UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE CIENCIAS ECONÓMICAS Y EMPRESARIALES

TESIS DOCTORAL

Modelos cuantitativos de adopción y uso de tecnologías de información y comunicación en España

Quantitative models of adoption and use of information and communication technologies in Spain

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

Ángel Eduardo Valarezo Unda

Director

Teodosio Pérez Amaral

Madrid

UNIVERSIDAD COMPLUTENSE MADRID

DECLARACIÓN DE AUTORÍA Y ORIGINALIDAD DE LA TESIS PRESENTADA PARA OBTENER EL TÍTULO DE DOCTOR

D. **ÁNGEL EDUARDO VALAREZO UNDA** _____, estudiante en el Programa de Doctorado en **ECONOMÍA** _____, de la Facultad de **CIENCIAS ECONÓMICAS Y EMPRESARIALES** de la Universidad Complutense de Madrid, como autor/a de la tesis presentada para la obtención del título de Doctor y titulada:

MODELOS CUANTITATIVOS DE ADOPCIÓN Y USO DE TECNOLOGÍAS DE LA INFORMACIÓN Y COMUNICACIÓN EN ESPAÑA _____
_____

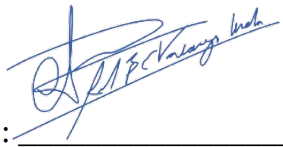
y dirigida por: **TEODOSIO PÉREZ AMARAL** _____
_____

DECLARO QUE:

La tesis es una obra original que no infringe los derechos de propiedad intelectual ni los derechos de propiedad industrial u otros, de acuerdo con el ordenamiento jurídico vigente, en particular, la Ley de Propiedad Intelectual (R.D. legislativo 1/1996, de 12 de abril, por el que se aprueba el texto refundido de la Ley de Propiedad Intelectual, modificado por la Ley 2/2019, de 1 de marzo, regularizado, aclarando y armonizando las disposiciones legales vigentes sobre la materia), en particular, las disposiciones referidas al derecho de cita.

Del mismo modo, asumo frente a la Universidad cualquier responsabilidad que pudiera derivarse de la autoría o falta de originalidad del contenido de la tesis presentada de conformidad con el ordenamiento jurídico vigente.

En Madrid, a 16 de septiembre de 2019



Fdo.: _____

Esta DECLARACIÓN DE AUTORÍA Y ORIGINALIDAD debe ser insertada en la primera página de la tesis presentada para la obtención del título de Doctor.

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE CIENCIAS ECONÓMICAS Y EMPRESARIALES

Departamento de Análisis Económico

Instituto Complutense de Análisis Económico | ICAE

Tesis Doctoral

MODELOS CUANTITATIVOS DE ADOPCIÓN Y USO DE TECNOLOGÍAS DE INFORMACIÓN Y COMUNICACIÓN EN ESPAÑA

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

REALIZADA Y PRESENTADA POR

Ángel Eduardo Valarezo Unda

Director

Teodosio Pérez Amaral

Madrid, 2019

COMPLUTENSE UNIVERSITY OF MADRID

FACULTY OF ECONOMICS AND BUSINESS

Economic Analysis Department

Complutense Institute of Economic Analysis | ICAE

Doctoral Thesis

QUANTITATIVE MODELS OF ADOPTION AND USE OF INFORMATION AND COMMUNICATION TECHNOLOGIES IN SPAIN

SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY, BY

Ángel Eduardo Valarezo Unda

Advisor

Teodosio Pérez Amaral, Ph.D.

Madrid, 2019

Acknowledgments

My deepest gratitude, respect, and admiration to Professor Teodosio Pérez Amaral. I feel quite fortunate to have had the opportunity to learn from a such great academic and even better person.

Thank you, Professors Teresa Garín Muñoz, Rafael López and Iñigo Herquera. Being part of a high standard research group has been quite an opportunity for me. My gratitude for your patience and generosity.

I am especially grateful to Professor Luis Puch. It is a great honor that the grant that allowed me to work on this thesis is related to a project[1] led by Professor Puch.

My gratitude to the Economic Analysis Department and to the researchers of Instituto Complutense de Análisis Económico. Thanks for letting me be part and learn from you.

I am indebted for the valuable comments received in following congresses and seminars:

1. 26th ITS (International Telecommunications Society) European Conference, Madrid-Spain, 2015
2. 27th ITS European Conference, Cambridge-United Kingdom, 2016
3. 28th ITS European Conference, Passau-Germany, 2017
4. PRODECON-CM Workshop on Innovation and Industrial Organization, Madrid-Spain, 2017.
5. ICAE Ph.D. Student Workshop, Madrid-Spain, 2018

[1] Project: Technology, Human Capital, Innovation and Commerce (ECO2016-76818-C3-3-P).

6. 29th ITS European Conference, Trento-Italy, 2018
7. ICAE PhD Candidate Seminar, Madrid-Spain, 2019
8. 30th its European Conference, Helsinki Finland, 2019

Thanks to the Ministry of Economy of Spain and its "National Program for the Promotion of Talent and Its Employability" for granting the scholarship that allowed me full dedication research training since 2018.

The research presented in this thesis was 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.

To María Dolores. To my parents, Ivanova and Angel Isaac. And to my sisters, Svetlana, Rocío and Anamaría. Thank you for encouraging me, for being my heroes, for your patience and impatience, and above all, thank you for your time.

Agradecimientos

Mi más profunda gratitud, respeto y admiración al Profesor Teodosio Pérez Amaral. Me siento sumamente afortunado por haber tenido la oportunidad de aprender de un gran académico y, sobre todo, de una gran persona.

Gracias, profesores Teresa Garín Muñoz, Rafael López e Iñigo Herguera. Ser parte de un grupo de investigación de alto nivel, como el que conformáis, ha sido una experiencia formidable. Mi gratitud por vuestra paciencia y generosidad.

Estoy especialmente agradecido con el profesor Luis Puch. Es un gran honor que la beca que me ha permitido trabajar en esta tesis esté relacionada con un proyecto[2] liderado por el profesor Puch.

Mi agradecimiento al Departamento de Análisis Económico y los investigadores del Instituto Complutense de Análisis Económico. Gracias por permitirme formar parte y aprender de vosotros.

Agradezco los valiosos comentarios recibidos en los siguientes congresos y seminarios:

1. 26ª Conferencia Europea de la ITS (International Telecommunications Society), Madrid–España, 2015
2. 27ª Conferencia Europea de la ITS, Cambridge–Reino Unido, 2016
3. 28ª Conferencia Europea de la ITS, Passau–Alemania, 2017
4. Taller de innovación y organización industrial de PRODECON-CM, Madrid–España, 2017.

[2]Proyecto: tecnología, capital humano, innovación y Comercio (ECO2016-76818-C3-3-P).

5. ICAE PhD Student Workshop, Madrid–España, 2018
6. 29ª Conferencia Europea de la ITS, Trento–Italia, 2018
7. ICAE PhD Candidate Seminar, Madrid-España 2019.
8. 30ª Conferencia Europea de la ITS, Helsinki–Finlandia.

Gracias al Ministerio de Economía de España y su "programa nacional para la promoción del talento y su empleabilidad" por la concesión de la beca que me ha permitido desde 2018 dedicarme a tiempo completo a mi formación como investigador.

La investigación presentada en esta tesis ha sido financiada por la Comunidad Autónoma de Madrid. Proyecto: Finanzas, innovación y estrategias: análisis económico de la productividad empresarial y sus determinantes (PRODECON-CM HM S2015/HUM-3491), 2016-2018.

A María Dolores. A mis padres, Ivanova y Angel Isaac. Y a mis hermanas, Svetlana, Rocío y Anamaría. Gracias por animarme, por ser mis héroes, por vuestra paciencia e impaciencia, y, sobre todo, gracias por vuestro tiempo.

Table of Contents

List of tables

Chapter 1

Chapter 2

Chapter 3

Chapter 4

List of figures

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Abstract

The digital transformation involves profound changes, with economic, social, and political implications. People, organizations, and governments work, interact and try to thrive in increasingly dynamic and complex environments where information and communication technologies are acquiring a growing prominence.

Universities, private companies, governments, and supranational institutions are aware that the digital economy implies both threats and opportunities, and that the design and development of policies and strategies must be based on the rigorous and continuous analysis of the phenomenon in all its dimensions: technical, political, economic and social. The European Union, committed to foster the Digital Single Market, bases the design and execution of its strategies in a specific framework of studies and surveys on the evolution and impact of the digital economy on individuals and companies.

The data used in this research comes from the Survey on Equipment and Use of Information and Communication Technologies in Households in Spain, carried out annually since 2002 by the National Institute of Statistics (INE), following methodological recommendations and the questionnaire models of the Statistical Office of the European Union (EUROSTAT). This survey is representative of the national territory and comparable across countries of the European Union and other international organizations.

The research studies included in this thesis present the specification and estimation of several Quantitative Models of Adoption and Use of ICT (Information and Communication Technologies), where e-commerce and its determinants of adoption and use in Spain are the focus of this research. The main objective of this thesis is to propose quantitative models that allow to identify the individual determinants of adoption and use of digital services and information and communication technologies for particular purpose. The discussion of

the results and the conclusions allow to suggest specific policy recommendations to foster the individual adoption of ICT and digital services.

This memory is composed of four chapters. The first one is an introductory chapter and the following correspond to three empirical analysis works, in which the standard neoclassical utility maximization approach is used as a theoretical framework.

The first chapter consists of a general introduction on the challenges presented by the evolution of the digital economy and how the different actors are responding. The pillars on which the European Commission bases its strategy to reach a Digital Single Market and how the development of electronic commerce plays a priority role are detailed.

The second chapter analyzes the determinants of the individual decision to conduct cross-border e-commerce in Spain. Using logistic regression techniques, the 2016 survey on Equipment and Use of Information and Communication Technologies in Households in Spain is analyzed. The results suggest a positive relationship between the adoption of cross-border e-commerce and the following factors: being male, having high levels of education, computing and Internet-related skills, consulting the opinions of other users and having a foreign nationality. The conclusions highlight the importance of education, digital skills, online trust and the use of information about opinions and recommendations, proposing related measures to promote the use of cross-border electronic commerce.

The third chapter focuses on the study of the individual adoption of electronic commerce in Spain, analyzing the main socio-economic and demographic determinants. The analysis is carried out on a pool data set of 174,776 observations, corresponding to the surveys from 2008 to 2016 on Equipment and Use of Information and Communication Technologies in Households in Spain. The Linear Probability Model, the Logistic Regression Model, and the Heckman Selection Correction Model have been used. The results indicate that the adoption of electronic commerce is positively related to being a man, level of education, income, digital skills, being Spanish and having an employee's employment situation; while variables such as being a woman or age are negatively related. The interaction included in the model suggests that high levels of digital skills counteract the negative effects of some age groups. To promote the adoption of e-commerce, specific recommendations are

proposed for actions directed towards the demand side, government initiatives, and strategies by the private sector.

The fourth and final chapter analyzes the adoption of electronic commerce and the digital divide in Spain. To this end, a panel data set on individuals is used, with 133,420 observations for the period 2008-2016. This work substantially improves other existing ones in the international literature, due to the characteristics of the database, the inclusion of variables previously neglected, the economic model, and the variety of panel data analysis techniques that have been used. The evolution of digital divides is measured over time, and the impact of socioeconomic and demographic variables is quantified. Models with high explanatory power are obtained, in which the variables income and digital skills are highly significant. Age, education, gender, and geographic variables are also significant. Recommendations of effective and affordable policies directed to specific socioeconomic and demographic groups are derived.

