spain is in a medium-low position compared to other EU

countries: out of 30 countries, it is in 22nd position in households with access to the Internet at home, 12th in individuals interacting online with public authorities, 16th in participating in social networks and 20th in using online banking (European Commission, 2014).

Various theories have been used to explain the digital divide (see Riggins and Dewan, 2005, or Van Dijk, 2013, for a review). The most accepted is the Diffusion of Innovations (DOI) of Rogers (2003). Diffusion is the process by which an innovation is communicated through certain channels over time among the members of a social system. The spread of an innovation follows an S-shaped curve. Rogers identifies five categories of adopters: innovators, early adopters, early majority, late majority and laggards. DOI theory can be used to explain the digital divide, as in Hilbert (2011a). Under DOI theory, younger users are placed in the group of innovators or early adopters while elderly users are included in the group of laggards. Age is considered to be one of the most important determining factors to explain the digital divide. Loges and Jung (2001) question whether the digital divide is a generational phenomenon that will disappear in time. If so, age *per se* is not the relevant variable. However, the authors' analysis suggests that observable differences between the ways old and young people use their Internet connections may survive the present generation of elderly, concluding that the problem of the age gap cannot be solved by providing access through technology and training. Czaja and Lee (2007) argue that many interfaces are designed without accommodating the needs of the elderly population.

In accordance with the above literature, we can hypothesize:

H1. Older people are less likely to use Internet services than younger ones. However, the youngest users do not participate in activities such as e-government and e-commerce, so a non-linear relation is expected in more sophisticated Internet use.

Differences due to education are supported by DOI but also by the Knowledge Gap Theory, which affirms that knowledge is considered more important than other benefits, such as consumption and entertainment (Van Deursen and Van Dijk, 2014). For the advocates of this theory, the number of hours you are online is not so important as the type of activities carried out. According to Van Dijk (2009), Internet users with a low level of education perform fewer activities of information retrieval, text communication and financial services than users with medium and high education levels. Van Dijk (2012) also appeals to the Resources and Appropriation Theory to explain the digital divide. An important concept implicit in this theory is user skills, which can explain differences arising from education, given that people with higher education perform better in all skills than people with a lower educational background (Van Dijk, 2012). Hoffman et al (2000) argues that more skill is

required to shop online than to search, which explains why education is an important factor in activities that require the application of some knowledge.

This leads us to propose our second hypothesis:

H2. Highly educated people are more likely to use Internet services than less educated people, especially in uses that require the application of knowledge.

DOI theory also explains how differences in income level may influence Internet usage, considering that financial status helps to absorb innovation quickly (Rogers, 2003) and can influence physical and material access (Van Dijk, 2012). In this sense, Chinn and Fairlie (2006) find that the digital divide between different countries is mainly explained by income differentials. Fuchs (2008) also states that income inequality is a factor that influences the level of the digital divide. Other authors have used the Internet Consumption Theory to explain differences in Internet diffusion, where the level of income plays an important role in the equilibrium for consumers who try to maximize their utilities within the constraints of the income budget (Zhang, 2013).

In line with the above literature, the following hypothesis is proposed:

H3. People with a high household income are more likely to use Internet than those with a low household income, especially in Internet uses that involve monetary transaction such as e-commerce.

The use of Internet is also related to job type. The Technology Acceptance Model (TAM) of Davis (1989) has been used to explain the digital divide, as in Porter and Donthu (2005) and Sipior et al. (2011). According to TAM, the perceived ease of use and perceived usefulness can influence the attitude toward using and the actual use of a technology. TAM theory can also explain differences in perceived usefulness by the type of job, as different situations can lead to different perceived usefulness. Each labor group may perceive Internet as a useful tool for their work in different ways. Hence, full time workers and students will take more advantage of Internet and perceive it to be more useful than those with a less intensive labor situation, pensioners and housewives. In fact, Van Dijk (2009) finds big differences in physical access between European students, employees, and the self-employed on the one hand and European unemployed, retired, and inactive people on the other. This leads us to propose the following hypothesis:

H4. Active labor groups, such as fully employed workers and students are more likely to use Internet services than pensioners and housewives.

The gender digital divide has been widely studied and has been explained though

sociological theories, such as the Social Cognitive theory. This theory argues that gender-based behavior is developed through repeated modeling of prototypical behavior associated with some gender models (Bandura, 1986). Females have been associated with lower computer usage and stereotype attitudes (Felter, 1985; Hunt and Bohlin, 1993). Cooper (2006) affirms that research published in the last 20 years draws the conclusion that females are at a disadvantage when learning about computers, which can be explained by the socialization patterns of boys and girls associated with the stereotype of computers as toys for boys. However, Hilbert (2011b) affirms that, although some claim that women are rather technophobic and many studies find differences between men and women, these differences often disappear when other variables are controlled for. Van Dijk (2009) states that the gender gap in Europe has closed for the youngest age groups of 16-24, but not for older age groups.

