



Instituto  
Complutense  
de Análisis  
Económico

# Adoption of e-commerce by individuals and digital divide: Evidence from Spain

Ángel Valarezo

Instituto Complutense de Análisis Económico (ICAE),  
Universidad Complutense de Madrid (UCM)

Rafael López

ICAE, UCM

Teodosio Pérez-Amaral

ICAE, UCM

## Abstract

E-commerce penetration rates are distant among those groups of individuals with the lowest and the highest levels of online shopping adoption. This is an indicator of digital divide, having negative effects in terms of untapped opportunities for people, companies and the whole economy. Key socioeconomic and demographic determinants of adoption of e-commerce are explored, analyzing a dataset of 174,776 observations for the period 2008-2017 in Spain. The empirical analysis is based on a standard neoclassical utility maximization framework. Linear probability model, logistic regression, and Heckman's sample selection correction model have been used. The results suggest that e-commerce adoption is positively related with being male, having higher levels of education, income and digital skills, being Spanish, and being employed; while being female, older and belonging to a household of two or more members have negative effects. An interaction between digital skills and age has been introduced in the model, where high digital skills seem to have a positive influence, partly counteracting the lower odds for some age groups. Policy recommendations related to demand and supply measures are suggested to foster the adoption of e-commerce.

**Keywords** e-commerce, digital divide, linear probability model, logistic regression, Heckman's sample selection correction, polychoric correlation, digital skills, time and regional dummies, pool data, utility maximization framework.

**JEL Classification** C25, D11, O33.

**Working Paper nº 1919**  
March, 2019



UNIVERSIDAD  
COMPLUTENSE  
MADRID

ISSN: 2341-2356

WEB DE LA COLECCIÓN: <http://www.ucm.es/fundamentos-analisis-economico2/documentos-de-trabajo-del-icae> Working papers are in draft form and are distributed for discussion. It may not be reproduced without permission of the author/s.

## **Adoption of e-commerce by individuals and digital divide: Evidence from Spain**

**02/2019**

Ángel Valarezo

Instituto Complutense de Análisis Económico (ICAE),  
Universidad Complutense de Madrid (UCM)

Rafael López  
ICAE, UCM

Teodosio Pérez-Amaral  
ICAE, UCM

### **ABSTRACT**

E-commerce penetration rates are distant among those groups of individuals with the lowest and the highest levels of online shopping adoption. This is an indicator of digital divide, having negative effects in terms of untapped opportunities for people, companies and the whole economy. Key socioeconomic and demographic determinants of adoption of e-commerce are explored, analyzing a dataset of 174,776 observations for the period 2008-2017 in Spain. The empirical analysis is based on a standard neoclassical utility maximization framework. Linear probability model, logistic regression, and Heckman's sample selection correction model have been used. The results suggest that e-commerce adoption is positively related with being male, having higher levels of education, income and digital skills, being Spanish, and being employed; while being female, older and belonging to a household of two or more members have negative effects. An interaction between digital skills and age has been introduced in the model, where high digital skills seem to have a positive influence, partly counteracting the lower odds for some age groups. Policy recommendations related to demand and supply measures are suggested to foster the adoption of e-commerce.

**Keywords:** e-commerce, digital divide, linear probability model, logistic regression, Heckman's sample selection correction, polychoric correlation, digital skills, time and regional dummies, pool data, utility maximization framework.

Acknowledgements: Paper presented at the 29<sup>th</sup> ITS European Conference, Trento-Italy, 2018.

This research is funded by the Autonomous Community of Madrid - Spain. Project: Finance, Innovation and Strategies: Economic Analysis of Business Productivity and its Determinants (PRODECON-CM HM S2015/HUM-3491), 2016-2018.

## 1. Introduction

A standard definition of e-commerce (Eurostat, 2018a) and INE (2018b) is the placing of orders of goods or services via the Internet, excluding orders via manually typed e-mails or text messages. Electronic payment or delivery are not required for an e-commerce transaction. Only purchases made for personal reasons are considered.

According to the European Commission's Digital Economy and Society Index (DESI) of 2018 (European Commission, 2018a), Spain ranks 10<sup>th</sup> among the 28 EU Member States, belonging to the group of medium performing countries. This index tracks the evolution of EU countries in digital competitiveness, assessing their digitization around five dimensions or policy areas: Connectivity<sup>1</sup>, Human Capital<sup>2</sup> (or digital skills), Use of Internet Services by citizens<sup>3</sup>, Integration of Digital Technology<sup>4</sup> and Digital Public Services<sup>5</sup>. Spain has improved in almost all domains, except for Human Capital. Although Spanish citizens perform a wide variety of Internet activities, Spain fell back from rank 16 to rank 17 in the Internet use section. Online shopping is below the EU-28<sup>6</sup> average and several indicators suggest it is due to a weak demand side, especially on the private user side.

The modeling of e-commerce adoption based on disaggregate measures of socio-economic factors on an individual level has received scarce attention. Much of the research about e-commerce adoption has been done from business and technological standpoints, focusing on the process of how the customer decides to shift from offline to online shopping. The aim is usually to identify the main characteristics and determinants, which can be targeted by marketing strategies or by technical implementations. Surveys and theoretical models have been specially designed or adapted to capture behavioral determinants and moderator effects on the decision to buy online.

The effects of socio-economic dimensions on the adoption of e-commerce is closely related to the concept of "digital divide", the unequal access of individuals to digital technology (Cerno & Pérez-Amaral, 2006b). As Srinuan & Bohlin (2011) and Demoussis & Giannakopoulos (2006), among others, pointed

---

<sup>1</sup> Fixed broadband, mobile broadband and prices.

<sup>2</sup> Internet use, basic and advanced digital skills.

<sup>3</sup> Citizens' use of content, communication and online transactions.

<sup>4</sup> Business digitisation and e-commerce

<sup>5</sup> eGovernment and eHealth.

<sup>6</sup> EU-28: Member States of the European Union.

out, digital divide is a multifaceted phenomenon that does not depend only on technological determinism. Socio-economic, institutional and psychological factors have been included in these studies.

Much of the purely online supply is composed by very high value-added products and services, so not using e-commerce can end up being a source of social exclusion. Helsper & van Deursen (2015) found that the Internet remains more beneficial for those with a higher social status, meaning that in terms of social implications existing offline inequalities could potentially be exacerbated.

The aim of this paper is to identify the main socio-economic and demographic factors that influence an individual's adoption of online shopping. This will be the basis for discussing public policies and private strategies to promote the use of e-commerce and bridging the digital divide by specifically targeting those groups with the lowest penetration rates of e-commerce and other ICTs services. In order to account for the determinants on the individual decision of becoming an online buyer for private use, descriptive analysis and quantitative models are formulated.

The pool dataset used in this paper has a total of 174,776 observations and was constructed using the annual representative survey on ICT usage by households and individuals (years 2008 to 2017) carried out by the National Institute of Statistics of Spain (INE).

The rest of the paper is organized as follows. Section 2 presents a literature review. Section 3 describes the evolution and current situation of adoption of e-commerce in Spain. Section 4 presents the data to be used and its main characteristics. In Section 5 the estimations of the linear probability, logistic regression and Heckman selection models are presented, showing the effects of the selected explanatory variables on the probability of adopting e-commerce. Section 6 contains conclusions, and policy recommendations.

## **2. Literature review**

### **2.1 Models of ICT Adoption**

Adoption of technology, online consumer behavior, and the specific decision of adopting e-commerce have been studied based on several frameworks. By far, the most used approaches are applications, adaptations and/or unifications of models and theories of individual acceptance and intention. Among the most sounded theoretical backgrounds are: Innovation Diffusion Theory (IDT), Rogers (2003), Theory of Reasoned Action (TRA), Ajzen & Fishbein (1977, 1980), Theory of Planned Behavior (TPB), Ajzen (1991), Technology Acceptance Model, TAM (Davis, et al., 1989; F. D. Davis & Warshaw, 1989; Fred D. Davis, 1989, 1993) , Theoretical Extension of the Technology Acceptance Model (TAM2), Venkatesh & Davis (2000), Decomposed Theory of Planned Behavior (DTPB), Taylor & Todd (1995), Expectation-

Confirmation Theory (ECT), Oliver (1980), and the Unified Theory of Acceptance and Use of Technology (UTAUT), Venkatesh, et al. (2003).

Most of the rich body of theory and applied research about ICT's diffusion, are based on IDT, TRA, TAM and TPB, which includes variables that affect an individual's motivation to accept a new technology and helps to explain the decision-making process of doing so. If the goal is to go further than the initial acceptance, meaning the study of post-purchase behavior, ECT is widely used to analyze the individual to continue using a technology or performing transactions. The contributions of the above-mentioned literature are considered, as precedents of this study.

## **2.2 Digital divide**

Not long ago the main indicators of digital divide were related with the "haves" and "haves nots", however NTIA<sup>7</sup> (1995) stated the necessity to go beyond the traditional focus on telephone penetration and infrastructure access. At that time, the importance of studying computer and modem penetration according to socioeconomic, demographic and geographic variables was clear. A few years later several works (DiMaggio & Hargittai, 2001; Hargittai, 2001; Norris, 2001; and OECD, 2001; among others), pointed out that the research community should further expand their focus to the full range of digital inequality, not only accounting for equipment and Internet access, but also for digital skills, technological evolution and the widening scope of Internet usage, without avoiding the complexity of characterizing individuals at different levels of analysis.

The shift from studying the inequalities of access to the differences in the extent of use brought new conceptual dimensions. Hargittai (2001), among others, suggested that there is a "second-level of digital divide" where the people's online skills to achieve Internet related tasks (as finding information online) play a key role, instead of just considering whether someone is or is not an Internet user. A "third-level of digital divide" has been proposed, as Helsper and van Deursen (2015) accounted for the link between inequality and the achievement of tangible outcomes of Internet use.