Resumen

La transformación digital supone profundos cambios, económicos, sociales y políticos. Las personas, organizaciones y gobiernos funcionan, interactúan y tratan de prosperar en entornos cada vez más dinámicos y complejos donde las tecnologías de la información y la comunicación adquieren un creciente protagonismo.

Universidades, empresa privada, gobiernos e instituciones supranacionales son conscientes de que la economía digital implica tanto amenazas como oportunidades, y que el diseño y desarrollo de políticas y estrategias ha de estar basado en el análisis riguroso y continuo del fenómeno en todas sus dimensiones: técnica, política, económica y social. La Unión Europea, en su interés por alcanzar un Mercado Único Digital, fundamenta el diseño y ejecución de sus estrategias en un marco específico de estudios y encuestas sobre la evolución e impacto de la economía digital en los individuos y empresas.

Los datos utilizados en esta investigación provienen de la Encuesta sobre Equipamiento y Uso de Tecnologías de Información y comunicación en los Hogares en España, realizada anualmente desde el 2002 por el Instituto Nacional de Estadística (INE), siguiendo las recomendaciones metodológicas y los modelos de cuestionarios de la Oficina de Estadística de la Unión Europea (EUROSTAT). Esta encuesta es representativa del territorio nacional y comparable en países de la Unión Europea y otros ámbitos internacionales.

En los trabajos de investigación presentados en esta tesis se han especificado y estimado varios Modelos Cuantitativos de Adopción y Uso de Tecnologías de la Información y Comunicación, siendo el principal objeto de estudio el comercio electrónico y sus determinantes de adopción y uso en España. El objetivo principal de esta tesis es el de proponer modelos cuantitativos que permitan identificar los determinantes individuales de adopción y uso de servicios digitales y TIC (tecnologías de la información y comunicación)

para fines particulares. Los resultados y conclusiones de la investigación realizada permiten sugerir recomendaciones de medidas específicas para fomentar la adopción de servicios TIC por parte de los individuos.

Esta memoria está compuesta por cuatro capítulos. El primero es un capítulo introductorio y los siguientes corresponden a tres trabajos de análisis empírico, en los que se utiliza como base el marco teórico neoclásico estándar de maximización de la utilidad.

El primer capítulo consiste en una introducción general sobre los desafíos que presenta la evolución de la economía digital y cómo están respondiendo los distintos actores. Se detallan los pilares sobre los que la Comisión Europea cimienta su estrategia para alcanzar un Mercado Único Digital y cómo el desarrollo del comercio electrónico juega un rol prioritario.

En el segundo capítulo se analizan los factores determinantes de la decisión individual de realizar comercio electrónico transfronterizo en España. Utilizando técnicas de regresión logística, se analiza la encuesta de 2016 sobre Equipamiento y Uso de Tecnologías de Información y Comunicación en los Hogares en España. Los resultados sugieren una relación positiva entre adopción de comercio electrónico transfronterizo y los siguientes factores: ser hombre, mayor nivel de educación, habilidades informáticas y relacionadas con Internet, frecuencia de consulta de opiniones de otros usuarios y tener nacionalidad extranjera. En las conclusiones se destaca la importancia de la educación, las habilidades digitales, la confianza online y el uso de información sobre opiniones y recomendaciones, proponiendo medidas relacionadas para promover el uso del comercio electrónico transfronterizo.

El tercer capítulo se centra en el estudio de la adopción individual del comercio electrónico en España, analizando los principales determinantes socioeconómicos y demográficos. El análisis se lleva a cabo sobre un pool de datos de 174.776 observaciones, correspondiente a las encuestas de 2008 a 2016 sobre Equipamiento y Uso de Tecnologías de la Información y Comunicación en los Hogares en España. Se han estimado los siguientes modelos: Modelo Lineal de Probabilidad, el Modelo de Regresión Logística y el Modelo de Corrección de Selección de Heckman. Los resultados indican que la adopción del comercio electrónico está relacionada positivamente con ser hombre, nivel de educación, ingresos,

habilidades digitales, ser español y estar en situación laboral de empleado; mientras que variables como ser mujer o la edad están relacionados negativamente. La interacción incluida en el modelo sugiere que niveles altos de habilidades digitales contrarrestan los efectos negativos de algunos grupos de edad. Para fomentar la adopción del comercio electrónico se proponen recomendaciones específicas de acciones dirigidas al lado de la demanda, así como iniciativas gubernamentales y estrategias por parte del sector privado.

En el cuarto y último capítulo se analiza la adopción del comercio electrónico y la brecha digital en España. Para ello se utiliza un panel de datos sobre individuos, con 133.420 observaciones para el periodo 2008-2016. Este trabajo mejora sustancialmente otros existentes en la literatura internacional, debido a las características de la base de datos, la inclusión de variables previamente omitidas, el modelo económico y a la variedad de técnicas de análisis de datos de panel empleadas. Se mide la evolución de las brechas digitales a lo largo del tiempo y se cuantifica el impacto de las variables socioeconómicas y demográficas en la adopción del comercio electrónico. Se obtienen modelos con alto poder explicativo, en los que las variables renta y habilidades digitales son altamente significativas. Edad, educación, género y variables geográficas también son significativas. Se derivan recomendaciones de políticas efectivas y asequibles dirigidas a grupos socioeconómicos y demográficos específicos.

"Everyone takes the limits of his own vision for the limits of the world."

—Arthur Schopenhauer

"Most long-range forecasts of what is technically feasible in future time periods dramatically underestimate the power of future developments because they are based on what I call the "intuitive linear" view of history rather than the 'historical exponential' view."

—Ray Kurzweil

CHAPTER 1:
General introduction

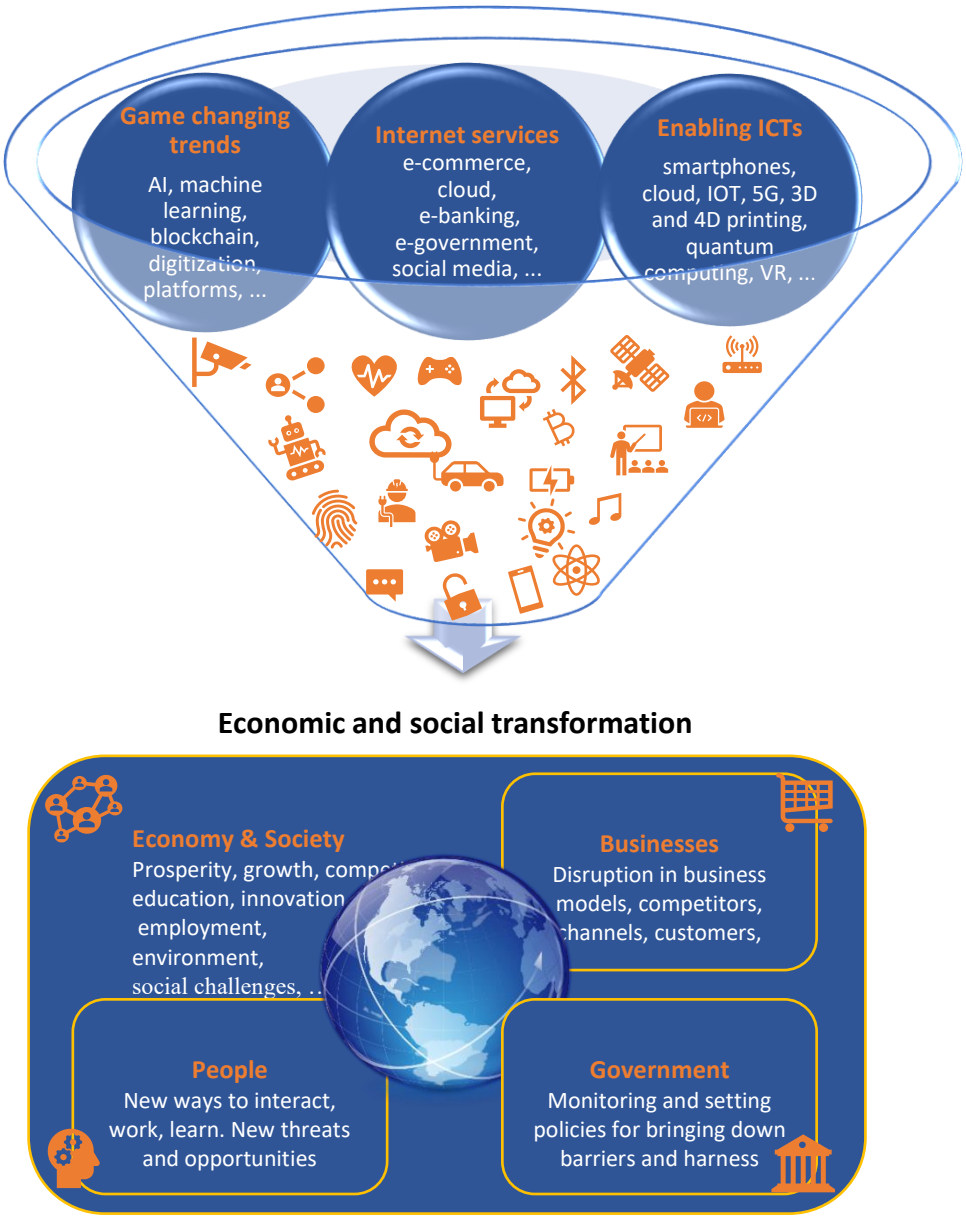
Digital economy and society

General-purpose technologies[3] have been the main engine of social development in human history. Overcoming the physical limitations of humans and domesticated animals with the introduction of steam power started a profound transformation of humanity. And now, thanks to the exponential increasing of processing power, communication speed, and storage capacity, that are the technological ground of today's digital transformation, we are experiencing the time when boundaries imposed by human brainpower are getting conquered with the help of Information and Communication Technologies (ICT) through digital devices, networks, and Internet-based services and applications (Brynjolfsson & McAfee, 2014).

A wide range of research is taking place to understand the digital economy, exploring how ICT is transforming the world and how large is its potential to continue being a leading global force of change. Technology is embedded in almost everything we do or relate with, affecting how we experience life, study, work, do businesses, communicate with others and spend our spare time. This pervasiveness influences the economy and society through complex dynamics that are challenging to be analysed, managed and ruled. There are, at least, two broad general goals at analysing this digital revolution from a socioeconomic perspective. The first one is to make sure we are harnessing the opportunities for individuals and society, and the other one is to understand how technology may raise inequalities and how this challenge can be tackled.

[3] Bresnahan, T.F., Trajtenberg, M. (1996) argue that the main features of general-purpose technologies should be pervasiveness, improvement and innovation spawning.

Figure 1:
Digital economy and society



Source: Own elaboration

ICT's evolving role in the economy is related not only with telecommunications and electronic equipment but with all the digital and online services that the Internet and related technologies enable to be developed. E-commerce, e-banking, e-government, e-learning, information search, social networks, among others, are just some of the visible pieces of the puzzle. Economics of ICT needs to go beyond the hardware and software part, widening

the scope of analysis to the implications of these new digital technologies for individuals, businesses, society and for the future of humanity (Bostrom, 2009).