We propose the following hypothesis:

H5. Gender is related to the use of Internet services.

The rural-urban digital divide was important in the early stages of Internet, but it has narrowed, as is shown in Warren (2007) and Vicente and López (2011). Urban Density theory and Global Village theory are used to explain the evolution of the rural-urban digital divide (Forman et al., 2005). Urban Density theory argues that Internet was first diffused through urban areas with complementary technical and knowledge resources that lowered the costs of investing in new frontier technology (Forman et al., 2005). Global Village theory asserts that Internet technology helps lower communication costs and breaks down geographic boundaries. Forman et al. (2005) find some evidence supporting Global Village theory in the diffusion of basic Internet access, but the pattern of adoption of frontier Internet technologies supports Urban Density theory.

We can define the following hypothesis:

H6. People living in urban environments use Internet services more than those living in rural environments, especially in more sophisticated uses.

3. EMPIRICAL STUDY

3.1 Sample and data

The paper is based on survey data (N=2,304) collected in the Autonomous region of Aragon, Northern Spain, in November 2013. Respondents were selected using random tri-ethnic sampling. In the analysis of the results, this method allows the differentiation of groups defined by belonging to: one of the two sexes, a rural or urban environment and an age group.

By design, the survey has an expected error margin of +/- 2.00% for the total sample, with 95% confidence.

The survey includes 63 items on the use of Internet. The items are grouped into five categories, allowing the construction of five indices: generic use (25 items on different Internet uses), e-commerce (14 items on products purchased), social network use (12 items on the use made of social networks), use of Internet on mobile (6 items), use of e-government (6 items). Finally, a sixth index was obtained as the sum of the 63 variables measuring the level of overall use. Table 1 shows the variables used to construct the indexes.

*** Table 1 ***

3.2 Exploratory analysis and hypotheses testing

The main characteristics of the data are presented in Table 2. There are five age groups, 4 levels of education, 4 levels of household income, 8 types de employment, rural and urban environments, and gender. The table shows the result of a t-test of difference of means.

All the socio-economic factors associated with the digital divide show statistically significant difference with the sign predicted by theory. Age is a very important factor. A linear relation can be observed between age and the use of social networks and the mobile: the youngest are the biggest users of both, confirming our first hypothesis. But the relation is not linear when considering online shopping, e-administration and the generic use of Internet. The biggest Internet shoppers are the group of 25-34 years old, who are also the biggest users of Internet and those who use e-government most. The higher the level of education, the greater the level of Internet use, which supports our second hypothesis. Two groups can be observed: those of medium or high levels and those of low level or uneducated. With respect to income the group with less than 1,100 €/month stands out from the rest (H3 accepted). The highest level of Internet use is found among student workers. The digital divide affects the retired and homemakers (H4). In contrast to other empirical studies, such as Van Dijk (2009), the values for Spanish unemployed people are even higher than for other groups, such as autonomous or part-time workers. The reason is that, in Spain, the period analyzed corresponds to a severe crisis situation in which unemployment reached 24% in total and 54% among young people, meaning that many unemployed people are highly qualified and not marginalized. Men obtain higher scores than women except in the use of social networks (H5 accepted). The urban environment scores higher than the rural. According to the exploratory study carried out, the 6 hypotheses proposed are accepted within the data analyzed.