## **2.3 Internet use and e-commerce adoption**

The determinants of Internet use and online shopping in Spain have been empirically analyzed before, based on the INE's survey on ICT equipment and usage by households and individuals. The studies have mainly used cross sectional data and pools of cross sections. One of the early studies in Spain about the

---

<sup>7</sup> National Telecommunications and Information Administration

demand of Internet access and use was carried out by Cerno and Pérez-Amaral (2006); two models were estimated using data from 2003, one for broadband access at home and the other for Internet use intensity. Lera-López, et al. (2011) explored the impact of socio-economic, demographic and regional factors to explain not only Internet use, but also the frequency of use by individuals. Cerno & Pérez-Amaral (2009) analyzed the survey for 2003, characterizing the e-consumer profile, for both the number of purchases and the expenditure. Garín Muñoz and Pérez-Amaral (2011) used data from 2007 to model the adoption and use of e-commerce identifying the effects of key socio-demographic variables, attitudes and beliefs. Covering the period 2004-2009, Pérez-Hernández and Sánchez-Mangas (2011) focused on having Internet at home and its implications for e-commerce adoption. A working paper of Alonso and Arellano (2015) explored the heterogeneity and diffusion of the digital economy in Spain from 2003 to 2014, estimating the effects of innovation and adoption of Internet, e-commerce and e-banking. During the 2008-2014 period, Correa, et al. (2015) also constructed a multiple-stage model to identify the importance of the individual determinants on the online purchasing decision. The dataset of 2016 has been studied by Garín-Muñoz, et al. (2019), to model the individual adoption of e-commerce, e-banking and e-government; and Valarezo, et al. (2018), to explore the determinants of the individual decision to perform cross-border e-commerce for private ends.

The above cited empirical studies have mostly specified binary response models. Demographic and socioeconomic characteristics of household and individuals have been used and interpreted. Although most of the works that have dealt with INE's survey on ICT account for the explanatory factors, they don't analyze specific sub categories. Besides controlling for aggregate time effects, the current paper analyzes an individual's economic incentives and determinants of e-commerce adoption, categorizing the explanatory variables as follows: sociodemographic (which includes, gender, age, population size, number of household members, nationality<sup>8</sup>), skills (attained levels of education, levels of digital competence), economic (employment situation, level of household income), and geography and time (region of residence and year).

### **3. Adoption of e-commerce in Spain**

In 2017, 82.7% of Spanish households had broadband connection and 85% of the Spanish population (between 16 and 74 years<sup>9</sup>) used Internet at least once in the last twelve months, although only 69.0% did

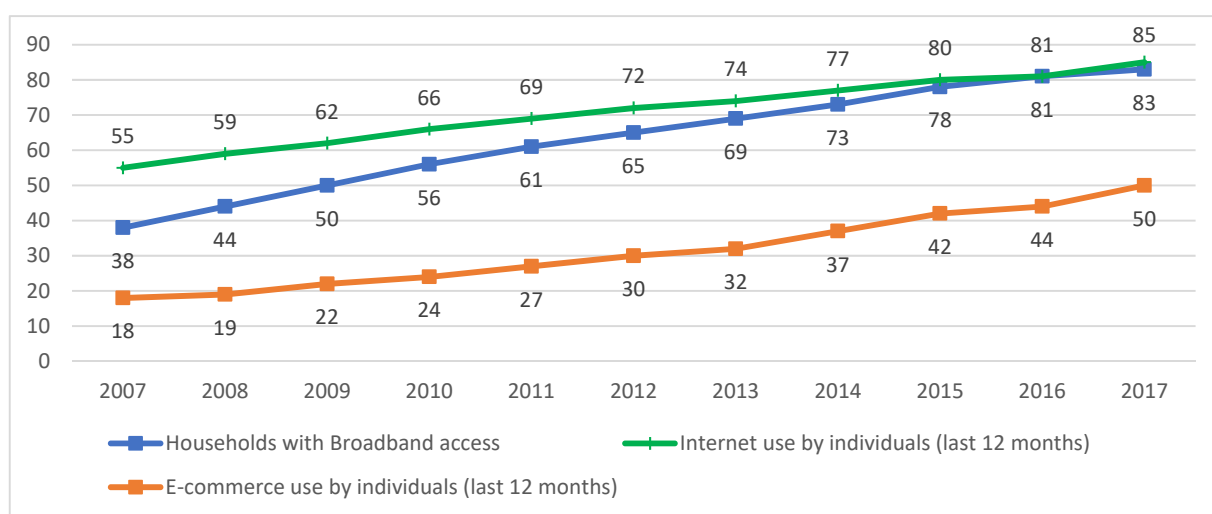
---

<sup>8</sup> Having or not Spanish nationality.

<sup>9</sup> For comparability with the data published by EUROSTAT, INES's ICT survey results are referred to households with at least one person between 16 to 74 years old.

so on a daily basis. While e-commerce penetration rate reached 40%, namely four out of 10 people bought online for private use at least once in the last 3 months. It goes up to 49.9% if the considered period is extended to 12 months (Figure 1). Spain still lags behind the EU average of e-commerce adoption, not to mention when the comparison is made with respect to the best performers which include: Finland, Germany, Iceland, Netherlands, Luxembourg, Denmark and United Kingdom (UK), all above 70%, considering those who bought online in the last 12 months.

*Figure 1*  
Households with private broadband access at home, penetration rates of internet use, and e-commerce. As percentage of people aged 16 to 74 (2007–2017)



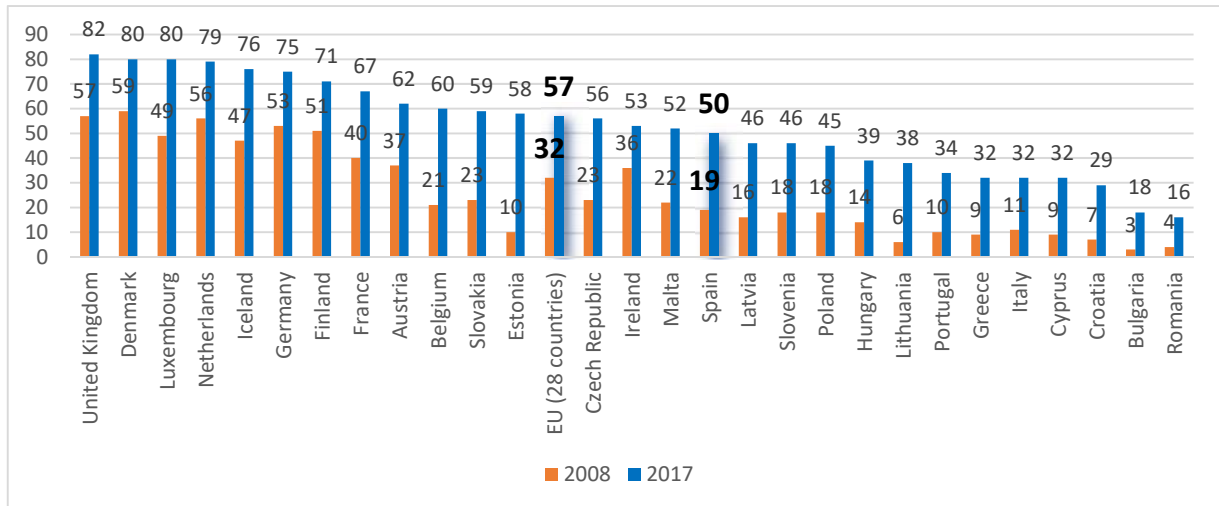
Relevant indicators on Europe’s digital performance are summarized by DESI, developed within the Digital Single Market strategy of the European Commission (2018).

According to the country specific profile report (European Commission, 2018a) Spain belongs to the medium performance group of countries<sup>10</sup> and ranks 10<sup>th</sup> among the 28 EU Member States in DESI 2018, improving two positions with respect to 2017<sup>11</sup>. However, this position represents an average that hides important differences when comparing the most and least advanced EU countries. This is the case of the Spanish penetration rate of online shopping by individuals for private use, ranking 17<sup>th</sup>, 4 places behind the EU average (Figure 2).

<sup>10</sup> DESI’s medium-performing countries are Spain, Austria, Lithuania, Germany, Slovenia, Portugal, Czech Republic, France and Latvia.

<sup>11</sup> Small increases of penetration rates are observed in 2018. Households with broadband connection reached up to 86.1%, matching Internet use. E-commerce adoption has also increased up to 43.5%.

Figure 2  
Penetration rates of e-commerce in the EU-28 as percentage of people aged 16 to 74 (2008 and 2017)



Aggregate data seems to suggest that the penetration rates of e-commerce, as other Internet services, in Spain have had a healthy evolution throughout the last few years, which might imply that the digital divide is narrowing. However, when we consider specifically the evolution of penetration rates by different variables and their categories, several divides are widening when the differences between groups with the lowest and largest penetration rates in 2008 and 2017 are measured in absolute values<sup>12</sup>.

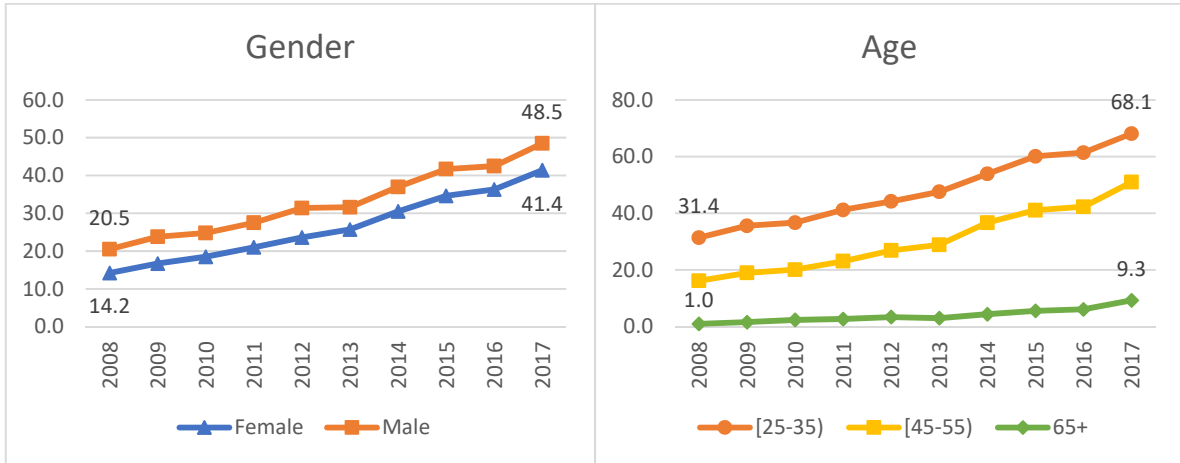
The distance between male and female has increased slightly in 2017 with respect to 2008. It is more evident when we compare the age groups between the highest and the lower levels of adoption of online shopping. The group aged 25 to 35 years old goes from 31.4% in 2008 to 68.1% in 2017, while those aged over 65 with 1% of penetration rate in 2008 reached up to 9.3% in 2017 (Figure 3).