Exponential advances in computation power, networks, cloud technologies, artificial intelligence (AI), robotics, advanced analytics, and digitization are fuelling a wide variety of trends and transformations that pose pressing issues on the economy and society. Economic growth, productivity, inequality, data and information privacy, digital business models, analytics and experimentation in social sciences are among the key areas that business, government, and academia need to focus on. *Figure 1* depicts how different elements produced by enabling technologies, digital services and game-changing trends pass through a wide-mouth funnel to the economy and society, evidencing the urgency of adaptation of individuals, businesses, and policymakers.

Digital Single Market

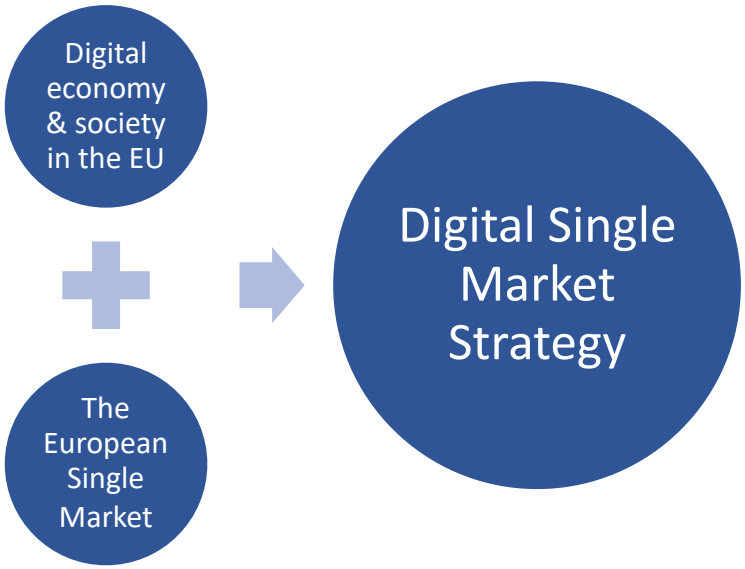
Academia, private companies, governments, international and supranational organizations have a long record demonstrating their interest and commitment in shaping the future of the digital economy and society. Economic policies, business strategies, and coordinated initiatives need to be grounded in rigorous analysis, which requires continuous monitoring of ICT and its implications across different dimensions (Manyika et al., 2016).

World-leading universities have specific units of research and teaching specialized in the study of digital technologies and their implications. That is the case, for instance, of University of Oxford, University of Cambridge, Massachusetts Institute of Technology and Harvard University. The following are some of the most relevant academic research initiatives and networks: Oxford Internet Institute, Cambridge Digital Humanities, Cambridge Big Data, Cambridge Trust & Technology Initiative, MIT Initiative on the Digital Economy and Digital Initiative at Harvard Business School.

In the private sector, the tech behemoths of the Internet, which are responsible for transforming entire sectors and represent serious challenges for policymakers and

regulation, are important actors at non-profit and socially responsible initiatives. Companies like Google, Facebook, Amazon, and Microsoft, have direct or coordinated participation through their foundations and commissioned programs to study and promote specific actions aiming to have a positive impact on society through technology. Google.org (2019), for instance, with a 1 billion dollars commitment to contribute with 1 million employee volunteer hours, is focused on closing educations gap, prepare people for the changing nature of work, granting non-profits organizations to use data science and find innovative ways to advance in social inclusion.

Figure 2:
EU Digital Single Market strategy



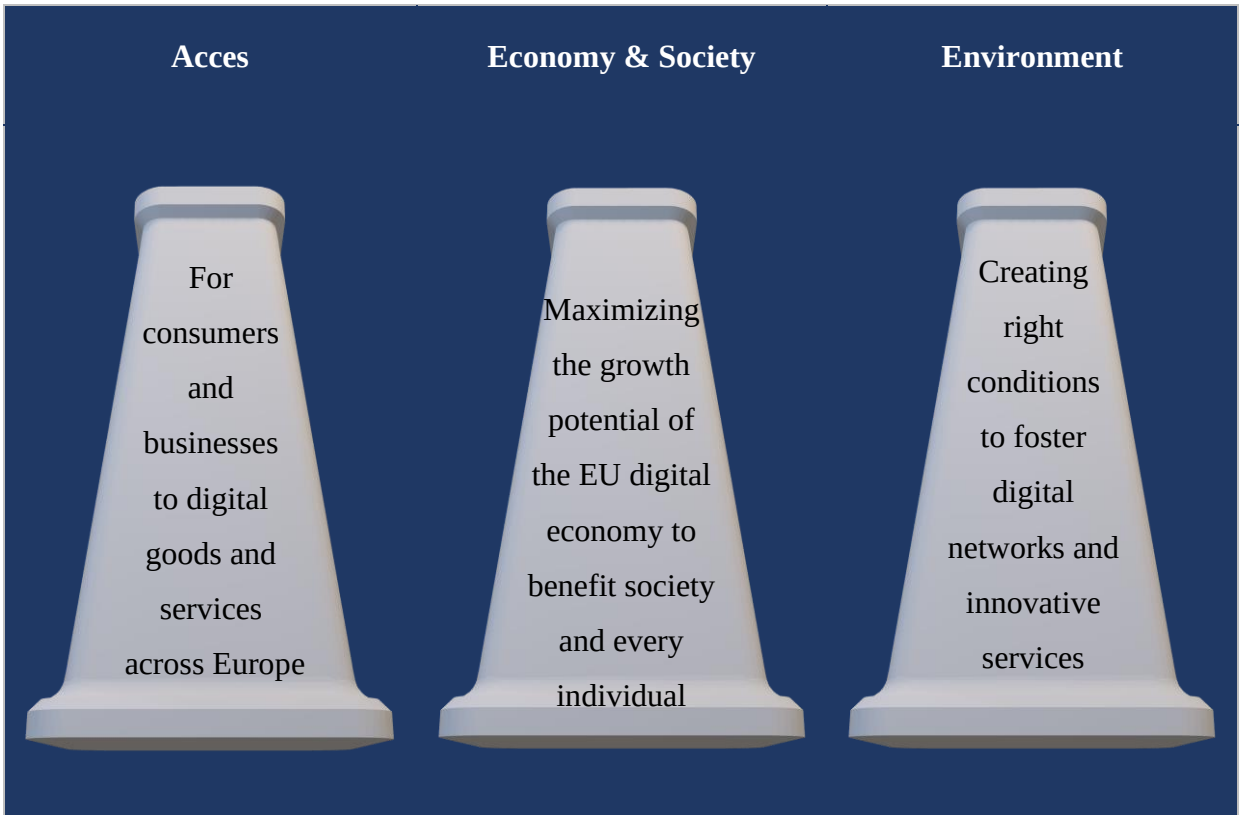
Source: Own elaboration based on European Commission, 2015b.

Governments, supranational, and International organizations soon recognized the power of emerging technologies to change economies at astonishing speed and scale. That is the case in the European Union (EU). The EU's Single Market, a territory without neither internal borders nor regulatory barriers to the free movement of goods and services, has given way to the Digital Single Marketing Strategy (DSM). The goal of the European Commission with this strategy is to harness the opportunities of digital technologies bringing down

online barriers that disincentive people and businesses to get full advantage of digital development (European Commission, 2015b). *Figure 2*.

Having one market for goods but 28 national digital markets give a clear picture of the high cost of non-Europe DSM. "Mapping the Cost of Non-Europe", a European Parliament (2017) Study, estimates in 415€ billion yearly the increase in terms of EU GDP (2016) if all barriers for a truly digital single market were eliminated. To develop the DSM across European countries the European Commission grounds its strategy on three pillars: Access, Environment, and Economy and Society (European Commission, 2015c). *Figure 3*.

Figure 3:
The three pillars of the European Digital Single Market



Source: Own elaboration based on European Comission, 2015b.

The first pillar, related to the demand side, is oriented to ensure better access to digital goods and services for individuals and businesses. The actions needed to build up this pillar are related with: easing e-commerce, review Consumer Protection Regulation, affordable

cross-border e-commerce, end unjustified geo-blocking, modern European copyright, review of the Satellite and Cable Directive, simpler VAT administrative regime and common VAT threshold, and antitrust competition inquiry into the e-commerce sector.

Focusing on the supply side, a better business environment for digital networks and services is the second pillar. The specific actions have to do with review, update and create, if needed, rules and directives on the following topics: EU telecoms, promotion of Internet connectivity in local communities and public spaces, plan on 5G communication network, audio-visual media framework, public-private partnerships for increased cybersecurity, high common level of network and information security, e-Privacy Directive, assessment of online platforms, and proposals to tackle illegal content.

The third pillar seeks to maximize the growth potential of the digital economy and its benefits for society. Its main targeted actions are a European Cloud, free flow of non-personal data, digitizing European Industry, European interoperability framework, a new e-Government action plan, and a review of Public Sector Information directive.

Monitoring the digital economy and society

Policy developments require a framework of data sources to monitor European digital economy evolution, information and communication technologies, and their impact. The major sources of data have been: 1) The Eurostat Survey on ICT Usage in Households and by Individuals; 2) The Eurostat Survey on ICT usage and e-commerce in Enterprises; 3) Eurostat secondary statistics indicators on the ICT sector and ICT specialists; 4) Data supplied by the National Regulatory Authorities for the telecommunications market; 5) Periodical studies commissioned by the European Commission when official statistics are not available; and 6) Ad hoc surveys on specific issues (European Commission, 2015a).

The data and information gathered needs to be policy-driven and contextualized. The design of questionnaires and topics to cover have to be constantly adapted to the current needs, while always monitoring key themes as: the ICT sector as a supplier of general-

purpose technologies, broadband as key infrastructure, the digitization of the economy, e-commerce at the core of the Digital Single Market, trust, security, data protection and privacy, Internet usage by citizens, consumption of online content, online public services, digital skills and ICT occupations (European Commission, 2015a).

The Survey on ICT Usage in Households and by Individual is the data source used in this research to study equipment availability, adoption and use of information and communication technologies at a dwelling, household, and individual level. *Table 1* shows the topics that have been incorporated into the analysis carried out within the present research project.

Table 1:
Analysed topics of the Spanish survey on ICT usage in households and by individual 2008-2017

☑	ICT usage in households and by individuals Survey
☑	Connection to the Internet and computer use
☐	Internet use by children
☑	Internet use
☑	Digital skills
☑	E-commerce
☑	E-government
☑	ICT trust, security, and privacy
☐	ICT usage at work
☑	Regional ICT statistics
Note: ☑ Checked if the data has been used in this research.	

The evolving situation of ICT requires that Eurostat's model of survey questionnaires changes every year. *Table 2* shows the different sections that Spanish surveys on ICT usage in households and by individuals have had from 2008 to 2018. A blank cell means that in a given year that topic has not been part of the survey. Each row gives information about the number of questions related to a specific topic for each year. Total questions are not equal to total variables. Several questions have several response options, generating different variables from one single question.