	% (N)	Generic use	E-commerce use	Social-network use	Internet on mobile use	E-government use	Total Overall use of Internet
Full sample	100% (2.304)	9.66	0.91	2.70	2.25	1.82	17.35
Age							
15-24 year-olds	12.8% (296)	13.26***	0.91	6.04***	4.02***	1.78	26.01***
25-34 year-olds	18.2% (420)	14.41***	1.64***	5.00***	3.89***	2.65***	27.59***
35-44 year-olds	21.3% (491)	12.08***	1.22***	2.96	2.89***	2.43***	21.58***
45-54 year-olds	20.1% (463)	9.11**	0.87	1.43***	1.48**	1.97**	14.86***
55-64 year-olds	15.5% (356)	4.89***	0.33***	0.48***	0.63***	1.10***	7.43***
65-74 year-olds	12.1% (278)	1.39***	0.04***	0.22***	0.13***	0.24***	2.01***
Education							
Uneducated	1.3% (31)	3.61***	0.33***	0.90***	0.81***	0.58***	6.23***
Low level	36.3% (836)	5.60***	0.33***	1.67***	1.32***	0.80***	9.71***
Medium level	37.6% (867)	10.92***	0.95	3.14***	2.62***	2.05***	19.68***
High level	24.6% (567)	14.06***	1.73***	3.67***	3.15***	3.07***	25.69***
Don't know	0.1% (3)	5.67	0.00	0.00***	0.00***	0.67	6.33
Household income							
Less than 1,100 €/month	19.4% (448)	5.60***	0.43***	1.59***	1.16***	.98***	9.76***
1,100-1,800 €/month	20.8% (480)	10.01*	0.90*	2.47	2.06	1.85	17.29
1,800-2,700 €/month	11.9% (275)	12.63***	1.59***	3.10***	2.89***	2.64***	22.84***
More than 2,700 €/month	6.4% (147)	13.96***	2.01***	3.20***	3.15***	3.17***	25.48***
Don't know	41.4% (954)	9.88	0.77***	3.15***	2.54***	1.77	18.11**
Employ							
Full-time worker	33.9% (780)	12.00***	1.32***	2.93**	2.74***	2.51***	21.49***
Retired	15.3% (352)	3.89***	0.30***	0.64***	0.54***	0.78***	6.15***
Unemployed	15.0% (345)	11.94***	1.06	3.58***	2.72***	2.07***	21.38***
Homemaker	12.7% (292)	4.36***	0.32***	0.85***	0.79***	0.68***	7.00***
Student	9.7% (223)	13.04***	0.97	5.87***	3.90***	1.75	25.53***
Autonomous worker	9.1% (210)	9.81	0.93	2.57	2.40	2.23***	17.94
Part-time worker	3.3% (76)	10.54	0.79	2.93	2.84**	1.87	18.97
Student worker	1.0% (24)	15.13***	1.04	6.83***	4.29***	2.38	29.67***
Don't know	0.1% (2)	9.66	0.91	2.70	2.25	1.83	17.35
Gender							
Men	51.0% (1.175)	10.16***	0.92	2.69	2.31	2.04***	18.11***
Women	49.0% (1.129)	9.14***	0.89	2.72	2.19	1.60***	16.55***
Ambit							
Urban	73.2% (1.686)	10.24***	0.99***	2.83***	2.41***	1.93***	18.40***
Rural	26.8% (618)	8.07***	0.67***	2.37***	1.82***	1.53***	14.47***

Table 2. Data characteristics: Mean and T-test results significant (***) significant at 1%; ** significant at 5%; * significant at 10%)

3.3 Factors explaining the digital divide. Multivariate regression

A single indicator is not enough to reflect the multidimensional nature of the digital divide. Six regressions have been performed, taking each of the six indices as the dependent variable. The socio-economic factors have been introduced as independent variables. To investigate the non-linear relation of age, the square of age is included as a variable. The principle of parsimony has been applied and when faced with two similar possible models, the simplest is preferred because it includes fewer explanatory variables. The following table

shows the final results of each regression.

Dependent Variable	Model 1 <i>Generic use</i>	Model 2 <i>E-commerce use</i>	Model 3 <i>Internet on mobile use</i>	Model 4 <i>Social-network use</i>	Model 5 <i>E-government use</i>	Model 6 <i>Total use of Internet</i>
Constant	8.006***	0.269	4.980***	7.222***	-0.409**	16.398***
Age	-	-	-0.088***	-0.125***	-	-
Age*age	-0.003***	-0.001***	-	-	-0.001***	-0.005***
Education level	2.040***	0.325***	0.192***	0.370***	0.699***	3.519***
Household income	1.090***	0.291***	0.340***	-	0.289***	1.987***
Homemaker	-1.818***	-	-	-	-	-2.989***
Part-time worker	-	-0.566**	-	-	-	-3.126**
Full-time worker	-	-	-	-	0.424***	-
Autonomous worker	-	-	-	-	0.474***	-
Unemployed	-	-	-	-	0.364***	-
Gender (ref. men)	-	-	-	-	0.297***	-
Ambit (ref. urban)	-	-	-	-	-	-
Adjusted R2	0.529	0.180	0.388	0.395	0.391	0.562
N	2,304	2,304	2,304	2,304	2,304	2,304

Table 3: Regression analysis coefficients. *** significant at 1%; ** significant at 5%; * significant at 10%.

The table reveals that age is the most important variable. It is included as an explanatory variable in all the models and it is always significant at 1%. The older the person, the lower the use of social networks and of the mobile, the relationship being linear. The square of age is significant in the use of e-government, e-commerce and generic use of Internet. This corroborates that the youngest users are not always the biggest users. The next most important factor is the educational level, which is also included as an explanatory variable in all the indices. The third factor that explains the level of use of the technologies is the income level, in all but the use of social networks. The type of employment is also important, especially in homemakers whose coefficient has a negative sign when explaining the level of overall use of the Internet. This group shows an important digital divide and, as most of them are women, this explains the gender digital divide to a great extent. Part-time work also explains the digital divide and appears as an explanatory variable in the level of e-commerce and in that of overall use. Three types of employment explain the higher use of e-government: full-time work, autonomous and unemployed, the latter possibly for administrative tasks related to unemployment or social security. Gender only appears as an explanatory variable in the level of use of e-government. The rural environment does not enter as an explanatory variable in the parsimonious regression models.