Figure 4 shows the evolution of individual penetration rates of e-commerce for the main categories of Education and Digital Skills. The trend for both dimensions is similar, the distances between categories are wider in 2017 than in 2008. The group with primary or less education goes from 1.4% in 2008 to 6.5% in 2017, while those with a master or PhD went from 20.5% to 83.5%. Individuals with low digital skills went from 9.8% in 2008 to 12.3% in 2017, while those with a very high level went from 69.2% in 2008 to 89.8% in 2017.

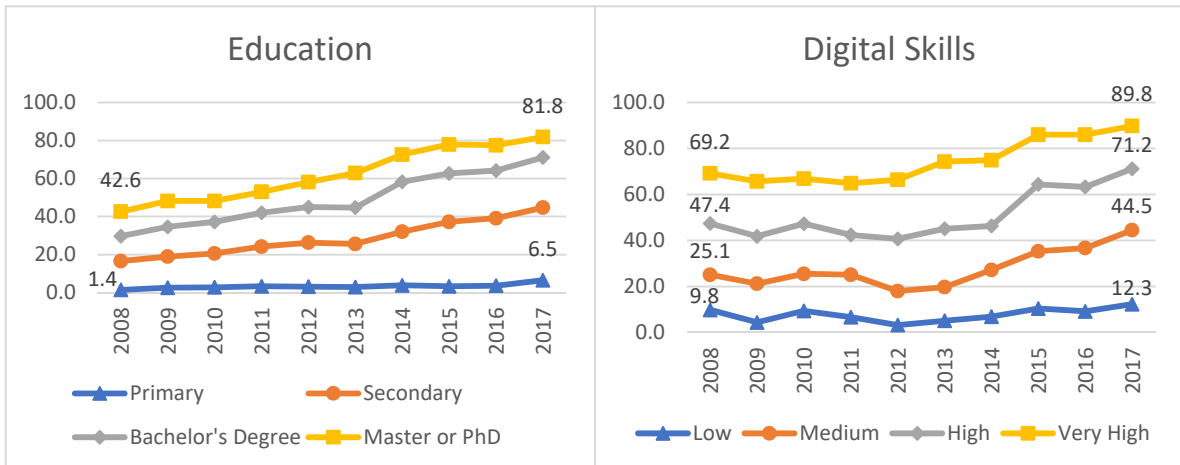
<sup>12</sup> All comparisons assessed in absolute values show greater distances. On the other hand, relative distances (the difference between the lowest and the highest penetration rate, as a percentage of the highest penetration rate) show a rapid growth in e-commerce adoption, but slow narrowing of the digital divide.



*Figure 3  
Evolution of individual penetration rates of e-commerce in Spain (2008-2017): Gender and Age*

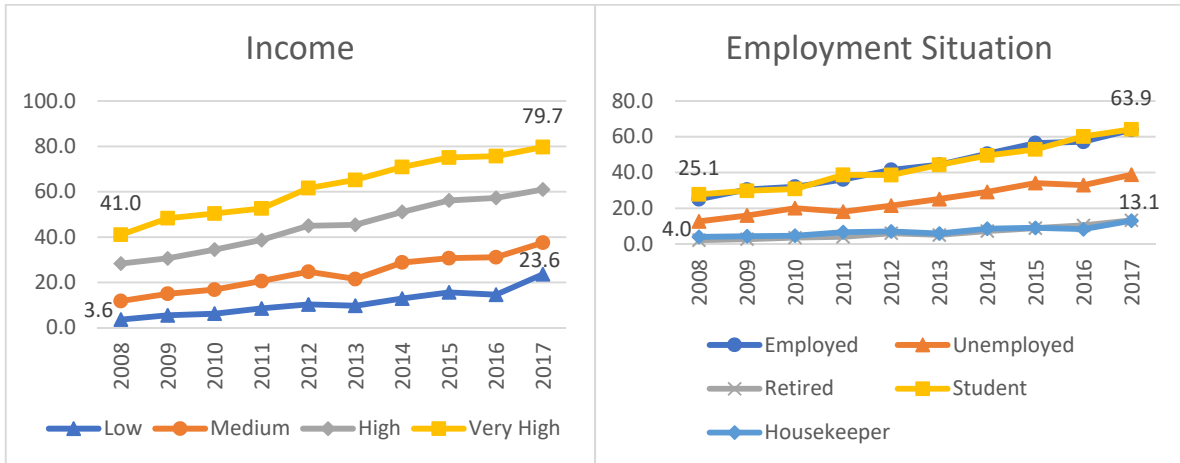


*Figure 4  
Evolution of individual penetration rates of e-commerce in Spain (2008-2017):  
by Education and Digital Skills*



Income and Employment situation also reflect a positive evolution in terms of growth of penetration rates, but not in terms of bridging the digital divide. 3.6% of individuals with low income bought online in 2008, reaching 41% in 2017; and the group of people with very high income went from 23.6% in 2008 to 79.7% in 2017. Unemployed and retired had a penetration rate of 4% in 2008 and 25.1% in 2017, whilst employed went from 13.1% in 2008 to 63.9% in 2017 (Figure 5).

*Figure 5*  
*Evolution of individual penetration rates of e-commerce in Spain (2008-2017):*  
*Income and Employment Situation*



An aggregate perspective suggests a positive evolution of e-commerce adoption in Spain. But, contrasting among groups of categories for different factors reveals that growth in adoption is not yet reducing the distance in absolute values between individuals with lowest penetration rates compared to those with the highest, on the contrary it is widening. And, as expected, most of the distances are slowly narrowing when relative values are compared (*Table 1*).

#### 4. The ICT-H survey. Pooled data

The dataset of 174,776 observations comes from pooling the annual cross sections from 2008 to 2017 of the survey on Equipment and Use of Information Technologies in Households (ICT-H), carried out annually by the Spanish National Statistical Institute (INE, 2018). The survey is a panel-based study, that collects information about ICTs equipment and usage in Spanish households, following the methodological recommendations of Eurostat.

As information and communication technologies change continuously, the questionnaire<sup>13</sup> undergoes slight changes every year based on Eurostat's annual model questionnaires. The survey gathers information about the following subjects: access to and use of ICTs by individuals and/or in households, use of the Internet by individuals and/or in households, ICT security and trust, ICT competence and skills, barriers to use of ICT and the Internet, perceived effects to ICT usage on individual and/or on households,

<sup>13</sup> From 2008 to 2017 used questionnaires have included between 50 and 70 questions, many of them with several possible answers, meaning more than 200 variables, depending on the composition of the questionnaires for each year.

use of e-government and ubiquitous connectivity. The data allows to classify households by region of residence, geographical location, population density, type of household and by household monthly net income. Individuals can also be classified by gender, country of birth, country of citizenship, educational level, occupation, employment situation, age and legal or de facto marital status (Eurostat, 2018b).

Since this paper examines the socio-economic and demographic factors that influence the odds of becoming an online shopper, only some of the questions from the questionnaire are used for the statistical description and for the specification and estimation of the models. Table 1 shows the penetration rates as a percentage of individuals aged 16 and above.

Overall penetration rates of e-commerce in Spain has more than doubled in a decade, going from 17.3% in 2008, to 44.9%. As expected, the rate is higher when calculated as a percentage of individuals aged 16 to 74, who used the Internet within the last three months, or even within the last year.

Differences of adoption are observable within specific demographic factors (Gender, Age, Habitat and number of Household Members). Attributes as being male, being between 16 to 45 years old, living in urban areas of more than 100,000 inhabitants and belonging to households with three or more members, are associated with higher penetration rates of e-commerce.

The higher the level of education, digital skills and income, the higher the penetration rates of online shopping. These are the variables where the gaps between groups of each category are more evident. In 2017 only 6.5% of people with primary or less education have bought online, while 81.8% did so in the case of those who hold a Master's or a PhD degree. In the case of Digital Skills, the penetration is 89.8% for very high level, compared with 12.3% for low level. For low, medium, high and very high income, the penetrations rates were 23.6%, 37.5%, 61.0% and 79.7%, respectively.

Comparing regions, Madrid, Navarra, Baleares, Cantabria and Cataluña are the top five autonomous communities with the highest penetration rates. País Vasco, that was 4<sup>th</sup> in 2008, dropped to the 9<sup>th</sup> position in 2017; and the 4<sup>th</sup> place now is occupied by Cantabria, that was 7<sup>th</sup> in 2008. It is remarkable how the autonomous city of Ceuta (in northern Africa) has moved from 17<sup>th</sup> to 7<sup>th</sup>, going from 12.6% (in 2008) up to 48.6% (in 2017) of e-commerce adoption, above the national average.

Differences within categories suggest that there are significant divides in e-commerce adoption. Penetration rates diverge across educational, digital skills and income levels. The same happens with gender, age, nationality, household members, habitat and regions.