Table 2:
Sections of the questionnaires. Spanish surveys on ICT usage in household and by individuals 2008-2018

Topics	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
Identification and selection	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Household's ICT equipment	13	13	13	3	3	3	2	2	2	2	2
Household's Internet access	4	4	4	4	3	3	4	3	3	3	2
Internet use by children 10-15	8	8	8	8	5	5	5	5	5	5	5
Use of mobile and computer	6	6	5	5	6	4	5	4		4	
Use of the Internet	9	11	10	10	7	10	7	11	12	10	11
Use of advanced services	9										
Mobile Internet and connectivity					11						
Use of cloud services							11				
Use of e-Government				3	3	7	3	3	3	3	3
Use of shared economy services										2	4
Use of e-commerce	**5**	**13**	**6**	**7**	**6**	**6**	**6**	**9**	**11**	**12**	**7**
Electronic ID		3	2				3				
Internet Security			7				2	9			
Privacy, security, and trust on the Internet									9	2	7
Computer skills			1	6		3		2	2	2	2
Use of ICT at work											13
Socio-demographic characteristics	6	9	7	7	7	9	9	7	9	9	11
Total questions	**60**	**67**	**63**	**53**	**51**	**50**	**57**	**55**	**56**	**54**	**67**
Total Variables	**296**	**324**	**308**	**305**	**235**	**258**	**260**	**256**	**257**	**236**	**236**
Total Common Variables	**123**	**130**	**132**	**141**	**143**	**150**	**154**	**157**	**165**	**170**	**236**

Individual access and use of new technologies in Spain

Not all the barriers EU citizens face when using online tools and services are being taken down by policies derived from DSM strategy, as can be seen in the Europe's Digital Progress Report (DPR) 2017 and the Digital Economy and Society Index (DESI) and its country profiles about progress in digitalization (European Commission, 2018a; European Commission, 2018b). Many of them may remain unidentified and need to be analysed at an individual level. The research presented in this thesis analyses the penetration rates of current Internet services and models the adoption by individuals, contributing to enriching the pathways of economic analysis of adoption and usage of the forthcoming new digital services.

The research presented in this thesis is part of a broader project that aims to study individual access and use of new technologies, using microdata from the Spanish Survey on Equipment and Use of Information and Communication Technologies in Households. *Table 3* shows the topics already covered (blue background) in papers that are published or have been sent to publish; and (yellow background) those that are in the development phase.

Centering attention on cross-border e-commerce, the European Parliament study estimates that, if all barriers were eliminated, the EU value-added could reach 204 billion euros yearly.

As the DSM strategy is one of the Juncker Commission's ten political priorities[4], e-commerce is one of the cornerstones of the DSM strategy. Several measures have been already taken and others are under review. The addressed topics include ending unjustified geoblocking in the EU, making cross-border parcel deliveries cheaper, protecting the rights

[4]Prior to his election as President of the European Commission, Jean-Claude Juncker set out ten political priorities for his 2014-2019 mandate. 1) A new boost for jobs, growth and investment; 2) A connected digital single market; 3) A resilient energy union with a forward-looking climate change policy; 4) A deeper and fairer internal market with a strengthened industrial base; 5) A deeper and fairer economic and monetary union; 6) A balanced and progressive trade policy to harness globalization; 7) An area of justice and fundamental rights based on mutual trust; 8) Towards a new policy migration; 9) A stronger global actor; and, 10) A Union of democratic change.

of online consumers, and facilitating access to audio-visual services (European Parliament, 2019).

Table 3:
Type of analysis carried out with the ICT usage in household and by individuals survey

Cross-sectional data	Pooled data	Panel data	Models
E-commerce Adoption (Private goods and services) S, L	*E-commerce Adoption (Private goods and services) S, LPM, L, H (**Chapter 3**)	*E-Commerce and Digital Divide (Private goods and services) S, D, L, H (**Chapter 4**)	**S**tatic **D**ynamic **LPM**[5] **L**ogistic **T**obit **H**eckman …
E-banking Adoption (Private financial) S, L		E-commerce Expenditure (Private goods and services) S, D, T	
E-government Adoption (Public services) S, L		E-tourism (Private goods and services) S, L	
*Cross-border e-commerce Adoption (Private goods and services) S, L (**Chapter 2**)		Gender and digital divide (Private goods and services) S, D, L	

Identifying barriers and the individual determinants of adoption of e-commerce is not a straightforward task and it can hardly be approximated by analysing descriptive statistics of aggregated data only. To figure out the drivers and barriers of online shopping adoption for private use a more sophisticated approach is needed. First, using an economic model that allows to include the socio-economic and demographic characteristics of the

[5] Linear Probability Model

individual, interacting perhaps with characteristics of the available options. Then, specifying and estimating models of individual adoption and usage of e-commerce for private use, where having limited dependent variables, relevant categorical explanatory variables are being controlled.

E-commerce adoption in Spain

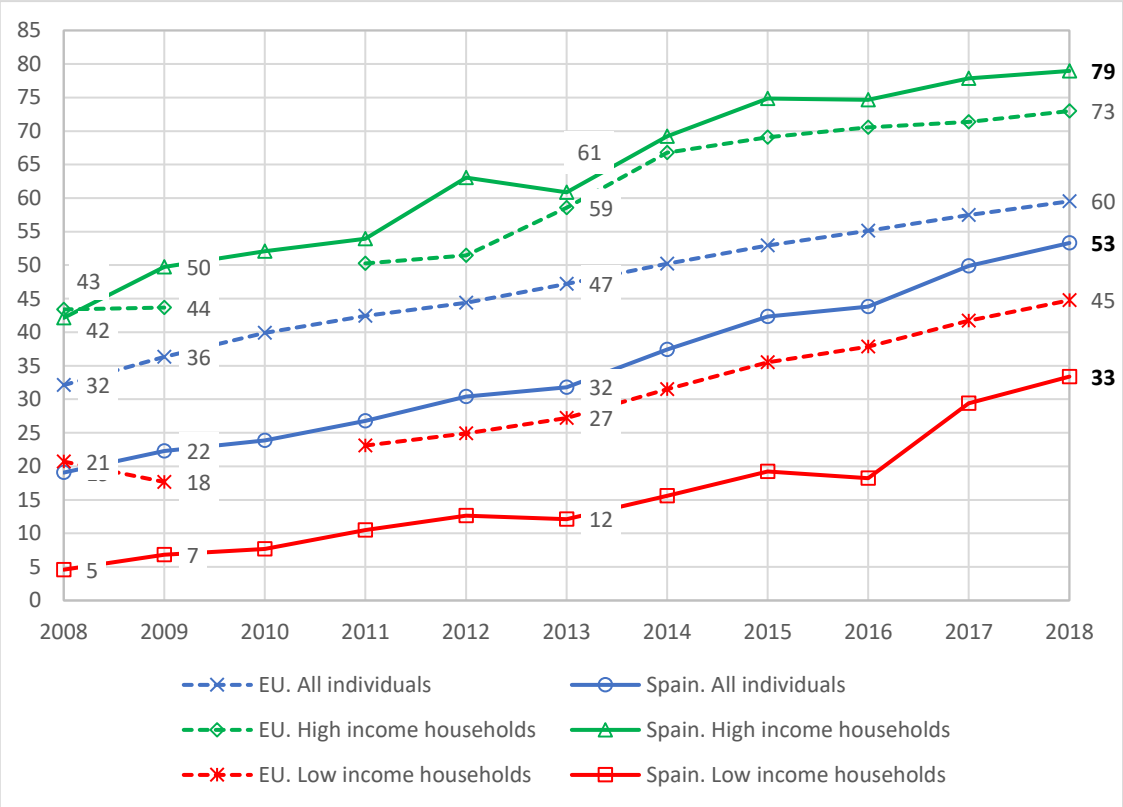
According to official data from Eurostat and the Spanish National Statistical Institute (INE), 53% of individuals aged 16-74 in Spain has bought online for private use during the last 12 months, approaching the EU average of 60%. The observed trend from 2008 to 2018 seems to indicate that Spain is moving towards closing the gap. However, when we compare the average e-commerce adoption trajectories with the adoption trajectories across different demographic groups the situation is not clear. *Figure 4*.

Levels of adoption of individuals with high income in Spain have remained above the average of their Europeans counterparts, reaching 79% and 73% respectively in 2018. While those individuals in Spain with the lowest income are far from getting close to the European average e-commerce adoption of those with low income, 33% and 45% respectively for 2018.

Income levels are just an example of the possible and necessary breakdowns to learn from the available data in order to carry out a proper analysis for monitoring the evolution of the Spanish and European Digital Economy and Society.

The following chapters present three empirical studies where the individual consumers adoption of e-commerce in Spain is modelled, analysing signs, significance and the magnitude of effects of the socioeconomic and demographic determinants. Policy recommendations to bridge the digital divide and to foster the individual adoption of e-commerce and other ICT and digital services are derived.

Figure 4:
E-commerce adoption in Spain. 2008-2018[6]



Source: Own elaboration based on INE, 2018. Survey on Equipment and Use of Information and Communication Technologies in Households.

[6] **Source:** Eurostat, Table isoc_ec_ibuy: Internet purchases by individuals. **Definition:** Individuals carrying out this activity over the internet in the last 12 months, for private use. **Breakdown**: All Individuals (aged 16-74). **Unit of measure**: Percentage of individuals.

References

Bostrom, N. (2009). The Future of Humanity. 29. New Waves in Philosophy of Technology. Palgrave Macmillan UK, 186-215.

Bresnahan, T. F., & Trajtenberg, M. (1995). General purpose technologies 'Engines of growth'?. Journal of econometrics, 65(1), 83-108.

Brynjolfsson, E., & McAfee, A. (2014). The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies. New York: WW Norton & Co.

European Commission. (2015a). A new Monitoring Framework for the Digital Economy and Society [Text]. Retrieved from https://ec.europa.eu/digital-single-market/en/news/new-monitoring-framework-digital-economy-and-society

European Commission. (2015b). Why we need a Digital Single Market. Text. Retrieved from https://ec.europa.eu/commission/publications/why-we-need-digital-single-market_en

European Commission. (2015c, May 6). A Digital Single Market Strategy for Europe - COM(2015) 192 final [Text]. Retrieved May 20, 2019, from Digital Single Market - European Commission website: https://ec.europa.eu/digital-single-market/en/news/digital-single-market-strategy-europe-com2015-192-final

European Commission. (2018a). The Digital Economy and Society Index (DESI). Retrieved November 11, 2018, from https://ec.europa.eu/digital-single-market/en/desi

European Commission. (2018b). 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 Parliament. (2017). Mapping the Cost of Non-Europe, 2014-19 - Fourth edition - Think Tank. Retrieved from European Parliamentary Research Service website: http://www.europarl.europa.eu/thinktank/en/document.html?reference=EPRS_STU(2017)603239

European Parliament. (2019). The Juncker Commission's ten priorities: An end-of-term assessment - Think Tank. Retrieved May 22, 2019, from http://www.europarl.europa.eu/thinktank/es/document.html?reference=EPRS_IDA%282019%29637943

Google.org. (2019). Our $1 billion commitment to create more opportunity for everyone. Retrieved May 5, 2019, from Google.org website: https://www.google.org/billion-commitment-to-create-more-opportunity/

INE (2018). "Survey on Equipment and Use of Information and Communication Technologies in Households". https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176741&menu=ultiDatos&idp=1254735976608

Manyika, J., Lund, S., Bughin, J., Woetzel, J., Stamenov, K., & Dhingra, D. (2016). Digital globalization: The new era of global flows. McKinsey Global Institute.