In summary, the level of overall use is explained by five variables: the square of age,

educational level, income level, homemakers and part-time work. But the level of use of social networks is explained by only two variables: age and educational level. **Income level is not an explanatory variable of the use of social networks**; in this respect, it is an egalitarian technology. However, Internet shopping is explained by income level because it is an activity that requires money. For the level of use of e-government, as well as age, education and income, it is necessary to take into account both type of job and gender.

3.4 Identification of the digitally excluded

Patterns of Internet usage are very different among groups of users. Van Deursen and Van Dijk (2014) identify the following Internet usage types: information, news, personal development, social interaction, leisure, commercial transaction and gaming. Multidimensional data should be analyzed by using multivariate statistical tools. In our paper, the identification of patterns of Internet usage has been performed through PCA and the identification of the digitally excluded through cluster analysis. PCA allows us to reduce the dimensionality of a set of data obtaining a few components that explain, to a great extent, the variability of the data. To interpret the meaning of each principal component and the different zones of the map, we use biplot-PCA, a technique that represents the observations and variables simultaneously, the latter plotted in the form of vectors.

Spanish public policy actions to reduce the digital divide are mainly aimed at the following collectives: women, people in rural environments and the unemployed. However, it is important to carry out deeper studies because not all the members of these groups show digital divide.

We obtained 71 subgroups from crossing all the variables: gender and educational level, gender and income level, gender with age and type of work, and gender with environment and age. An example of a subgroup is “woman from rural environment, aged 25 to 34”. For each of these 71 subgroups we calculate the average for the 63 questions of the survey. In this way, we have a multivariate table with 71 cases and 63 variables.

The upper right-hand side of Figure 1 is related to the variables of Internet shopping and professional uses such as e-banking and e-administration. The lower right-hand side of the figure explains variables such as the use of social networks, the downloading of games and listening to music. The figure allows us to distinguish between “online shopping and professional use” and “recreational and social Internet use”.