*Table 1*  
*Individual penetration rates of e-commerce in Spain (2008 and 2017)*

Characteristics	Category	Relative distance 2018	Penetration rates		Relative distance 2017
			2008	2017	
GENDER	Female	30.7%	14.2	41.4	14.6%
	Male		20.5	48.5	
AGE	<25	96.8%	25.1	64.5	↓ 86.3%
	[25,35)		31.4	68.1	
	[35,45)		21.3	62.3	
	[45,55)		16.2	51.0	
	[55,65)		7.2	31.8	
	65+		1.0	9.3	
EDUCATION	Primary	96.7%	1.4	6.5	↓ 92.1%
	Secondary		16.6	44.7	
	Bachelor's Degree		29.7	71.1	
	Master or PhD		42.6	81.8	
DIGITAL SKILLS	Low	85.8%	9.8	12.3	↑ 86,3%
	Medium		25.1	44.5	
	High		47.4	71.2	
	Very High		69.2	89.8	
HABITAT	<20,000	41.4%	13.0	38.7	↓ 24.4%
	20,000-100,000		16.3	45.4	
	100,000-500,000		18.2	42.2	
	>500,000		22.2	51.2	
HOUSEHOLD MEMBERS	1	42.3%	12.4	31.1	↑ 47.0%
	2		14.6	32.5	
	3		18.1	51.2	
	4		21.5	58.7	
	5+		14.9	43.0	
NATIONALITY	Foreigner	18.1%	14.5	41.0	↓ 9.4%
	Spanish		17.7	45.3	
EMPLOYMENT SITUATION	Employed	85.6%	25.1	63.9	↓ 79.6%
	Unemployed		12.7	38.8	
	Retired		2.2	13.3	
	Student		27.8	64.2	
	Housekeeper		4.0	13.1	
	Other		11.2	27.6	
INCOME	Low	91.2%	3.6	23.6	↓ 70.4%
	Medium		11.8	37.5	
	High		28.3	61.0	
	Very High		41.0	79.7	

*Table 1. (Continued)*  
*Individual penetration rates of e-commerce in Spain (2008 and 2017)*

Characteristics	Category	Relative distance 2018	Penetration rates		Relative distance 2017
			2008	2017	
AUTONOMOUS COMMUNITY	Andalucía	49.8%	13.1	41.0	↓33.2%
	Aragón		18.7	49.1	
	Asturias		16.1	44.3	
	Baleares		24.3	52.9	
	Canarias		12.2	37.6	
	Cantabria		18.6	49.8	
	Castilla y León		14.6	39.4	
	Castilla La Mancha		12.6	41.7	
	Cataluña		21.2	49.1	
	Valencia		15.3	43.7	
	Extremadura		14.4	37.1	
	Galicia		13.9	36.0	
	Madrid		23.8	53.9	
	Murcia		14.2	39.2	
	Navarra		20.1	53.2	
	País Vasco		21.1	46.7	
La Rioja	18.3	44.0			
AUTONOMOUS CITIES	Ceuta	29.6%	12.6	48.6	↓0.8%
	Melilla		17.9	47.8	
TOTAL			17.3	44.9	
NOTES. Weighted data: The sampling weight correspond to the number of subjects in the population represented by each observation. Penetration rates as a percentage of individuals aged 16 and up. An individual is considered as an online buyer in any of the categories if he or she has carried out an online purchase for private use within the last 12 months. Relative distances are the difference between the lowest and the highest penetration rates, as a percentage of the highest penetration rate.					

## 5 Theoretical and Empirical Models

Adopting e-commerce is an individual economic decision that will depend mostly on economic conditions as well as other sociodemographic and geographic variables. Like most economic decisions related to individual consumption it depends on income, cost and cost of substitutes/complements, and size of the network (supply and customer base). A discrete choice framework is adopted, derived from utility theory, following an approach similar to Varian (2002) for modeling the demand for bandwidth, and Demoussis & Giannakopoulos (2006) used to model the individual decision to adopt the Internet.



	Habitat:	4 groups
	Household Members:	5 groups
	Nationality:	2 groups: 1 if Spanish, 0 if Foreign
<b>Individual skills</b>	Education:	4 levels of study
	Digital Skills:	4 levels
<b>Economic</b>	Employment Situation:	6 groups
	Income:	4 groups, monthly net income
<b>Geographic and Time</b>	Yearly Dummies:	1 for each year
	Regional Dummies:	17 Autonomous Communities

Since Gender, Nationality, Yearly Dummies and Regional Dummies are binary variables, no categorization was needed. The rest of the variables were categorized as shown in *Table 1*.

Socio Demographic variables can be interpreted as follows: Gender allows for a gap between males and females; Age lowers or increases the cost, depending on the considered age ranges; Habitat increases or reduces the benefits, conditional on the size of the populations to which the individual belongs; Household Members is related with the benefits and also with the costs (that may be due to a learning effect); and, Nationality might signal effects dependent on costs (e.g., different language) and benefits (e.g., access to a wider range of products and services). Variables such as Education and Digital Skills<sup>14</sup> are expected to diminish the costs of using Internet services; Economic variables, Income and Employment Situation, are supposed to increase the benefits. And, Yearly and Regional Dummies point out benefits and costs across regions, and also those associated with changes (e.g., entry of new suppliers, legislation and institutional measures to foster e-commerce) that can occur over the years associated to belonging to specific regions.

A Linear Probability Model (LPM), depicted in *Table 2 (1)* and *Figure 6*, has been estimated in the first place, allowing to measure the change in the probability of an individual adopting e-commerce when the observed explanatory variables change, holding other factors fixed.

Since the dependent variable is dichotomous, a Logistic Regression Model (LRM), *Table 2 (2)*, has been estimated<sup>15</sup>. And to correct a possible sample selection bias, a Heckman Selection Model (HSM), *Table 2*

---

<sup>14</sup> Digital Skills is a self-elaborated index, based on the answers where the respondent declares whether he or she used specific Internet services and/or performed specific computer and Internet related tasks.

<sup>15</sup> All three models are estimated using Stata 15.

(3), has also been estimated. For all three models the specification is equivalent in terms of the considered dependent and explanatory variables and the corresponding base categories<sup>16</sup>.

---

<sup>16</sup> Base categories for each variable have been set as follow: Gender, Male; Age, 16 to 24 years old; Education, Primary or less; Digital Skills, Low; Habitat, lower than 20,000 inhabitants; Household Members, one member; Nationality, Foreigner; Employment Situation, Employed; Income, Low; Year, 2008; and Autonomous Community, Andalucía. The interaction term 'Digital Skills  $\times$  Age' uses the same base categories already fixed for interacted explanatory variables.



*Table 2*  
*Models of adoption of e-commerce by individual internet users. Linear Probability Model (LPM), Logistic Regression Model (LRM) and Heckman Selection Model (HSM). Pool data (2008-2017)*

		(1) E-commerce adoption. Linear Probability Model		(2) E-commerce adoption. Logistic Regression Model		(3) E-commerce adoption. Heckman Selection Model	
		Coef.	t	Odds ratios	z	Coef.	z
<b>Gender</b>	Female						
	Male	<b>0.04</b>	<b>10.33</b>	<b>1.29</b>	<b>10.40</b>	<b>0.04</b>	<b>10.2</b>
<b>Age</b>	16-24						
	25-34	-0.03	-1.89	0.82	-0.98	-0.03	-1.5
	35-44	<b>-0.09</b>	<b>-5.05</b>	<b>0.52</b>	<b>-3.49</b>	<b>-0.08</b>	<b>-4.64</b>
	45-54	<b>-0.11</b>	<b>-6.39</b>	<b>0.46</b>	<b>-4.13</b>	<b>-0.11</b>	<b>-5.86</b>
	55-64	<b>-0.16</b>	<b>-9.19</b>	<b>0.31</b>	<b>-6.00</b>	<b>-0.15</b>	<b>-8.35</b>
	> 65	<b>-0.19</b>	<b>-9.39</b>	<b>0.24</b>	<b>-6.38</b>	<b>-0.18</b>	<b>-8.37</b>
<b>Education</b>	Primary or less						
	Secondary	<b>0.04</b>	<b>5.77</b>	<b>1.45</b>	<b>6.36</b>	<b>0.04</b>	<b>5.13</b>
	Bachelor	<b>0.09</b>	<b>11.13</b>	<b>1.90</b>	<b>10.25</b>	<b>0.09</b>	<b>10.15</b>
	Master/PhD	<b>0.13</b>	<b>14.88</b>	<b>2.34</b>	<b>13.45</b>	<b>0.12</b>	<b>13.52</b>
<b>Digital Skills</b>	Low						
	Medium	<b>0.11</b>	<b>5.51</b>	<b>2.48</b>	<b>4.78</b>	<b>0.11</b>	<b>5.21</b>
	High	<b>0.26</b>	<b>12.65</b>	<b>4.88</b>	<b>8.71</b>	<b>0.25</b>	<b>12.2</b>
	Very high	<b>0.47</b>	<b>23.89</b>	<b>12.52</b>	<b>13.84</b>	<b>0.47</b>	<b>23.19</b>
<b>Habitat</b>	<20000						
	20000-100000	-0.01	-1.44	0.95	-1.57	-0.01	-1.78
	100000-500000	<b>-0.04</b>	<b>-5.07</b>	<b>0.78</b>	<b>-5.24</b>	<b>-0.04</b>	<b>-5.38</b>
	>500000	0.00	0.86	1.02	0.73	0	0.57
<b>Household Members</b>	One						
	Two	<b>-0.01</b>	<b>-2.10</b>	0.93	-1.85	<b>-0.02</b>	<b>-2.35</b>
	Three	<b>-0.04</b>	<b>-6.72</b>	<b>0.78</b>	<b>-6.36</b>	<b>-0.05</b>	<b>-6.86</b>
	Four	<b>-0.05</b>	<b>-7.00</b>	<b>0.76</b>	<b>-6.71</b>	<b>-0.05</b>	<b>-7.26</b>
	Five or more	<b>-0.08</b>	<b>-8.60</b>	<b>0.62</b>	<b>-8.34</b>	<b>-0.08</b>	<b>-8.72</b>
<b>Nationality</b>	Foreigner						
	Spanish	<b>0.03</b>	<b>3.30</b>	<b>1.16</b>	<b>2.82</b>	<b>0.03</b>	<b>3.09</b>
<b>Employment Situation</b>	Employed						
	Unemployed	<b>-0.05</b>	<b>-8.19</b>	<b>0.74</b>	<b>-7.80</b>	<b>-0.05</b>	<b>-7.96</b>
	Retired	<b>-0.04</b>	<b>-3.18</b>	<b>0.82</b>	<b>-2.72</b>	<b>-0.03</b>	<b>-2.49</b>
	Student	<b>-0.06</b>	<b>-4.71</b>	<b>0.73</b>	<b>-5.00</b>	<b>-0.06</b>	<b>-5.05</b>
	Housekeeper	<b>-0.02</b>	<b>-2.58</b>	<b>0.84</b>	<b>-2.65</b>	<b>-0.02</b>	<b>-1.91</b>
	Other	<b>-0.02</b>	<b>-1.60</b>	0.90	-1.36	<b>-0.02</b>	<b>-1.16</b>
<b>Income</b>	Low						
	Medium	<b>0.04</b>	<b>6.69</b>	<b>1.32</b>	<b>6.97</b>	<b>0.04</b>	<b>5.8</b>
	Medium-high	<b>0.11</b>	<b>15.78</b>	<b>1.91</b>	<b>15.14</b>	<b>0.1</b>	<b>14.45</b>
	High	<b>0.17</b>	<b>20.28</b>	<b>2.71</b>	<b>19.45</b>	<b>0.16</b>	<b>18.92</b>
<b>Digital Skills</b> × <b>Age</b> <sup>17</sup>	High × 55-64	<b>0.14</b>	<b>5.64</b>	<b>2.40</b>	<b>4.15</b>	<b>0.14</b>	<b>5.44</b>
	High × 65 or more	<b>0.12</b>	<b>3.83</b>	<b>2.40</b>	<b>3.61</b>	<b>0.11</b>	<b>3.59</b>
	Very high × 55-64	<b>0.08</b>	<b>3.50</b>	<b>2.04</b>	<b>3.30</b>	<b>0.08</b>	<b>3.26</b>
	Very high × 65 or more	<b>0.14</b>	<b>4.54</b>	<b>3.07</b>	<b>4.27</b>	<b>0.13</b>	<b>4.19</b>

<sup>17</sup> Complete estimates for the interaction term are shown in the Appendix, *Table A-1*.