CHAPTER 2:
Drivers and barriers to cross-border e-commerce: evidence from Spanish individual behavior[7]

[7] Published in Telecommunications Policy, 42 (6), 464-473; with Pérez-Amaral, T., Garín-Muñoz T., Herguera García, I., & López, R.

Abstract

This paper explores the determinants of the individual´s decision to perform cross-border e-commerce (CBeC). The European Union (EU) is especially interested in the promotion of CBeC because it is an important tool in its strategy to achieve the Digital Single Market in Europe. In this paper, official data is used from a representative survey of 16,209 individuals on ICT usage by households and individuals that was carried out in Spain by the National Institute of Statistics (INE) for the year 2016. Using a standard neoclassical utility maximization framework, and logistic regression techniques, the results show that being a male is positively related to the probability of practicing CBeC. Education is positively and significantly related to the probability of being involved in CBeC with EU countries. Computer and Internet Skills are significant and positive factors in explaining CBeC (either with EU countries or with the rest of the world). The variable "how often the consumer sees other customer reviews before buying online", has a positive effect. Foreign nationality also increases the likelihood of using CBeC. To promote CBeC in Spain measures towards developing digital skills, Internet trust and use of online information reviews of goods and services are discussed.

Keywords: cross-border e-commerce; domestic e-commerce; digital skills; customer reviews; online opinion seeker; logistic model.

JEL classification: C25; D11; O33

Acknowledgments: Paper presented at the 28th ITS European Conference, Passau-Germany, 2017. This research has been 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.

Determinantes del comercio electrónico transfronterizo: evidencia del comportamiento individual en España [8]

Resumen

Este trabajo explora los determinantes de la decisión individual de realizar comercio electrónico transfronterizo (CBeC). La Unión Europea (EU) está especialmente interesada en la promoción del CBeC debido a que es una herramienta indispensable en su estrategia para alcanzar el Mercado Único Digital en Europa. En este trabajo se han utilizado datos oficiales que provienen de una encuesta representativa de 16.209 individuos, sobre el uso de tecnologías de la información y comunicación en hogares y por individuos, llevada a cabo por el Instituto Nacional de Estadística (INE) para el año 2016. Utilizando el marco neoclásico de maximización de la utilidad y técnicas de regresión logística, los resultados muestran que ser hombre está relacionado positivamente con la probabilidad de realizar CBeC con países de la Unión Europea. Las habilidades informáticas y aquellas relacionadas con el uso de Internet son factores significativos y positivos para explicar CBeC (tanto con países de la Unión Europea, como con países del resto del mundo). La variable "con qué frecuencia un consumidor ve opiniones de otros consumidores antes de comprar", tiene un efecto positivo. Tener nacionalidad extranjera también incrementa la probabilidad de realizar CBeC. Para promover el CBeC en España se proponen medidas dirigidas a desarrollar las habilidades digitales de los individuos, confianza en Internet y el uso de información de opiniones sobre bienes y servicios.

[8] Published in Telecommunications Policy, 42 (6), 464-473; with Pérez-Amaral, T., Garín-Muñoz T., Herguera García, I., & López, R.

Introduction

Being able to shop for goods and services online offers customers several benefits, many of which did not exist until recently. Some of them are: access to an increasing variety of products and sellers from all over the world, a decrease in the asymmetry of information, a reduction in search costs, an ability for comparison, the fostering of competition between sellers, it saves time, it allows better spending planning, it gives voice to the individual consumer through the possibility of public customer reviews and social media sharing, and as an online purchase is a computer-mediated transaction it improves its traceability and makes more efficient contracts possible. In the case of cross-border e-commerce, it is worth mentioning that different studies conclude that consumers also benefit from online commerce because of the substantial reduction in distance-related trade costs (Martens, 2013).

However, there are still barriers that deter consumers from taking full advantage of e-commerce. Besides limitations related to Internet penetration and equipment of Information and Communication Technologies (ICT), others emerge from internet users' concerns and perceptions. Some of the main barriers for domestic and cross-border e-commerce are highlighted by Gefen (2000), Gomez-Herrera *et al*. (2014), Cardona *et al*. (2015) and PayPal (2016): little familiarity and trust in the vendor, crossing language barriers, need of a secure way to pay, cost-efficiency of parcel delivery, and costs shown in different currency, among others. Governments and the main stakeholders on the supply side are aware of the concerns of the consumers and the benefits that could come from dealing with them.

Given the importance of e-commerce within the Digital Single Market (DSM) strategy, the EU has engaged in opening digital opportunities for people and businesses, particularly in

boosting cross-border e-commerce through a wide range of initiatives. To deal with these existing barriers, the EU consults with several expert groups, reports, and studies to ground their strategies, decisions, and proposals for the future. The European Commission (EC) wants to encourage a wider choice of goods and services online, better consumer protection, more transparent and affordable parcel delivery across borders, privacy and personal data protection, as well as improving the Value Added Tax (VAT) environment for e-commerce businesses (European Commission, 2017).

At the country level, Spain, like other countries in the EU, has recently improved its Internet penetration at an affordable price. This a necessary condition, but not a sufficient one to generalize its use (Garín Muñoz *et al*., 2017). In Spain only 25% of individuals used the Internet to make an online purchase from other countries (EU or other countries from the rest of the world) in the last year.

In relation to the supply side, most of the attention goes to the biggest players worldwide, Amazon and Alibaba, which are not only leading e-commerce in their own countries and abroad, but transforming and influencing entire industries, businesses and technologies such as retail, logistics and transportation, manufacturing, cloud computing, entertainment, digital advertising, social media, search engine business, Internet of things (IoT), automation, robotics, artificial intelligence, payments, video streaming, and physical shops. As *The Economist* (2017) states in its special report "E-commerce: The New Bazaar", the two giants are paving the way to become true conglomerates.

The decline of traditional retailing in the U.S. and in some of Europe's biggest markets shows this industry as one of the most affected. The share of e-commerce in the total retail trade was 8.5% worldwide and 10% in the U.S. in 2016, still far from 18% of South Korea. Due to the prospects towards automation, robotization and artificial intelligence, it is clear that the destruction of jobs—which are difficult to compensate by the creation of new ones in e-commerce related activities—is an imminent threat.

Economies of scale, patient shareholders (in the case of Amazon) and well-identified core investments, are some of the main ingredients to become an undisputed leader of e-commerce. Although it seems difficult to compete against the giants, what is happening in China could be an example for Western economies to follow. To challenge Alibaba's

dominance, its main competitors, JD and Tencent have joined forces (T*he Economist*, 2017).

In this paper, we use representative data on 16,209 individuals to study the determinants driving an individual to perform cross-border e-commerce (CBeC) in Spain. Besides the general CBeC (i.e. when an individual buys online from a country other than Spain), we distinguish two more types of CBeC: whether the foreign vendor belongs to the European Union, CBeC_EU; or to the Rest of the World (outside the EU), CBeC_RW. Notice that for a particular individual these are non-exclusive types of CBeC: she may buy from vendors in the EU, from vendors in the Rest of the World, or from both. Three separate logistic models are built to assess the determinants for each of the three types: CBeC, CBeC_EU and CBeC_RW.

The rest of the paper is structured as follows. Section 2 presents a literature review. Section 3 describes the current situation of CBeC in Spain and other countries within the EU. Section 4 describes the main characteristics of the data to be used. In Section 5 the three logistic models are presented, showing how the selected variables influence the probability of buying online outside Spain. Section 6 concludes with the main conclusions, the policy recommendations, the caveats, and further research.

Literature review

Previous research

The EC has devoted significant effort to fostering CBeC, while it has received relatively little attention in academic literature. Most of the EC literature and policy measures on cross border e-commerce deal with the supply side, while here, we concentrate on the demand side.

Correa *et al*. (2015) use ICT-H data up to 2014 to estimate a model of e-commerce in Spain using a logit specification. Determinants of individual adoption of e-commerce in Spain are studied by Garín-Muñoz and Perez-Amaral (2011), using the 2007 ICT-H data. A

multinomial logit specification (MNL) is used to estimate relevant effects over the dependent variable (the type of participation in online shopping in the last three months).

Previous studies (Cardona *et al*., 2015; Falk & Hagsten, 2015; Manyika *et al*., 2016; Martens, 2013) highlight the impact that e-commerce and cross-border e-commerce already have on the economy and their potential growth and future broader positive effects. The challenges and opportunities for both sides of the market (supply and demand), the increase of price competition, the gains in efficiency in the retail sector, the positive effects on production in other sectors, the benefits for individuals and household consumers, labour productivity and GDP growth, are just some of the implications that support the strong commitment of the EC to remain actively developing and implementing policies to foster the Digital Single Market.

European Commission reports and guidelines

The European Commission in its European Digital Single Market strategy (DSM) states the intention to reduce barriers that firms and consumers face when entering into a cross-border contract in the EU (European Commission, 2017). In its DSM strategy, the following measures have been adopted during 2016 (European Commission, 2016):

- Eliminating geo-blocking to facilitate the purchasing of any goods or services across the EU without any interference or restriction on any consumer. The purpose is to eliminate any undue discrimination against consumers in relation to the prices or contracting conditions because of the place of residence or nationality of the end-user. The geo-blocking involves the reform of several Directives since it affects not only goods but services and contents as well.

- Making parcel delivery more affordable is another objective of the EU (European Commission, 2012). Specifically, the EU wants to make prices and conditions for contracting parcel delivery more transparent, so that prices reflect real underlying differences in costs of providing the service. In a survey conducted by the European Commission, it was found that a substantial proportion of internet users did not finalize an online contract because of fears of high delivery costs. As the next step within a long process of studies, consultations, proposals, debates, and negotiations,

in October 2017 the Transport and Tourism Committee of the European Parliament approved draft rules which aim to improve quality and pricing of cross-border parcel delivery services (European Parliament, 2017).

- The European Commission promotes harmonized consumer protection across the EU and simplification in the burden of different VAT rates among countries. Suppliers of goods and services that need simplified rules when trans-border trade occurs inside the EU.

In addition, the EC has published guidelines on the implementation of the Directive on unfair commercial practices (European Parliament and Council, 2005), to implement EU-wide common principles so that consumers can trust any EU provider with the same ease as a domestic online supplier. National Consumer Protection Authorities shall coordinate among the EU's practices and monitor any after-sales conditions attached to contracts provided trans-border that may involve undue discrimination.

Europe's Digital Progress Report (DPR) 2017 combines information from the Digital Economy and Society Index (DESI) and qualitative information about the country-specific policies, to build country profiles about their progress in digitization: connectivity, human capital, use of the internet, integration of digital technology and digital public services. Digital Public Services is where Spain has obtained its greatest progress, while lower levels of growth on digital skills and a weak demand weigh down its position in relation to the other member states (European Commission, 2017).