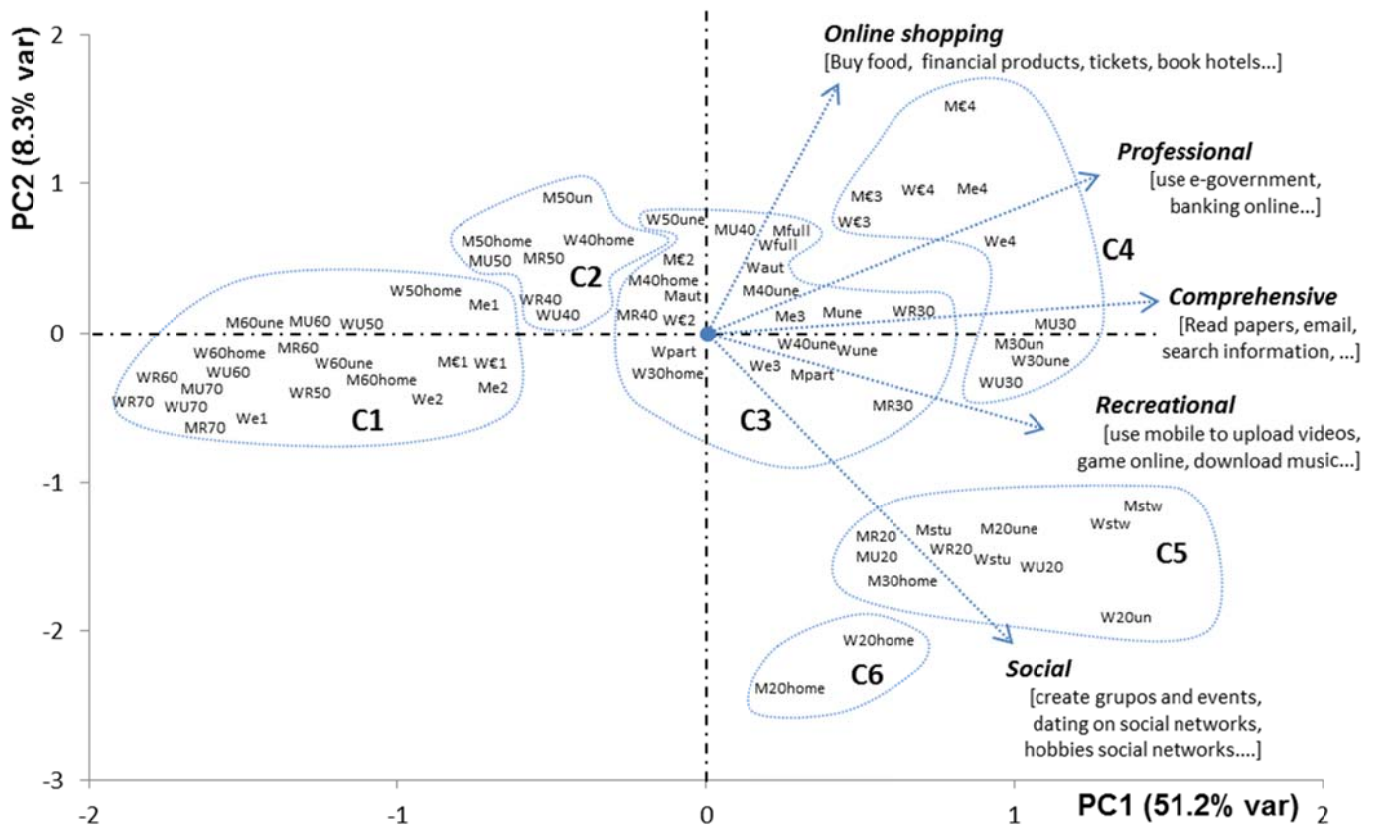


Figure 1. Results of Biplot-Principal Component Analysis, and Cluster Analysis.

Acronyms: **W** women; **M** men; **R** rural; **U** urban;

€1 less than 1,100€ **€2** 1,100-1,800 € **€3** 1,800-2,700€ **€4** more than 2,700€

E1 uneducated; **E2** low education; **E3** middle education; **E4** high education;

20 15-24 years; **30** 24-35 years; **40** 35-44 years; **50** 45-54 years; **60** 55-64 years; **70** 65-74 years

une unemployed; **part** part time job; **aut** autonomous; **full** full time job; **stu** student; **stuw** student worker; **home** homemaker

Figure 2 shows the dendrogram resulting from carrying out a hierarchic cluster analysis. Ward's method has been taken as the grouping algorithm and Euclidean distance as the measure of similarity. Six clusters, C1, C2, C3, C4, C5 and C6, have been identified. The results of the cluster analysis have been superimposed on Figure 1.*

Cluster 1 identifies the collectives affected by digital exclusion. In Figure 1, they are located on the extreme left of the map. They include people over the age of 55, those with the lowest incomes and those without studies or of only primary education, regardless of gender. There is no gender gap in people over 55: both men and women of this age, on average, show the lowest levels of Internet use. A gender gap does exist in the 45-54 age group, especially in the homemakers collective, which in Spain consists mainly of women. Unemployed people and homemakers of over 55 also belong in this cluster. In the group of digitally excluded, there is no distinction between rural and urban environments, both being equally affected. Efforts and actions of digital literacy to eradicate the digital divide should be carried out among these collectives.