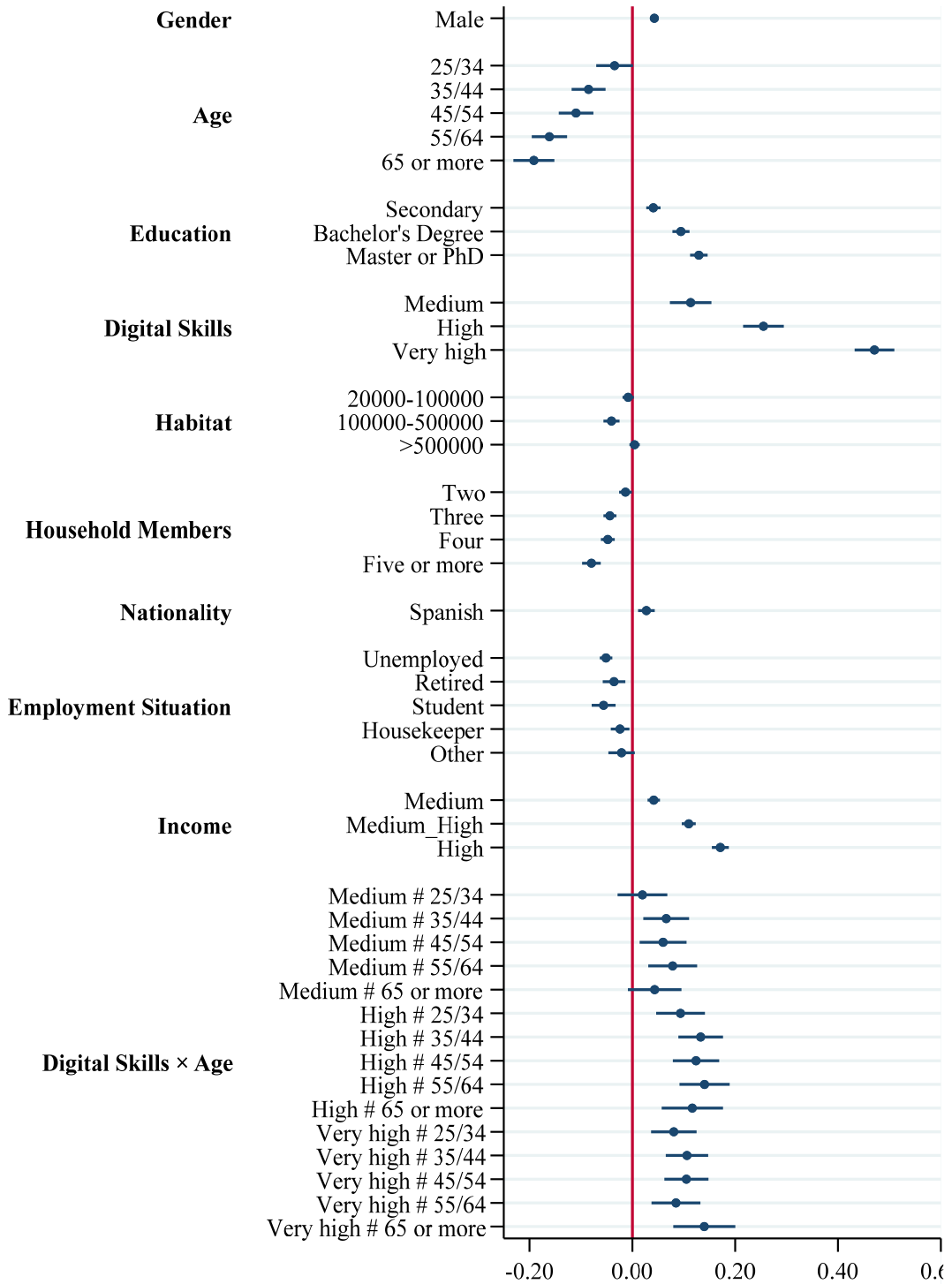
Table 2 (Continued)

Models of adoption of e-commerce by individual internet users. Linear Probability Model (LPM), Logistic Regression Model (LRM) and Heckman Selection Model (HSM). Pool data (2008-2017)

		(1) E-commerce adoption. Linear Probability Model		(2) E-commerce adoption. Logistic Regression Model		(3) E-commerce adoption. Heckman Selection Model	
		Coef.	t	Odds ratios	z	Coef.	z
<b>Year</b>	2008						
	2009	-0.01	-1.58	<b>0.87</b>	<b>-2.41</b>	-0.02	-1.82
	2010	<b>0.03</b>	<b>2.64</b>	<b>1.14</b>	<b>2.32</b>	<b>0.02</b>	<b>2.32</b>
	2011	0.02	1.54	1.05	0.78	0.01	1.15
	2012	0.01	1.35	1.06	1.01	0.01	0.74
	2013	<b>0.04</b>	<b>3.81</b>	<b>1.23</b>	<b>3.72</b>	<b>0.03</b>	<b>3.32</b>
	2014	<b>0.09</b>	<b>9.29</b>	<b>1.61</b>	<b>8.43</b>	<b>0.08</b>	<b>8.46</b>
	2015	<b>0.18</b>	<b>19.19</b>	<b>2.89</b>	<b>18.68</b>	<b>0.17</b>	<b>18.14</b>
	2016	<b>0.18</b>	<b>19.21</b>	<b>2.83</b>	<b>18.56</b>	<b>0.17</b>	<b>18.15</b>
2017	<b>0.24</b>	<b>26.31</b>	<b>4.17</b>	<b>25.81</b>	<b>0.23</b>	<b>24.71</b>	
<b>Autonomous Community</b>	Andalucía						
	Aragón	0.01	1.43	1.10	1.56	0.01	-1.82
	Asturias	<b>0.03</b>	<b>3.44</b>	<b>1.22</b>	<b>3.65</b>	<b>0.03</b>	<b>2.32</b>
	Baleares	<b>0.09</b>	<b>7.46</b>	<b>1.68</b>	<b>7.49</b>	0.08	1.15
	Canarias	<b>-0.05</b>	<b>-4.46</b>	<b>0.75</b>	<b>-4.31</b>	-0.05	0.74
	Cantabria	<b>0.05</b>	<b>5.26</b>	<b>1.40</b>	<b>5.42</b>	<b>0.05</b>	<b>3.32</b>
	Castilla la Mancha	0.02	1.79	1.11	1.68	<b>0.02</b>	<b>8.46</b>
	Castilla León	0.00	-0.05	1.01	0.19	<b>0</b>	<b>18.14</b>
	Cataluña	<b>0.05</b>	<b>5.68</b>	<b>1.32</b>	<b>5.68</b>	<b>0.05</b>	<b>18.15</b>
	Extremadura	0.01	1.07	1.06	0.98	<b>0.01</b>	<b>24.71</b>
	Galicia	0.00	0.00	1.00	0.09	0	-1.82
	La Rioja	<b>0.03</b>	<b>2.59</b>	<b>1.21</b>	<b>2.72</b>	<b>0.03</b>	<b>2.32</b>
	Madrid	<b>0.03</b>	<b>3.07</b>	<b>1.17</b>	<b>3.20</b>	0.02	1.15
	Navarra	<b>0.05</b>	<b>5.47</b>	<b>1.35</b>	<b>5.47</b>	0.05	0.74
	País Vasco	<b>0.05</b>	<b>5.82</b>	<b>1.39</b>	<b>5.86</b>	<b>0.05</b>	<b>3.32</b>
	Murcia	<b>-0.02</b>	<b>-2.52</b>	<b>0.87</b>	<b>-2.33</b>	<b>-0.02</b>	<b>8.46</b>
Valencia	<b>-0.02</b>	<b>-1.99</b>	0.91	-1.81	<b>-0.02</b>	<b>18.14</b>	
Ceuta	-0.04	-1.55	0.78	-1.64	<b>-0.04</b>	<b>18.15</b>	
Melilla	-0.02	-0.58	0.88	-0.73	<b>-0.01</b>	<b>24.71</b>	
<b>Constant</b>		0.01	0.44	0.05	-15.36	0.04	1.65
<b>N. observations</b>		<b>75,960</b>		<b>75,960</b>		<b>77,362</b>	
<b>F</b>		<b>556.10 DF: 70</b>		<b>10581.81 DF: 70</b>		<b>29332.2 DF: 70</b>	
<b>Wald <math>\chi^2</math></b>		<b>0.3280</b>		<b>10581.81 DF: 70</b>		<b>29332.2 DF: 70</b>	
<b>R<sup>2</sup></b>		<b>0.3280</b>		<b>0.2759</b>		<b>54.96 DF: 1</b>	
<b>Pseudo R<sup>2</sup></b>		<b>0.3280</b>		<b>0.2759</b>		<b>54.96 DF: 1</b>	
<b>Correctly classified</b>		<b>0.3280</b>		<b>75.95%</b>		<b>54.96 DF: 1</b>	
<b>Wald <math>\chi^2</math>: <math>H_0</math>: independent equations</b>		<b>0.3280</b>		<b>75.95%</b>		<b>54.96 DF: 1</b>	

Notes: Coefficients and statistics (t and z) significant at the 5% are represented in bold. Linear Probability (1), Logistic Regression (2) and Heckman Selection (3) models are estimated. Weighted estimations. In all three models, estimations refer to those Internet users (people who have used the Internet, at least once, within the last year) who have purchased online (or not) during the last year. Robust estimates. The R<sup>2</sup> and the pseudo R<sup>2</sup> are measures of the in-sample goodness of fit. The relevant statistics for testing the significance of the estimated coefficients is the F test and the  $\chi^2$  test of joint significance of all slope coefficients, which are high and have very low p-values: below 0.001 in all three cases. The percentages of correct classifications are above 75%. A polychoric correlation matrix has been computed to explore the possible collinearity among the regressors of the model. As shown in the Appendix, Table A 2, correlations are moderate.