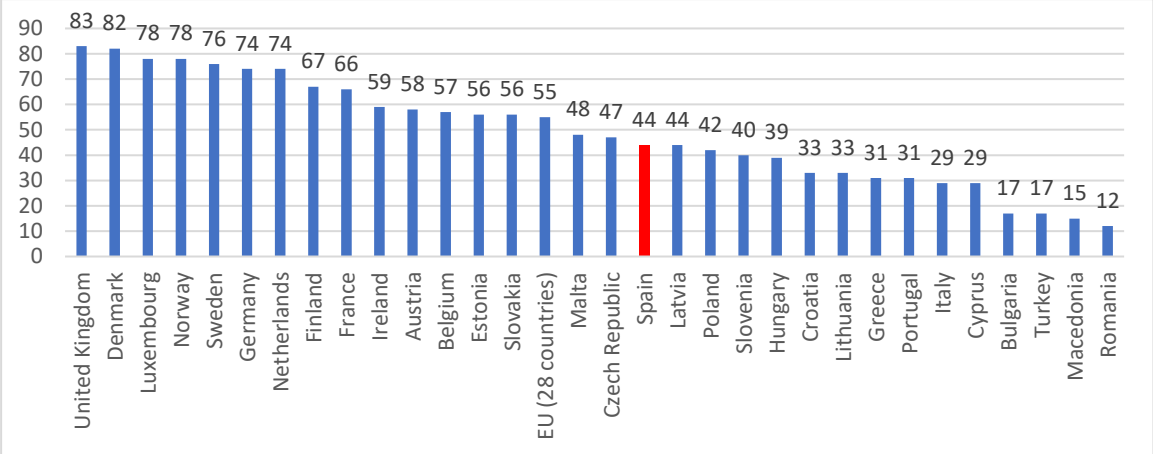
Based on the 2016 ICT-H data, ONTSI (Observatorio Nacional de Telecomunicaciones y de la Sociedad de la Información) presented a study for Spain. The research found that 82.7% of people accessed the Internet at some time and that almost 23 million people aged 16 to 74 connect to Internet every day. The highest digital capacities coincide with younger people and those with higher education while there is also a direct link between high digital capacities and the employed or student population. The lowest values of technological capacities correspond to older people, with low income and low education levels, and who are either pensioners or dedicated to housework (ONTSI, 2016).

E-commerce and cross-border transactions in Spain

Eurostat data on Information and Communication Technologies (ICT) usage in households and by individuals show that in 2008, 19% of the individuals aged 16 to 74 bought online in Spain, while nine years later the proportion rose to 44%. On the other hand, in 2008 only 5% of the individuals purchased online from another country in the EU. Nine years later the proportion has more than quadrupled. While internet penetration rose during the same period, online purchasing participation rose at a much faster rate.

In the EU-28 an average of 84% of the population (aged 16 to 74) had used the Internet in the last 12 months, and 55% had ordered goods or services via the Internet in 2016, as seen in *Figure 1*. In Spain the proportion of Internet users is 81% and 44% used Internet for e-commerce in the last year; a proportion that goes down to 35% of those who purchased online in the last 3 months is considered (Eurostat, 2017).

Figure 1:
Penetration rates of e-commerce in the EU-28 countries. 2016



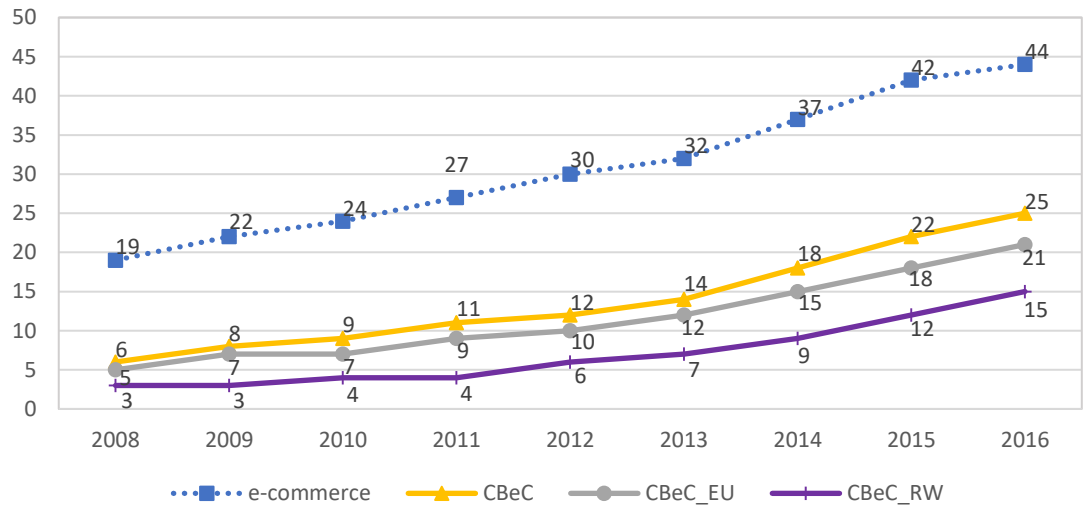
Source: Own elaboration based on data from Eurostat, 2017. Penetration rates as the percentage of people aged 16 to 74.

Figure 2 summarizes the evolution of e-commerce penetration rates in Spain among people aged 16 to 74. Distances between the compared categories are roughly maintained year

after year. The highest penetration is for CBeC (any country abroad, 25%), followed by CBeC_EU (21%), and CBeC_RW (15%)[9].

Figure 2 shows how the behavior of residents in Spain has evolved during the period 2008-2016 in relation to E-commerce. A significant improvement in all the categories is observed along the 9-year period. However, *Figure 3* shows that Spain is still below the leading EU countries in terms of transactions with other EU countries.

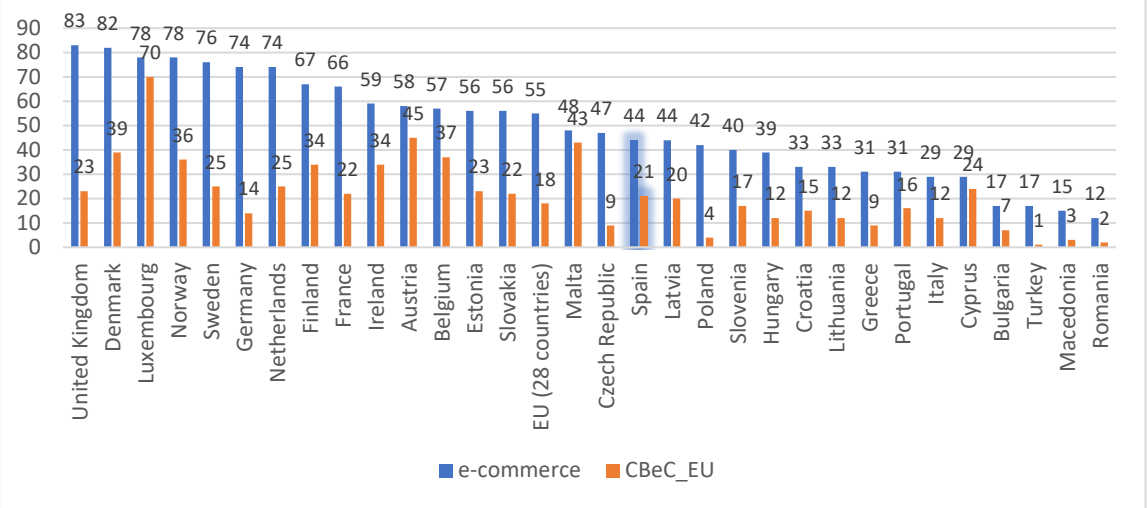
Figure 2:
Evolution of the penetration rates of E-commerce by the geographical location of the seller. 2008-2016



Source: Own elaboration based on data from Eurostat, 2017. Penetration rates as percentages of people aged 16 to 74.

[9] Question 40 of the survey asks each individual who has bought online products or services in the last twelve months if: (a) They have bought from national vendors, (b) They have bought from other countries of the European Union, (c) They have bought from vendors from the rest of the World, (d) The country of origin of the vendors is unknown. These four options are not exclusive of each other. An individual performs CBeC_EU when checking option (b); CBeC_RW when checking (c); and CBeC when checking (b) or (c).

Figure 3:
Penetration rates of e-commerce and cross-border e-commerce with other EU countries. 2016



Source: Own elaboration based on data from Eurostat, 2017. Penetration rates as the percentage of people aged 16 to 74.

The above information suggests that residents in Spain are more open than the average to make purchases in other EU countries, but there is still plenty of room to improve.

The ICT-H survey data

The data in this paper comes from the survey of 2016 on Equipment and Use of Information and Communication Technologies in Households (ICT-H) carried out by the Spanish National Statistical Institute (INE, 2016). Since the survey follows the methodological recommendation of the Statistical Office of the European Communities (Eurostat), it is comparable among EU member States which follow the same standards.

INE's annual continuous survey 2016 ICT-H is a household panel study. It has been carried out since 2002, but it is only homogeneous from 2006 onwards. The data are available online at www.ine.es (INE, 2016).

From an initial sample size of 23,887 households in 2016, there remain a total of 16,209 valid respondents after processing. They represent 16,029,823 dwellings and 34,389,822 people. The aim of the survey is focusing on people aged 10 (residents in family dwellings) and above and collecting data about ICT equipment and services at home.

The statistics in this paper are calculated using the respective weight for each individual. Applying the weight corrects a possible sample selection bias. The design weights are computed as the inverse of the inclusion probabilities and then scaled such that their sum equals the net sample size (Garín Muñoz *et al*., 2017). In doing so, inferences can be made more accurately than in similar works in which this procedure is not used (Pfeffermann, 1993).

Some previous works focused on CBeC (supply or demand side) using ad-hoc cross-section questionnaires, but were usually not representative, while our sample is representative by design.

For comparison and given that in order to be cataloged within any of the three types of CBeC it is necessary to pass the first stage (i.e., to buy online), Table 1 shows the penetration rates for eBuyers who purchased online during the last year, in addition to the types of CBeC that are the main objective of this paper.

The concern of whether or not an e-commerce buyer can identify where is she buying from, is partially mitigated by a report about consumer perceptions of CBeC in the EU (Cardona, Duch-Brown & Martens, 2015), based on data of a survey carried out in the EU-28 in 2015, they find that 77% of the respondents correctly assess whether a website is located domestically or not.

In our 2016 sample, as shown in *Table 1*, the last row of penetration rates by socio-economic and demographic characteristics suggest that there are differences between the distinct groups that are analyzed. 53.3% of people who used Internet within the last year bought online, while only 30.7% did so outside the borders, and 25.4% bought from other European countries.

Weighted data in *Table 1* shows that there is still a gender gap in the adoption of e-commerce where the greatest difference is observed for the three types of CBeC, where the distance in terms of penetration rates between males and females are 7.1 (CBeC), 7.2 (CBeC_EU), and 6.1 (CBeC_RW) respectively.

Penetration rates are higher for the younger groups and tend to decline for older groups. Notably, education is related to the penetration rates. For instance, the penetration rate of CBeC for the group which has a level of education with a master's or Ph.D. degree is more than 8 times higher than for the group whose level of education is primary or below.

In relation to household size, no discernible differences are observed between groups of different sizes, while the differences are bigger when it comes to the number of inhabitants of the population, where the higher the habitat size the higher the penetration rate. Penetration rates of diverse types of e-commerce are more similar among Spanish and non-Spanish nationals.

Being employed, being a student or having a higher income, as might be expected, is related to higher participation rates. It is remarkable that within occupation comparisons the highest penetration of CBeC goes for the group classified as a student, and the lowest penetration for all types of e-commerce users is for the group of housekeepers (representing a total of 2,772,032 people: 17,237 men and 2,754,795 women).