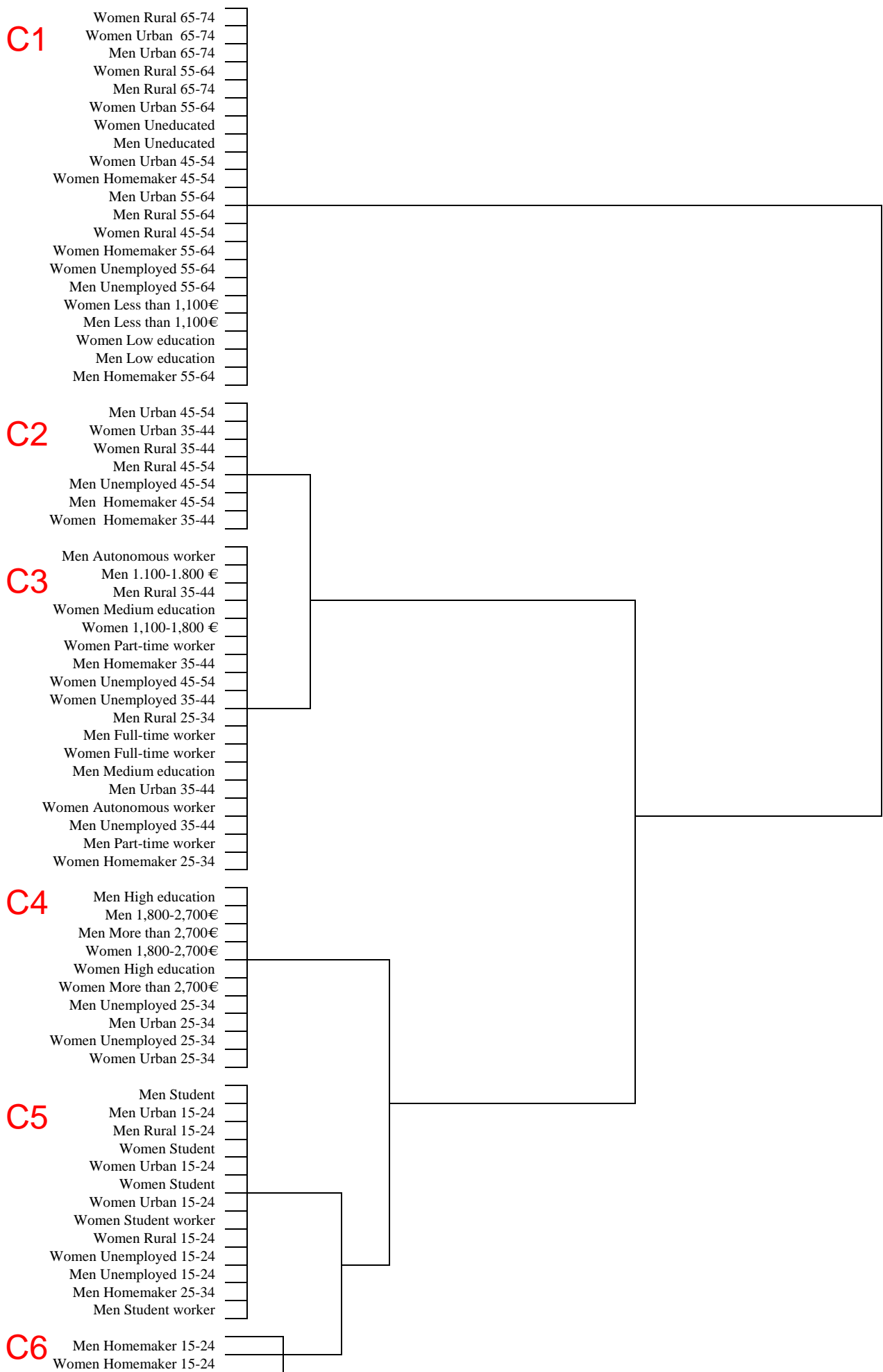


Figure 2. Hierarchical cluster analysis of 71 groups, using Ward's method.

The following cluster, C2, is made up of men between 45 and 54 and women between 35 and 44, both regardless of whether they reside in urban or rural areas. It also includes homemaker men of 45-54 and homemaker women of 35-44, as well as unemployed men of 45-54. In Figure 1, they are located to the left in the zone of very little Internet use. These groups also need specific formative actions because they are in danger of digital exclusion.

The third cluster, C3, contains people with a medium educational level, those with an income between 1,100-1,800€, full-time, part-time and autonomous workers, unemployed people between 35 and 44, women homemakers between 25 and 44, unemployed women between 45 and 54 and people aged 25 to 34 from a rural environment. This cluster occupies the center of Figure 1 and represents the Spanish average. It does not show the digital divide but does not make full use of the possibilities of Internet.

The fourth cluster, C4, is made up of people with a university education, independent of gender, age and other conditions; people with the highest income levels, also regardless of age; and all urban men and women between 25 and 34, regardless of education and income. It is the cluster that uses Internet most widely. We can highlight that unemployed people between 25 and 44 also belong to this cluster; in fact, they are the only labor group included in it. The cluster is located on the right of the map in the zone of maximum use and shows that, between men and women with incomes over 1,800 and those with university education, there is no gender gap.

Cluster C5 is very homogeneous, consisting of young people and students: young people are a separate group with respect to Internet use, which is coherent with Ahn (2011) who finds that youth are unique subpopulations. In Figure 1 the cluster occupies the lower part of the map and is outstanding in the second principal component characterized by an intensive use of the mobile telephone and participation in social networks, downloading games and listening to music. The higher the level of studies, the greater the use of Internet. Those who study and work, due to their good economic situation, as well as recreational and social uses, also shop on Internet and use e-banking. In this cluster, there is no gender gap, nor any difference between rural and urban environments: young people use Internet and technologies regardless of gender and environment.

Summing up, we can highlight the following: a) unemployed people up to 34 have a high level of Internet usage, which is explained because in Spain this group is highly qualified professionally; b) the gender digital gap does not affect the oldest and youngest groups, but only those of intermediate age; c) with respect to the type of work, one of the groups that

shows most from the digital divide is the homemaker, who in Spain are mainly women; d) we observe a low use of interest in rural environments, but mainly in older groups of the population; e) Internet use among young people is intensive but limited, especially in the case of NEETs, where it is exclusively for leisure. Efforts and actions for digital training to eradicate the digital divide should be directed at these groups.

4. CONCLUSIONS

This paper explores the use of Internet, including social networks, e-government services and e-commerce from the perspective of the digital divide. The data has been collected from a survey (N=2,304), carried out in Spain and including 63 questions about Internet use. Age is the factor that most explains the level of Internet use but the relationship between age and advanced Internet uses is not linear because young people use the Internet a lot but hardly take any advantage of the advanced functions, such as shopping and e-government. Education level is the second most important factor in explaining the digital divide. The third most important factor is income level, especially when dealing with activities which involved money, such as Internet shopping and e-banking.