Figure 6  
 Estimated Coefficients Linear Probability Model. Pool data (2008–2017)<sup>18</sup>



<sup>18</sup> Notes: point estimates are represented by dots; confidence intervals are represented by solid lines.

### **Lineal Probability Model**

The linear probability model is a first approach to modeling the relationship between a binary variable and its determinants. It suffers from well known limitations like predicted probabilities that may lie outside the [0, 1] interval. Nevertheless, it is recognized as a reliable indication of signs and significance of estimated coefficients. Moreover, it provides a straight-forward interpretation of the estimated coefficients in terms of differences of predicted probabilities.

Each significant coefficient of the independent variables in the LPM estimation represents changes in the probability that a person adopts e-commerce for private use (E-commerce = 1), everything else held constant. Being male increases the probability of buying online by 0.04. Higher levels of education are related with higher probabilities of adoption, going up to 0.09 for Bachelor and 0.13 for Master or PhD. A Spanish person has a probability 0.03 higher than a foreigner. Any other employment situation than being employed, decreases the probability of becoming an online buyer. The higher the income, the higher the probability: 0.04, 0.06 and 0.17, for Medium, Medium-high and High levels, respectively.

The year dummies seem to be especially relevant for the last five years, increasing the probability of an individual buying online from 0.04 in 2013 up to 0.24 in 2017. Something similar happens with the regional dummies, where living in Canarias lowers the probability by 0.05, while living in Balears increases by 0.09 the probability of adopting e-commerce.

In the LPM estimates, the interaction between Digital Skills and Age is statistically significant. The interaction term allows us to infer how the effect of Age on the e-commerce adoption is moderated by the level of Digital Skills<sup>19</sup>. High Digital Skills seems to have a positive influence, partly counteracting the negative effects on probability of e-commerce adoption of some age groups. The parameters on the interacted original variables, Age and Digital Skills, have no direct interpretation. Coefficients for each Age group represent the effect of Age on e-commerce adoption for an Internet user with Low Digital Skills; that is why all Age groups have negative effects. It is similar at analyzing Digital Skills, its coefficients represent the effect of Digital Skills on e-commerce adoption for an Internet user aged 16 to 24 years old.

### **Logistic Regression Model**

---

<sup>19</sup> If it is needed to be analysed the other way around: Interaction term allows us to infer how the effect of Digital Skills on the e-commerce adoption depends on the magnitude of Age.

The LRM is an alternative to the LPM. One of the objectives is to obtain predicted probabilities in the interval [0, 1]. This is achieved by fitting a nonlinear model, the LRM, which is widely used in the literature.

Odds ratios<sup>20</sup> estimated from the LRM also allow to establish the relationship between observed variables and the odds of an individual to become an online shopper. In line with LPM, the outcomes of LRM confirm that being male, account for higher levels of education, being Spanish, being employed and having higher income levels are positively related with the adoption of e-commerce.

Odds ratios of the Interaction term in LRM are not straightforward interpretable, as the magnitude of the interaction effect in non-linear models does not equal the marginal effect of the interaction term (Ai & Norton, 2003).

Yearly dummies denote mostly positive and significant effects along time. Odds ratios for 2015 and 2016 are almost twice the odds ratio for 2013; and, for 2017 its odds ratio is more than 4 times higher when it is compared with the base year, 2008. Meaning that the odds of buying online have increased substantially along the years and suggesting unobserved factors, that have had important effects in the years 2015 to 2017.

As with LPM estimates, odds ratios of the regions reflect considerable differences among Spanish Autonomous Communities. Positive and significant effects: Baleares, 1.68; Cantabria, 1.40; País Vasco, 1.39; Navarra, 1.35; Cataluña, 1.32; Asturias, 1.22; La Rioja, 1.21; and Madrid, 1.17. Those negatively related with the odds of shopping online are: Canarias and Murcia, with odds ratios of 0.75 and 0.87, respectively.

The magnitude of the coefficient estimates in the LPM and those of the LRM are not directly comparable. However the signs are comparable and also the magnitudes of t and z statistics, which show a similar pattern in models (1) and (2).

### **Heckman Selection Model**

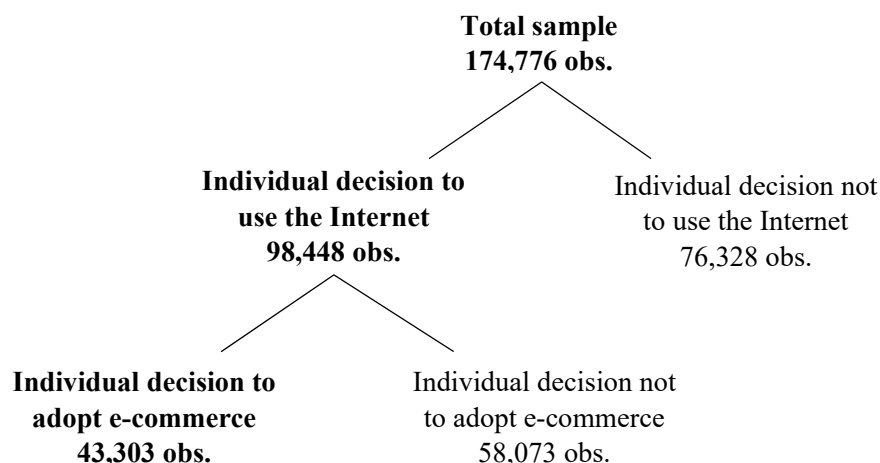
The use of e-commerce is only asked to those who answered having used the Internet at least once. Hence, models of e-commerce adoption are estimated with the resulting subset of the sample obtained by a previous decision individual made about using the Internet, *Figure 7*. As this might imply a selection bias,

---

<sup>20</sup> Odds ratios of the explanatory variables are depicted in *Figure A 1* and *Figure A 2*, in the Appendix section. Odds ratios greater than 1 describe a positive relationship, whereas if they are less than 1, they indicate a negative relationship.

Heckman’s Selection Model (HSM) is estimated<sup>21</sup>, as Cerno & Pérez-Amaral (2006a), Lera-López et al. (2011) and Pérez-Hernández & Sánchez-Mangas (2011), among others, did at modeling similar individual decision processes.

*Figure 7*  
*Decision process of the Internet and e-commerce adoption (2008–2017)*



Wald’s test for independent equations, with  $\chi^2 = 54.96$  and very low p-value, below 0.0001, suggests using Heckman’s method to estimate a model that allows for selection (Heckman, 1977) with this dataset. However, the magnitudes of the coefficients of the second stage estimation of HSM, *Table 2 (3)*, are comparable and remarkably similar to the results of the LPM, *Table 2 (1)*. Most of the categories of the explanatory variables coincide in their statistical significance in the three models. The main differences are observed in some of the regional variables (Balears, Canarias, Madrid and Navarra) that are no longer statistically significant, and others that become statistically significant, as Castilla la Mancha, Castilla Leon, Extremadura, Ceuta and Melilla.

As a selection equation<sup>22</sup> shown in table *Table A-3* of the appendix, a probit model of the decision of using the Internet has been estimated. All the explanatory variables in the second stage estimation also include as covariates in the first stage equation, except for Digital Skills, to avoid simultaneity, since Digital Skills account for some Internet related activities. Computer availability is used as exclusion restriction, as it is highly correlated with using Internet and weakly correlated with the adoption of e-commerce; besides, it might pick up the unobserved individual factors that affect the tendency to use Internet services.

<sup>21</sup> Stata – Heckman selection model: ‘heckman’ command and ‘maximum likelihood estimates’ option (which allows to include the sample elevation factor for all observations, as well as compute robust standard errors.).

<sup>22</sup> Estimation results of first stage of HSM: Appendix, *Table A-3*.

Except Gender and Autonomous Community, all the coefficients of the variables included in the selection model are statistically significant, at the 5%, to explain the individual decision of using the Internet.  $\chi^2$  test of joint significance of all slope coefficients is 2631.86, which is high and has a p-value below 0.001.

## 6. Conclusions

### a. General conclusions

When the proportion of e-buyers in Spain is compared with EU-28 average for 2017, it may seem that the gap is being bridged, but this is an illusion due to data aggregation. If we zoom in specific factors (Age, Education, Digital Skills, Habitat, Household Members, Nationality, Employment Situation, Income, Time and Regions), it has been improved overall, however, the distances between category groups within those factors still exist and are especially remarkable, as: younger than 25, primary vs. those with higher education, individuals with low digital skills vs. very high level, employed vs. any other employment situation, lower vs. higher income.

Previous works that have used cross sectional data from 2016, already pointed out how sociodemographic individual characteristics, and economic factors, explain the adoption of three Internet services (e-commerce, e-banking and e-government)<sup>23</sup>; and, many of them are statistically significant at explaining the individual decision to perform cross-border e-commerce<sup>24</sup>. The results of the present work reinforce the conclusions and allow to consider aggregate time effects and unobserved economic environment characteristics through the introduction of year dummies.

To the best of our knowledge, this is the first time the period 2008-2017 of this survey has been analyzed using pooled data to model and estimate the determinants of e-commerce adoption in Spain<sup>25</sup>.

Signs and statistical significance of the determinants mostly coincide for Linear Probability, Logistic Regression and Heckman models. The findings have implications for consumers, policy makers, and firms.