Table 1:
Penetration rates of e-commerce in Spain by categories (2016)

	e-commerce	CBeC	CBeC_EU	CBeC_RW
GENDER				
Male	55.2	34.2	28.9	21.5
Female	51.4	27.1	21.7	15.4
AGE				
16-24	58.7	37.9	27.5	23.5
25-34	63.5	37.7	33.2	24.1
35-44	60.1	35.2	29.4	21.0
45-54	49.2	27.2	23.2	16.1
55-64	40.3	20.0	16.1	10.7
65<	26.6	11.5	9.1	6.5
EDUCATION				
No studies/Primary	14.5	5.7	2.9	4.0
Secondary studies	45.5	24.7	18.9	15.2
Bachelor´s degree	67.4	39.3	34.0	23.5
Master or PhD	79.3	51.0	46.1	29.9
HABITAT				
> 500,000	58.6	35.7	30.7	21.6
100,000-500,000	54.8	30.6	24.4	19.7
20,000-100,000	51.4	29.1	23.6	17.9
<20.000	48.5	26.5	21.3	15.3
HOUSEHOLD MEMBERS	49.6	29.7	25.4	17.2
One	52.8	30.9	26.1	18.3
Two	53.1	29.3	23.8	18.4
Three	56.9	33.0	27.7	19.9
Four	47.3	28.1	21.3	16.4
Five or more				
NATIONALITY				
Spanish	54.6	30.9	25.6	18.8
Foreigner	42.7	29.1	23.8	16.3
OCCUPATION				
Employed	61.4	35.4	30.1	21.6
Unemployed	40.2	22.8	17.5	12.8
Retired	32.5	15.3	13.4	8.7
Student	60.6	39.5	30.7	23.7
Housekeeper	23.5	9.6	6.6	6.4
Other	35.2	19.3	14.2	13.1
INCOME				
<900	29.7	15.0	11.0	8.1
900-1599	44.6	24.4	18.8	16.0
1600-2499	61.5	34.8	28.6	21.6
2500-2999	73.4	44.1	38.5	27.0
3000<	79.7	50.8	43.8	28.8
TOTAL	53.3	30.7	25.4	18.5

NOTES. Weighted data: The sampling weight is the number of subjects in the population represented by each observation. Penetration rates as a percentage of individuals aged 16 to 74, who used the Internet within the last year. An individual is considered as an eBuyer in any of the categories if he or she has carried out an online purchase within the last 12 months.

Theoretical and empirical models

A sound basis for modelling the economic decisions made by the individuals is the discrete choice model framework that describes choices between two or more discrete alternatives. Such choices contrast with standard consumption models in which the quantity of each good is assumed to be a continuous variable. Instead of analyzing "how much" as in situations with continuous variables, discrete choice analysis tries to answer the question: "which one". Techniques such as logistic regression can be used for empirical modelling of discrete choice. Estimation of such models is usually done via maximum likelihood methods.

Discrete choice models analyze choices made by consumers among a finite set of alternatives. The models have been used to examine telecommunications equipment and services demands, among others. In the paragraphs below, the decision-making unit is assumed to be an individual. The discussion below follows Nobel Laureate in Economics 2000 McFadden, (2001) who is a pioneer in developing the theoretical basis for discrete choice.

Discrete choice models relate the choice made by each individual to the attributes of the individual and the attributes of the alternatives available to the person. The model approximates the probability that a person chooses a particular alternative. The models could be used to forecast how individual choices will change under variations in the attributes of the alternatives and/or socio-demographics.

In practice, we cannot observe all factors affecting individual decisions as their determinants are partially unobserved or measured with error. In particular, the discrete choice approach uses stochastic specifications to account for unobserved variables related to choose alternatives, taste heterogeneity, individual dynamics, and also heterogeneous choice alternatives.

A special case is the random utility maximization model (RUM) of Marschak (1959). Discrete choice models can be derived from utility theory and are useful because they give a precise meaning to the probabilities P_i, where i is the subscript for individuals.

U_i denotes the utility (or net benefit or well-being) that individual *i* obtains from choosing one alternative (buy or not buy). The behavior of the individual is utility-maximizing: individual *i* chooses the alternative yielding the highest utility. The choice of individual *i* is denoted by a dummy variable, y_i:

$$\boldsymbol{y_i} = \begin{cases} \mathbf{1} & \boldsymbol{U_i(buy) \geq U_i(not\ buy)} \\ \mathbf{0} & \boldsymbol{otherwise} \end{cases} \quad \textbf{(1)}$$

The choice made by individual *i* depends on many factors, some of which are observed and some of which are not. For convenience, a linear approximation can be expressed as

$$\boldsymbol{U_i = \beta\, z_i + \varepsilon_i} \quad \textbf{(2)}$$

where

- z_i is a vector of observed variables relating to individual *i* that depends on attributes of the alternative, interacted perhaps with attributes of the individual, s_i, such that it can be expressed $z_i = z\,(x_i, s_i)$ for some numerical function *z*,
- β is a corresponding vector of coefficients of the observed variables, and
- ε_i captures the impact of all unobserved factors that affect the person's choice.

The individuals considered in this paper are those e-commerce users who purchased (or not) online goods or services in the last year with ages between 16 to 74 years. Conditional on being e-commerce users, this paper considers three types of e-commerce: CBeC, CBeC_EU, and CBeC_RW. Each one is modeled using a binary dependent variable for each of the estimated empirical models. The models considered below are specific cases of the combination of models 1 and 2 above. Here each type of CBeC is modeled independently, which is a convenient simplification.

In our sample most of the individuals buying at least once outside the EU also buy within the EU; 71.84% of buyers outside the EU also buy within the EU. They are not mutually

exclusive events. The behavior of each of the three types of e-commerce is approximated using distinct logit models.

The dependent and explanatory variables of the cross-border e-commerce models are the following:

Dependent variables	
CBeC	1 if the individual buys online from sellers from a country other than Spain; 0 otherwise.
CBeC_EU	1 if the individual buys online from sellers from other EU countries; 0 otherwise.
CBeC_RW	1 if the individual buys online from sellers from other countries from the rest of the world; 0 otherwise.
Sociodemographic	
Gender:	1 if male; 0 if female.
Age	Considered groups: 16-24, 25-34, 35-44, 45-54, 55-64, and 65 or more years old.
Education	It is measured by the level of study: primary or less, secondary, bachelor's, and master's or PhD
Nationality	1 if foreigner; 0 if Spanish.
Individual skills	
PC Skills	None, low, medium and high. This index was built based on the computer-related specific tasks (weighted according to their degree of difficulty) the respondents declared they carried out.
Internet Skills	Low, Medium, High, Very high. This index was built based on the Internet-related tasks (27 items were used to

	determine the 4 levels of Internet Skills) the respondent declared they carried out.
Online Opinion Seeker	Hardly ever, Sometimes, Always. This variable is based on the frequency the respondents declared they used information from online reviews of other customers before buying online.
Risk	
Internet Trust	Low, Medium, High. These levels have been reported by the respondents.
Economic	
Income	Grouped by different ranges: less than 900, 900-1599, 1600-2499, 2500-2999, more than 3000 euros per month. These categories reflect the monthly household income.

Two of the explanatory variables are binary: Gender and Nationality; all the others are categorical.

The explanatory variables are classified as sociodemographic (Gender, Age and Nationality), individual skills (Education, PC skills, Internet skills, and Online Opinion Seeker), risk (Internet trust) and economic (Family Income).The individual skills variables play the role of reducing the costs of transactions, Internet trust is supposed to decrease the cost, income might increase the benefits, and the sociodemographic variables may have more ambiguous signs, except for Nationality which may signal higher benefits (access to a wider offer of products and services) and/or lower costs (foreign language, for instance).

Internet Trust measures the declared trust of an individual for using the Internet for any purpose in general. Online Opinion Seeker measures the ability of an individual to acquire and process reviews by other customers to make an informed decision about a specific supplier, good or service (Park *et al*., 2007). Internet Trust relates essentially to the trust on a specific channel, while Online Opinion Seeker relates to the trust on a specific supplier

or good. The variables above are used to model the occurrence of each of the e-commerce options using the logistic function.

The results of the estimation of the different models are contained in *Table 2*. The first column contains the explanatory variables used along with the different models. The next three columns contain the odds ratios of each category of each variable, with respect to the base category, and the z statistic of the individual significance of the odds ratio. The z is standard normal under the null of no effect, with a 95% confidence interval of plus and minus 1.96. Only odds ratios that are significantly different from 1 are shown in the table. This model is conditional on being an e-commerce user.

The columns under the header "CBeC" contain the estimations of the odds ratios for the model of buyers who purchase goods or services outside the national borders and its z-statistics.

The columns under the header "CBeC_EU" correspond to the odds ratios and z-statistics of the buyers that purchase in the European Union.

And the columns under the header "CBeC_RW" correspond to the odds ratios z-statistics of buyers that purchase in any other country outside the EU.

The choice of these specific models corresponds to the questions posed by the previous international literature. Several of them correspond to the purchases by households from foreign or national suppliers. The number of observations for each model may not be totally intuitive. Their variation across models depends on the availability of data for each variable used in each model within the ICT-H sample.

The comments for *Table 2* are organized by variables, to facilitate the comparability across models.

GENDER: The effect is higher for males than for females in the three models.

AGE: Someone's age does not affect the odds of making purchases abroad, as seen with being an e-commerce user in the model proposed by Garín *et al*. (2018).

Table 2:
Models for adoption of cross-border e-commerce

		CBeC		CBeC_EU		CBeC_RW	
		Odds ratios	**z**	**Odds ratios**	**z**	**Odds ratios**	**z**
GENDER	Female						
	Male	1.42	4.60	1.52	5.63	1.33	3.30
AGE	16-24						
	25-34	--	--	--	--	--	--
	35-44	--	--	--	--	--	--
	45-54	--	--	--	--	--	--
	55-64	--	--	--	--	--	--
	> 65	0.64	-2.29	--	--	--	--
EDUCATION	Primary/less						
	Secondary	--	--	2.17	2.79	--	--
	Bachelor	--	--	2.89	3.78	--	--
	Master/PhD	1.90	2.40	3.72	4.64	--	--
PC SKILLS	None						
	Low	1.60	2.20	--	--	2.58	3.18
	Medium	1.91	3.05	1.87	2.75	2.67	3.29
	High	1.98	3.12	1.90	2.80	3.39	4.08
INTERNET SKILLS	Low						
	Medium	--	--	--	--	--	--
	High	1.53	2.19	--	--	1.87	2.29
	Very high	2.55	4.53	2.59	4.30	2.84	3.71
INTERNET TRUST	Low						
	Medium	1.20	2.06	--	--	--	--
	High	1.65	3.61	--	--	1.73	3.52
ONLINE OPINION SEEKER	Hardly ever						
	Sometimes	1.65	3.75	1.90	4.58	1.41	1.98
	Always	1.97	5.78	2.12	6.15	2.10	4.81
INCOME	<900						
	900-1599	--	--	--	--	1.46	2.14
	1600-2499	--	--	--	--	--	--
	2500-2999	--	--	--	--	--	--
	>3000	--	--	--	--	--	--
NATIONALITY	Spanish					--	--
	Foreigner	1.86	3.35	1.61	2.72		
CONSTANT		0.14	-5.58	0.06	-7.63	0.03	-8.56
N. observations		5,576		5,576		5,587	
Wald χ^2		318.86 DF: 20		336.52 DF: 13		190.27 DF: 15	
Pseudo R^2		0.0745		0.0817		0.0632	
Correctly classified		63.77%		63.18%		69.09%	

Notes: Weighted logistic regression. Robust estimates. Reference category: female, 16-24 years, no studies or primary studies, no PC skills, low internet skills, low internet trust, hardly ever Online Opinion Seeker and net income levels below 900€ per month. Estimations refer to those who have purchased online during the last year. The pseudo-R^2 is a measure of the in-sample goodness of fit. However, the relevant statistics for testing the significance of the estimated coefficients is the χ^2 test of the joint significance of all slope coefficients, which is very high and has very low p-values: below 0.001 in all cases, and also the percentages of correct classifications, which are above 63%.