A principal component analysis shows that two components explain most of the variability of the sample. The first component can be interpreted as “comprehensive Internet use”. The second distinguishes between “online shopping and professional use” versus “recreational and social use of the Internet”. A cluster analysis has allowed us to identify the groups most affected by the digital divide, and even the digitally excluded. They include people over 55, those with the lowest incomes, and those without studies or only primary education, whatever their genders. All these groups need specific training to avoid their digital exclusion, homemakers are also greatly affected by the digital divide and, as they are mostly women, this explains to a great extent the gender gap. The level of Internet use of the unemployed on average is similar to that of full-time workers. This situation is a characteristic of unemployment caused by the crisis in Spain, where many unemployed people have high educational levels and good qualifications. Although women use Internet less than men, there is no gender gap between the young and those with good incomes and educational levels. Among older people there are no important differences with respect to gender: both sexes using Internet very little. A lower use of Internet is observed in rural environments but not among the young: the gap is connected with the aging of the population. In young people, Internet use is intense but limited; a special case being the NEETs who are great users of Internet but only in recreational and social aspects, thus in a certain way appears the digital divide.

ACKNOWLEDGMENTS

This work was supported by grants ECO2010-20228 and ECO2013-45568-R of the Spanish Ministry of Education and the European Regional Development Fund and by grant Ref. S-14/2 of the Government of Aragon.

REFERENCES

- Ahn, J. (2011). Digital divides and social network sites: Which students participate in social media?. *Journal of Educational Computing Research*, 45(2), 147-163.
- Attewell, P. (2001). The first and second digital divides. *Sociology of Education*, 74, 252–259.
- Bandura, A. (1986). *Social Foundations of Thought and Action*. Prentice-Hall, Englewood Cliffs, NJ.
- Bélanger, F., & Carter, L. (2009). The impact of the digital divide on e-government use. *Communications of the ACM*, 52(4), 132-135.
- Chinn, M. D., & Fairlie, R. W. (2006). The determinants of the global digital divide: a cross-country analysis of computer and Internet penetration. *Oxford Economic Papers*.
- Cooper, J. (2006). The digital divide: The special case of gender. *Journal of Computer Assisted Learning*, 22(5), 320-334.
- Czaja, S. J., & Lee, C. C. (2007). The impact of aging on access to technology. *Universal Access in the Information Society*, 5(4), 341-349.
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly*, 319-340.
- European Commission (2014). *Digital Agenda Scoreboard*, Retrieved at http://ec.europa.eu/information_society/digital-agenda/scoreboard/index_en.htm.
- Felter, M (1985). Sex differences on the California state-wide assessment of computer literacy. *Sex Roles*, 13, 182.-192
- Forman, C., Goldfarb, A., & Greenstein, S. (2005). Geographic location and the diffusion of Internet technology. *Electronic Commerce Research and Applications*, 4(1), 1-13.
- Fuchs, C. (2008). The role of income inequality in a multivariate cross-national analysis of the digital divide. *Social Science Computer Review*, 27(1), 41-58
- Gibbs, J., Kraemer, K. L., & Dedrick, J. (2003). Environment and policy factors shaping global e-commerce diffusion: A cross-country comparison. *The information society*, 19(1), 5-18.
- Helbig, N., Ramón Gil-García, J., & Ferro, E. (2009). Understanding the complexity of electronic government: Implications from the digital divide literature. *Government Information Quarterly*, 26(1), 89-97.

- Hilbert, M. (2011a). The end justifies the definition: The manifold outlooks on the digital divide and their practical usefulness for policy-making. *Telecommunications Policy*, 35(8), 715-736.
- Hilbert, M. (2011b). Digital gender divide or technologically empowered women in developing countries? A typical case of lies, damned lies, and statistics. *Women's Studies International Forum*, 34(6), 479-489.
- Hoffman, D. L., Novak, T. P., & Schlosser, A. (2000). The evolution of the digital divide: How gaps in Internet access may impact electronic commerce. *Journal of Computer-Mediated Communication*, 5(3), 0-0.
- Hunt, N.P. and Bohlin, R.M. (1993). Teacher education and students attitudes toward using computers. *Journal of research on Computing in Education*, 25(4), pp. 487-497.
- Loges, W. E., & Jung, J. Y. (2001). Exploring the digital divide. Internet connectedness and age. *Communication Research*, 28(4), 536-562.
- Klecun, E. (2008). Bringing lost sheep into the fold: questioning the discourse of the digital divide. *Information Technology & People*, 21(3), 267-282.
- Monsuwé, T. P., Dellaert, B. G., & De Ruyter, K. (2004). What drives consumers to shop online? A literature review. *International journal of service industry management*, 15(1), 102-121.
- Porter, C.E., & Donthu, N. (2006). Using the technology acceptance model to explain how attitudes determine Internet usage: the role of perceived access barriers and demographics. *Journal of Business Research*, 59, 999-1007.
- Rice, R. E., & Katz, J. E. (2003). Comparing Internet and mobile phone usage: digital divides of usage, adoption, and dropouts. *Telecommunications Policy*, 27(8), 597-623.
- Riggins, F. J., & Dewan, S. (2005). The digital divide: Current and future research directions. *Journal of the Association for information systems*, 6(12), 13.
- Rogers E.M. (2003), *Diffusion of Innovations*, (5th ed.) Free Press, New York
- Sipior, J. C., Ward, B. T., & Connolly, R. (2011). The digital divide and t-government in the United States: Using the technology acceptance model to understand usage. *European Journal of Information Systems*, 20(3), 308-328.
- Taipale, S. (2013). The use of e-government services and the Internet: The role of socio-demographic, economic and geographical predictors. *Telecommunications Policy*, 37(4), 413-422.
- Van Deursen, A. J., & Van Dijk, J. A. (2014). The digital divide shifts to differences in usage. *New Media & Society*, 16(3), 507-526.
- Van Dijk, J. A. (2009). *One Europe, digitally divided*. *Routledge handbook of Internet politics*, 288-305.