Some of the determinants that reduce the probability of buying online can hardly be directly affected by policy measures. Among others, these are the factors that have a negative impact: being female, age, not

---

<sup>23</sup> Garín Muñoz, et. al., (2018)

<sup>24</sup> Valarezo, et. al., (2018)

<sup>25</sup> Correa, et al. (2015) and Pérez-Hernández & Sánchez-Mangas (2011), estimated the determinants of buying online, using pooled data for periods 2008-2014 and 2004-2009, respectively.

being Spanish, any other employment situation than being employed, and belonging to Autonomous Communities of Canarias, Murcia and Valencia.

On the other hand, factors such as Education, Digital Skills and living in specific regions, have a positive influence and are more susceptible to be targeted by policy measures for reducing the e-commerce divide. And, as expected, Higher Income is also positively related with e-commerce adoption.

The interaction between Digital Skills and Age points out that high digital skills have a positive influence, partly counteracting the negative effect of some age groups.

### **b. Policy recommendations**

There are different types of actors interested in fostering e-commerce. Specific policy recommendations are proposed for demand side actions, to promote governmental initiatives and to address specific strategies by the private sector.

On the demand side the main goal for policy makers should be to narrow the existing digital divide between groups with low and high penetration rates of e-commerce. Considering that even small increases in Education and Digital Skills substantially improve the probability of buying online, the economic policy could be targeted to specific groups, especially to individuals with lower income, lower levels of digital skills and education, older people, unemployed, retired, housekeepers, and in general for those that belong to groups with lower odds of adopting e-commerce. Training on specific Digital Skills, technical support (online, by phone and in person) and diffusion through conventional channels may bring about the untapped opportunities of e-commerce.

From the government standpoint relevant initiatives can be carried out, such as: to promote the adoption of complementary services (e.g., e-government, e-health, e-learning, among others), to identify and re-edit successful supply side programs, to reduce transaction costs (e.g., red tape, trade and taxation barriers, geographic and linguistic barriers) and to foster the development of efficient e-commerce platforms. Measures of policy success should be based on adoption gains on those groups that are worst placed, instead of only accounting at an aggregate level.

The private sector can play an active role, adopting easy-to-use platforms, encouraging customer reviews, designing promotions and communication campaigns for groups with low penetration, and streamlining payment and transaction process. Technology already allows e-commerce companies to serve websites and highly customized user interfaces based on artificial intelligence systems trained with user data. It can



also adjust and customize front pages and web application user interfaces to match the needs of those groups with lower penetration rates. Another helpful initiative could be the implementation of in-store demonstration of online buying; and for those pure play e-commerce<sup>26</sup>, physical points of purchase equipped with display devices, where prospective customers can go through the online purchase process (as if they did it from their own devices) with the help of shopping assistants<sup>27</sup>.

### **c. Caveats**

Some of the limitations found in this study are related with the lack of specific information about barriers that stopped Internet users from buying online, since the ICT-H INE's survey was not specifically designed for this purpose. Controlling for individual (unobserved) effects would be useful and not having this represents another important caveat. Supply side data also may be useful to perform a complementary analysis to get richer context for results and conclusions.

### **d. Further research**

The next stage in the further research agenda will be analyzing dynamic models, using the full panel data set. Doing so we will be able to control for individual (unobserved) effects and account for possible simultaneity and dynamic relationships. Specific online transactions and digital activities that are performed or affect individuals also deserve attention such as: buying and selling online different types of goods and services, cloud computing, use of digital certificates, sharing economy services, digital transformation and the use of ICTs at work, among others.

---

<sup>26</sup> A pure play e-commerce is an e-commerce business that only sell through the Internet.

<sup>27</sup> A good example are the physical stores of the third largest Spanish e-commerce firm, Pcomponentes.com, where customers and prospective customers get advice and guide for later purchasing online.

## REFERENCES

- Ai, C., & Norton, E. C. (2003). Interaction terms in logit and probit models. *Economics Letters*, *80*(1), 123–129. [https://doi.org/10.1016/S0165-1765\(03\)00032-6](https://doi.org/10.1016/S0165-1765(03)00032-6)
- Ajzen, I. (1991). The theory of planned behavior”. *Organizational Behavior and Human Decision Processes*, *50* (2), 79–211.
- Ajzen, Icek, & Fishbein, M. (1977). Attitude-behavior relations: A theoretical analysis and review of empirical research. *Psychological Bulletin*, *84*(5), 888–918. <https://doi.org/10.1037/0033-2909.84.5.888>
- Ajzen, Icek, & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. Englewood Cliffs, N.J.: Prentice-Hall.
- Alonso, J., & Arellano, A. (2015). Heterogeneity and diffusion in the digital economy: Spain’s case. Retrieved November 21, 2018, from [https://www.bbva-research.com/wp-content/uploads/2015/11/15-28\\_WP-Ec\\_Digital\\_e.pdf](https://www.bbva-research.com/wp-content/uploads/2015/11/15-28_WP-Ec_Digital_e.pdf)
- Cerno, L., & Pérez-Amaral, T. (2006a). Demand for Internet Access and Use in Spain. In B. Preissl & J. Müller (Eds.), *Governance of Communication Networks: Connecting Societies and Markets with IT* (pp. 333–353). Heidelberg: Physica-Verlag HD. [https://doi.org/10.1007/3-7908-1746-5\\_18](https://doi.org/10.1007/3-7908-1746-5_18)
- Cerno, L., & Pérez-Amaral, T. (2006b). *Medición y determinantes de la brecha tecnológica en España*. Retrieved from <https://eprints.ucm.es/7909/1/0601.pdf>
- Cerno, L., & Pérez-Amaral, T. (2009). E-Commerce Use in Spain. In P. Curwen, J. Haucap, & B. Preissl (Eds.), *Telecommunication Markets: Drivers and Impediments* (pp. 157–172). Heidelberg: Physica-Verlag HD. [https://doi.org/10.1007/978-3-7908-2082-9\\_9](https://doi.org/10.1007/978-3-7908-2082-9_9)
- Correa, M., García, J. R., & Tabanera, A. (2015a). Comercio electrónico y hábitos de consumo en España: la importancia de la banca on-line. *Observatorio de Economía Digital, BBVA Research*. Retrieved from <https://www.bbva-research.com/publicaciones/comercio-electronico-y-habitos-de-consumo-en-espana-la-importancia-de-la-banca-on-line/>
- Correa, M., García, J. R., & Tabanera, A. (2015b). Comercio electrónico y hábitos de consumo en España: la importancia de la banca on-line. *Observatorio de Economía Digital, BBVA Research*.
- Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User acceptance of computer technology: a comparison of two theoretical models. *Management Sciences*, (35:8), 982–1003.
- Davis, F. D., & Warshaw, P. R. (1989). User acceptance of information technology: System characteristics, user perceptions, and behavioral impacts. *Management Sciences*, (35:8), 982–1003.
- Davis, Fred D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, *13*(3), 319–340. <https://doi.org/10.2307/249008>

- Davis, Fred D. (1993). User acceptance of information technology: system characteristics, user perceptions and behavioral impacts. *International Journal of Man-Machine Studies*, 38(3), 475–487. <https://doi.org/10.1006/imms.1993.1022>
- Demoussis, M., & Giannakopoulos, N. (2006). Facets of the digital divide in Europe: Determination and extent of internet use. *Economics of Innovation and New Technology*, 15(3), 235–246. <https://doi.org/10.1080/10438590500216016>
- DiMaggio, P., & Hargittai, E. (2001). *From the “Digital Divide” to “Digital Inequality”*: Studying Internet Use as Penetration Increases (Working Papers No. 47). Princeton University, Woodrow Wilson School of Public and International Affairs, Center for Arts and Cultural Policy Studies. Retrieved from <https://ideas.repec.org/p/pri/cpanda/workpap15.html.html>
- European Commission. (2018a). Digital Economy and Society Index (DESI) 2018 Country Report Spain. Retrieved November 11, 2018, from <https://ec.europa.eu/digital-single-market/en/countries-performance-digitisation>
- European Commission. (2018b). The Digital Economy and Society Index (DESI). Retrieved November 11, 2018, from <https://ec.europa.eu/digital-single-market/en/desi>
- Eurostat. (2018a). Glossary:E-commerce - Statistics Explained. Retrieved October 7, 2018, from <https://ec.europa.eu/eurostat/statistics-explained/index.php/Glossary:E-commerce>
- Eurostat. (2018b). ICT usage in households and by individuals. Retrieved October 6, 2018, from [https://ec.europa.eu/eurostat/cache/metadata/en/isoc\\_i\\_esms.htm](https://ec.europa.eu/eurostat/cache/metadata/en/isoc_i_esms.htm)
- Garín Muñoz, T., & Pérez-Amaral, T. (2011). Factores Determinantes del Comercio Electrónico en España. *Boletín Económico Del ICE*, 3016, 51-65.
- Garín-Muñoz, T., López, R., Pérez-Amaral, T., Herguera, I., & Valarezo, A. (2019). Models for individual adoption of eCommerce, eBanking and eGovernment in Spain. *Telecommunications Policy*, 43(1), 100–111. <https://doi.org/10.1016/j.telpol.2018.01.002>
- Hargittai, E. (2001). Second-Level Digital Divide: Mapping Differences in People’s Online Skills. *ArXiv:Cs/0109068*. Retrieved from <http://arxiv.org/abs/cs/0109068>
- Heckman, J. J. (1977). *Sample Selection Bias As a Specification Error (with an Application to the Estimation of Labor Supply Functions)* (Working Paper No. 172). National Bureau of Economic Research. <https://doi.org/10.3386/w0172>
- Helsper, E. J., & van Deursen, A. J. A. M. (2015). The Third-Level Digital Divide: Who Benefits Most from Being Online? In *Communication and Information Technologies Annual* (Vol. 10, pp. 29–52). Emerald Group Publishing Limited. <https://doi.org/10.1108/S2050-206020150000010002>
- INE. (2018). Encuesta sobre equipamiento y uso de tecnologías de información y comunicación en los hogares. Documentos metodológicos y cuestionarios. Retrieved October 7, 2018, from [http://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica\\_C&cid=1254736176741&menu=metodologia&idp=1254735976608](http://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176741&menu=metodologia&idp=1254735976608)

- Lera-López, F., Billon, M., & Gil, M. (2011). Determinants of Internet use in Spain. *Economics of Innovation and New Technology*, 20(2), 127–152. <https://doi.org/10.1080/10438590903378017>
- Norris, P., & Norris, M. L. in C. P. P. (2001). *Digital Divide: Civic Engagement, Information Poverty, and the Internet Worldwide*. Cambridge University Press.
- NTIA (National Telecommunications and Information Administration). (1995). Falling Through the Net: A Survey of the “Have Nots” in Rural and Urban America. Retrieved November 27, 2018, from <https://www.ntia.doc.gov/ntiahome/fallingthru.html>
- OECD. (2001). Understanding the Digital Divide. <https://doi.org/10.1787/236405667766>
- Oliver, R. L. (1980). A Cognitive Model of the Antecedents and Consequences of Satisfaction Decisions. *Journal of Marketing Research*, 17(4), 460–469. <https://doi.org/10.2307/3150499>
- Pérez-Hernández, J., & Sánchez-Mangas, R. (2011). To have or not to have Internet at home: Implications for online shopping. *Information Economics and Policy*, 23(3), 213–226. <https://doi.org/10.1016/j.infoecopol.2011.03.003>
- Rogers, E. M. (2003). *Diffusion of Innovations, 5th Edition* (Edición: 5th). New York: Simon & Schuster.
- Srinuan, C., & Bohlin, E. (2011). Understanding the digital divide: A literature survey and ways forward. Budapest: ITS. Retrieved from <https://www.econstor.eu/handle/10419/52191>
- Taylor, S., & Todd, P. A. (1995). Understanding Information Technology Usage: A Test of Competing Models. *Information Systems Research*, 6(2), 144–176. <https://doi.org/10.1287/isre.6.2.144>
- Valarezo, Á., Pérez-Amaral, T., Garín-Muñoz, T., Herguera García, I., & López, R. (2018). Drivers and barriers to cross-border e-commerce: Evidence from Spanish individual behavior. *Telecommunications Policy*, 42(6), 464–473. <https://doi.org/10.1016/j.telpol.2018.03.006>
- Varian, H. R. (2002). The Demand for Bandwidth: Evidence from the INDEX Project. In *Crandall, R.W., and Alleman, J.H. (Eds.) Broadband: Should We Regulate High-Speed Internet Access? Aei-Brookings Joint Center for Regulatory Studies, American Enterprise Institute* (pp. 39–56). Press.
- Venkatesh, V., & Davis, F. D. (2000). A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. *Management Science*, 46(2), 186–204. <https://doi.org/10.1287/mnsc.46.2.186.11926>
- Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly*, 27(3), 425–478. <https://doi.org/10.2307/30036540>

**Appendix: Supplementary tables and figures.**

*Table A-1*

*Models of adoption of e-commerce by individual internet users. Linear Probability Mode (LPM) and Logistic Regression. Complete estimates for the interaction term Digital Skills × Age. Pooled data (2008-2017)*

		(1) E-commerce adoption. Linear Probability Model		(2) E-commerce adoption. Logistic Regression		(3) E-commerce adoption. Heckman Selection Model	
		Coef.	t	Odds ratios	Z	Coef.	z
<b>Digital Skills × Age</b>	Medium × 16-24						
	Medium × 25-34	0.02	0.79	1.09	0.40	0.02	0.65
	Medium × 35-44	<b>0.07</b>	<b>2.90</b>	<b>1.65</b>	<b>2.43</b>	<b>0.07</b>	<b>2.85</b>
	Medium × 45-54	<b>0.06</b>	<b>2.57</b>	<b>1.57</b>	<b>2.16</b>	<b>0.06</b>	<b>2.57</b>
	Medium × 55-64	<b>0.08</b>	<b>3.23</b>	<b>1.89</b>	<b>2.94</b>	<b>0.08</b>	<b>3.17</b>
	Medium × 65 or more	0.04	1.63	<b>1.73</b>	<b>2.23</b>	0.04	1.57
	High × 16-24						
	High × 25-34	<b>0.09</b>	<b>3.88</b>	1.46	1.79	<b>0.09</b>	<b>3.57</b>
	High × 35-44	<b>0.13</b>	<b>6.00</b>	<b>2.12</b>	<b>3.79</b>	<b>0.13</b>	<b>5.83</b>
	High × 45-54	<b>0.12</b>	<b>5.38</b>	<b>2.01</b>	<b>3.47</b>	<b>0.12</b>	<b>5.26</b>
	High × 55-64	<b>0.14</b>	<b>5.64</b>	<b>2.40</b>	<b>4.15</b>	<b>0.14</b>	<b>5.44</b>
	High × 65 or more	<b>0.12</b>	<b>3.83</b>	<b>2.40</b>	<b>3.61</b>	<b>0.11</b>	<b>3.59</b>
	Very high 16 × 24						
	Very high × 25-34	<b>0.08</b>	<b>3.58</b>	<b>1.59</b>	<b>2.19</b>	<b>0.07</b>	<b>3.21</b>
	Very high × 35-44	<b>0.11</b>	<b>5.06</b>	<b>2.21</b>	<b>3.98</b>	<b>0.1</b>	<b>4.84</b>
	Very high × 45-54	<b>0.11</b>	<b>4.80</b>	<b>2.16</b>	<b>3.76</b>	<b>0.1</b>	<b>4.61</b>
	Very high × 55-64	<b>0.08</b>	<b>3.50</b>	<b>2.04</b>	<b>3.30</b>	<b>0.08</b>	<b>3.26</b>
Very high × 65 or more	<b>0.14</b>	<b>4.54</b>	<b>3.07</b>	<b>4.27</b>	<b>0.13</b>	<b>4.19</b>	

*Table A-2.*

*Polychoric correlations among selected independent variables*

	Gender	Age	Education	Digital Skills	Household Members	Income
Gender	1.00					
Age	0.02	1.00				
Education	-0.07	-0.02	1.00			
Digital Skills	0.07	-0.31	0.44	1.00		
Household Members	0.02	-0.32	-0.02	0.07	1.00	
Income	0.08	0.05	0.45	0.31	0.24	1.00

Note: for assessing possible multicollinearity. Computed polychoric correlations are moderate. Little sign of multicollinearity is found.

Table A-3  
 Model of Internet use. Heckman first stage estimates. Probit Regression Model.  
 Pooled data (2008-2017)

Internet use. Probit Regression Model - Heckman first stage estimates			
		Coef.	z
<b>Gender</b>	Female		
	Male	0.04	1.1
<b>Age</b>	16-24		
	25-34	<b>-0.18</b>	<b>-2.25</b>
	35-44	<b>-0.36</b>	<b>-4.54</b>
	45-54	<b>-0.56</b>	<b>-6.97</b>
	55-64	<b>-0.7</b>	<b>-8.44</b>
	> 65	<b>-0.87</b>	<b>-9.02</b>
<b>Education</b>	Primary or less		
	Secondary	<b>0.11</b>	<b>2.51</b>
	Bachelor	<b>0.33</b>	<b>5.62</b>
	Master/PhD	<b>0.59</b>	<b>9.09</b>
<b>Habitat</b>	<20000		
	20000-100000	<b>0.14</b>	<b>3.13</b>
	100000-500000	<b>0.2</b>	<b>3.24</b>
	>500000	<b>0.09</b>	<b>2.55</b>
<b>Household Members</b>	One		
	Two	-0.01	-0.32
	Three	<b>-0.09</b>	<b>-2.09</b>
	Four	-0.07	-1.27
	Five or more	-0.07	-0.91
<b>Nationality</b>	Foreigner		
	Spanish	0	0.02
<b>Employment Situation</b>	Employed		
	Unemployed	-0.09	-1.9
	Retired	<b>-0.39</b>	<b>-7.2</b>
	Student	<b>1.16</b>	<b>7.5</b>
	Housekeeper	<b>-0.36</b>	<b>-6.21</b>
	Other	<b>-0.34</b>	<b>-4.24</b>
<b>Income</b>	Low		
	Medium	<b>0.17</b>	<b>4.26</b>
	Medium-high	<b>0.33</b>	<b>6.68</b>
	High	<b>0.56</b>	<b>8.11</b>

		Coef.	z
<b>Computer availability</b>	No		
	Yes	<b>1.04</b>	<b>29.34</b>
<b>Year</b>	2008		
	2009	<b>0.17</b>	<b>2.53</b>
	2010	<b>0.2</b>	<b>3.18</b>
	2011	<b>0.26</b>	<b>3.79</b>
	2012	<b>0.45</b>	<b>6.35</b>
	2013	<b>0.3</b>	<b>4.4</b>
	2014	<b>0.55</b>	<b>7.84</b>
	2015	<b>0.57</b>	<b>7.58</b>
	2016	<b>0.59</b>	<b>8.67</b>
	2017	<b>0.94</b>	<b>12.84</b>
<b>Autonomous Community</b>	Andalucía		
	Aragón	0.07	0.92
	Asturias	-0.03	-0.45
	Balears	0.13	1.51
	Canarias	0	-0.01
	Cantabria	0	0.04
	Castilla la Mancha	0.03	0.42
	Castilla León	-0.04	-0.66
	Cataluña	0.01	0.12
	Extremadura	0.08	1.11
	Galicia	0.02	0.27
	La Rioja	0.05	0.57
	Madrid	0.05	0.75
	Navarra	-0.06	-0.92
	País Vasco	0.01	0.22
	Murcia	<b>-0.15</b>	<b>-2.09</b>
Valencia	0.05	0.8	
Ceuta	-0.04	-0.2	
Melilla	-0.1	-0.62	
<b>Constant</b>		0.86	7.48

<b>N. observations</b>	77,362
<b>Wald <math>\chi^2</math></b>	2631.86 DF: 53
<b>Pseudo R<sup>2</sup></b>	0.2363
<b>Correctly classified</b>	97.08%
<b>Wald <math>\chi^2</math>: <math>H_0</math>: independent equations</b>	54.96 DF: 1

Notes: Coefficients and statistics (t and z) significant at the 5% are represented in bold. Heckman Selection first step is estimated. Robust estimates. Pseudo R<sup>2</sup> is a measure of the in-sample goodness of fit. The relevant statistics for testing the significance of the estimated coefficients is the  $\chi^2$  test of joint significance of all slope coefficients equal to zero, which is high and has very low p-value, below 0.001. The percentages of correct classifications are above 97%.

Figure A-1  
Odds ratios of e-commerce. Pool (2008 – 2017)



Figure A-2  
Odds of e-commerce. Pool (2008-2017)