- Van Dijk, J. A. (2012). *The Evolution of the Digital Divide - The Digital Divide Turns to Inequality of Skills and Usage*. In: Digital Enlightenment Yearbook 2012. IOS Press, 57 - 78.
- Van Dijk, J. A. (2013). A theory of the digital divide. In Ragnedda and Muschert (Eds), *The Digital Divide*, Routledge, 29-52.
- Valadez, J. R., & Durán, R. P. (2007). Redefining the digital divide: Beyond access to computers and the Internet. *The High School Journal*, 90(3), 31-44.
- Vicente, M. R., & López, A. J. (2011). Assessing the regional digital divide across the European Union-27. *Telecommunications Policy*, 35(3), 220-237.
- Warren, M. (2007). The digital vicious cycle: Links between social disadvantage and digital exclusion in rural areas. *Telecommunications Policy*, 31(6), 374-388.
- Wei, K.K., Teo, H.H., Chan, H.C. and Tan, B.C.Y. (2011). Conceptualizing and Testing a Social Cognitive Model of the Digital Divide. *Information System Research*, 22, 11, pp. 170-187.
- Zhang, X. (2013). Income disparity and digital divide: The Internet Consumption Model and cross-country empirical research. *Telecommunications Policy*, 37(6), 515-529.

Internet activity	Variables
Generic use of Internet (25 items)	<i>To read newspapers</i> <i>To read emails</i> <i>To get information about courses</i> <i>To do online courses</i> <i>To make medical appointments</i> <i>To carry out e-banking activities</i> <i>To get information about jobs</i> <i>To get information about health</i> <i>To get information about culture</i> <i>To get information about shows</i> <i>To get information about society and economy</i> <i>To get information about history</i> <i>To get information about the weather</i> <i>To listen to music</i> <i>To download films and videos</i> <i>To listen to the radio or TV</i> <i>To read blogs</i> <i>To write in blogs</i> <i>To get information about holidays</i> <i>To get information about products</i> <i>To take shopping decisions</i> <i>To download software</i> <i>To download P2P file</i> <i>To make online call or video call</i> <i>To play or download games</i>
Social-network use (12 items)	<i>To upload photos or videos on social networks</i> <i>To set the level of privacy in social networks</i> <i>To send private messages</i> <i>To create events and groups online</i> <i>To use the mobile to access social-networks</i> <i>To play and to use applications</i> <i>To post comments</i> <i>To contact friends</i> <i>To use social networks about hobbies and interests</i> <i>For professional purposes</i> <i>To find recommendations about products</i> <i>Dating on social networks</i>
Use of e-government (6 items)	<i>To search for information in administrations</i> <i>To get forms from administrations</i> <i>To send forms to administrations</i> <i>To make payments to administrations</i> <i>To use digital signature in dealings with public administrations</i> <i>To use Electronic Identity Card in dealings with public administrations</i>
E-commerce use (14 items)	<i>To buy clothing, accessories and sports equipment</i> <i>To buy transport tickets</i> <i>To buy travel accommodation</i> <i>To buy electronic equipment</i> <i>To buy tickets for shows</i> <i>To buy computer equipment</i> <i>To buy books</i> <i>To buy products for home and car</i> <i>To buy food and cleaning products</i> <i>To buy films and music</i> <i>To buy pharmaceutical products</i> <i>To buy computer games</i> <i>To buy insurance and financial products</i> <i>To buy lottery</i>
Use of Internet on mobile (6 items)	<i>Use the mobile to read emails</i> <i>Use the mobile to browse Internet</i> <i>Use the mobile to send or upload photos or videos</i> <i>Use the mobile to download or watch TV or videos</i> <i>Use the mobile for e-banking</i> <i>Use the mobile for other activities</i>

Table 1. Internet activity variables.