EDUCATION: Higher levels of education are associated with higher odds ratios in the CBeC and CBeC_EU models. However, they are insignificant in the CBeC_RW model, suggesting a more homogeneous behavior across the diverse groups of education.

PC SKILLS: The level of skills using computers has a significant and positive effect on the three models.

INTERNET SKILLS: This variable is mostly significant with positive signs and relative magnitudes on the three models.

INTERNET TRUST: This variable has significant odds ratios in CBeC while for CBeC_RW model only the highest category is significant. This may be interpreted as what is important for becoming an e-commerce user is trust, but once this barrier is crossed, its effect is negligible, except for the high level of trust.

ONLINE OPINION SEEKER: This variable is relevant when deciding to become a cross-border e-commerce user (any of the three types). The larger the use of this type of information, the larger the odds of making purchases abroad (either in the EU or from the rest of the world).

INCOME: It is essentially insignificant across models and income segments. This suggests that once you are an e-commerce user other factors rather than income are driving your purchase behavior.

NATIONALITY: This variable is significant in CBeC and CBeC_EU models. This suggests that being a foreigner increases the odds of becoming a cross border eBuyer.

The numbers of observations vary between 5,587 and 5,576, which constitute reasonably large samples, which are representative of all Spanish-resident households, including the regional dimension.

The Wald χ^2 is a test of the joint significance of all the parameters in each equation. Under the null of insignificance of all the parameters, the statistic behaves as a χ^2 with a number of degrees of freedom equal to the number of estimated coefficients in each model. The computed values are highly significant in all the models considered.

The Pseudo R^2 is a measure of the goodness of fit. Higher values denote a better fit. The values in these models are significant enough to be able to estimate with sufficient precision the statistics of interest and to be able to make inference and policy recommendations since most of the parameters are individually and jointly significant. The correctly classified measure is another goodness of fit statistic which measures the percentage of cases correctly predicted by the model. The estimated percentages are reasonably high in the models[10] of *Table 2*.

The decision to purchase online is also influenced by the characteristics of the good/service being purchased, as well as a variety of unobserved individual characteristics and perceptions of each individual. Therefore, supply-side considerations would help to improve de the model. Recognizing unobserved individual heterogeneity is also relevant in this context.

However, the availability of relevant demand-side data is limited to the variables that we have discussed so far in this paper, while supply-side characteristics are not included in our sample.

The approach here is that in the data set, the large quantity of data makes up for the relatively low quality of the data on some variables. In this fashion, it is possible to estimate meaningfully the sign, size, and significance of each one of the effects of the variables included in the models.

Due to the possible presence of collinearity in the explanatory variables of the model, polychoric correlations among the independent variables have been computed, as shown in *Table 3*.

[10] The pseudo-R^2 is a measure of the in-sample goodness of fit and it seems somewhat low in these models. However in this case the relevant statistics for testing the joint significance of the estimated coefficients are the χ^2 tests of joint significance of all slope coefficients, which are highly significant in all cases: 318.86, 336.52 and 190.27 with very low p-values, which are below 0.001 in all cases. The percentages of correct classifications are also high with values of 63.77%, 63.18% and 63.09% respectively.

In *Table 3*, most of the computed polychoric correlations (Drasgow, 1988, Stata Corporation, 2017) are moderate. However, the estimated standard errors and associated LR statistics show that they are mostly significantly different from zero. Thus, it is likely that collinearity is partly responsible for the insignificance of AGE and INCOME.

Table 3:
Polychoric correlations among selected independent variables

	AGE	INCOME	PC SKILLS	INTERNET SKILLS	INTERNET TRUST
AGE	1.000				
INCOME	0.045	1.000			
PC SKILLS	-0.387	0.344	1.000		
INTERNET SKILLS	-0.391	0.312	0.729	1.000	
INTERNET TRUST	-0.103	0.157	0.246	0.261	1.000

It can be mentioned that preliminary results using pooled data 2014-2017 indicate that all independent variables, including age and income, are significant in the CBeC model.

It should be pointed out that the models presented in this paper are conditional on performing e-commerce. In the e-commerce model of Garín *et al.* (2018) the variables AGE and INCOME are both significant. This may suggest that they are relevant for adopting e-commerce, while they appear to be insignificant in the cross-border equations (their effect cannot be measured accurately with this sample, model and estimation technique).

Conclusions

This paper deals with the drivers and barriers for the individual adoption of cross-border e-commerce using data **from the 2016 Survey on ICT usage in households** INE (2017).

The paper estimates three models of cross-border e-commerce, conditional on the individual being a user of e-commerce. The three models are CBeC, CBeC_EU, and CBeC_RW. A summary of the main findings contained in *Table 2* goes as follows:

Being a Male is positively related to using cross-border e-commerce. **Age** is mostly significant with a negative effect on CBeC for those aged above 65. **Education** is positively related to CBeC_EU while it seems mostly insignificant in the other cases.

PC skills is significant and positive in all the models, especially in the CBeC model. However, **Internet skills** are mostly significant, especially at high and very high levels.

Trust in the Internet seems to be relevant and positive for becoming a cross-border eBuyer especially in the case of higher levels of trust.

When a potential customer is an **Online Opinions Seeker** he or she is more likely to do cross-border e-commerce. On the other hand, income seems to be mostly irrelevant. **Foreign** nationality increases the likelihood of using cross-border e-commerce in general and decreases the likelihood of a CBeC_RW.

The European Commission and the national authorities decided to promote cross-border e-commerce for a variety of reasons discussed in previous sections. For doing so, policy instruments are needed. Several variables of the models can hardly be manipulated to be used as policy instruments, like gender, age, education, income, and nationality. However, there are some variables on which one could act in an attempt to boost cross-border e-commerce such as PC skills, Internet skills, Internet trust, and online opinions seeker. And, even if it has not been considered as a variable, knowledge of foreign languages plays a decisive role in buying online abroad. Creative measures designed to improve them could be implemented. They should be targeted at population groups with less cross-border e-commerce penetration. One such example could be to use specific education measures

targeted in older people and housekeepers, who have lower penetration and more time to learn. These measures are in line with the EC (2017) report that points to deficiencies in the demand side in Spain, mainly centered in human capital while focusing on ensuring equal conditions for online buyers and sellers across Europe through homogenization of regulation.

The conclusions of this study require some caveats. The first is that several questions in the survey are not designed specifically for the topics discussed in this paper. Second, cross-border e-commerce may not be well approximated via surveys since many consumers may not know exactly where the website or the center for dispatching the merchandise is located. Another qualification is that static models are being used in situations in which the dynamics seem to be relevant. A further limitation is that only the demand side is analyzed. However, in the context of survey data of only one year, the supply is given, and it is treated as such throughout this paper.

The research agenda contains several issues to be analyzed next. The first one is to use the pool of all the available years to estimate richer sub-models that allow for heterogeneity within the sample. Furthermore, a full panel data set will be available shortly for 2008-2016. Allowing the use of dynamic panel data models to further characterize questions analyzed in the present paper, as well as additional ones like simultaneity and dynamic relationships. Next, it would be important to focus on other relevant issues that can be analyzed using the panel data of this survey, like the demand for IT equipment by households, demand for bundled services, and the demands for a variety of other services, both allowing for geographical and economic dimensions.

References

Cardona, M., Duch-Brown, N., Francois, J., Martens, B., & Yang, F. (2015). The Macro-Economic Impact of e-Commerce in the EU Digital Single Market. Institute for Prospective Technological Studies Digital Economy Working Paper 2015/09.

Cardona, M., Duch-Brown, N., & Martens, B. (2015). Consumer Perceptions of (Cross-border) eCommerce in the EU Digital Single Market. Institute for Prospective Technological Studies Digital Economy Working Paper (2015-06).

Correa, M., García, J. R., & Tabanera, A. (2015). Comercio electrónico y hábitos de consumo en España: la importancia de la banca on-line. Observatorio de Economía Digital, BBVA Research.

Drasgow, F. (1988). Polychoric and Polyserial Correlations. In Encyclopaedia of Statistical Sciences. John Wiley & Sons, Inc. https://doi.org/10.1002/0471667196.ess2014.pub2

European Commission. (2012). A coherent framework for building trust in the Digital Single Market for e-commerce and online services.

European Commission. (2016). A comprehensive approach to stimulating cross-border e-commerce for European citizens and businesses.

European Commission. (2017a). Europe´s Digital Progress Report (EDPR) 2017 country profile SpainEuropean Commission. Retrieved from https://ec.europa.eu/digital-single-market/en/news/europes-digital-progress-report-2017

European Commission. (2017b). Proposal for Regulation of the European Parliament and Council on cross border parcel delivery services.

European Commission. (2017c). Shaping the Digital Single Market. Retrieved November 17, 2017, from https://ec.europa.eu/digital-single-market/en/policies/shaping-digital-single-market

European Parliament. (2017, December 10). Better cross-border parcel deliveries to boost e-commerce and help consumers. Retrieved November 19, 2017, from http://www.europarl.europa.eu/news/en/press-room/20171009IPR85647/better-cross-border-parcel-deliveries-to-boost-e-commerce-and-help-consumers

European Parliament and Council. (2005). The Unfair Commercial Practices Directive.

Eurostat. (2017). Database - Eurostat. Retrieved July 22, 2017, from http://ec.europa.eu/eurostat/data/database

Falk, M., & Hagsten, E. (2015). E-commerce trends and impacts across Europe. International Journal of Production Economics, 170, 357–369. https://doi.org/10.1016/j.ijpe.2015.10.003

Garín Muñoz, T., Lopez, R., Perez Amaral, T., Herguera García, I., & Valarezo Unda, A. (2017). Models for Individual Adoption of eCommerce, eBanking and eGovernment in Spain (SSRN Scholarly Paper No. ID 3014108). Social Science Research Network. Retrieved from https://papers.ssrn.com/abstract=3014108

Garín Muñoz, T., & Pérez-Amaral, T. (2011). Factores Determinantes del Comercio Electrónico en España. Boletín Económico Del ICE, 3016, 51-65.

Gefen, D. (2000). E-commerce: the role of familiarity and trust. Omega, 28(6), 725–737. https://doi.org/10.1016/S0305-0483(00)00021-9

Gomez-Herrera, E., Martens, B., & Turlea, G. (2014). The drivers and impediments for cross-border e-commerce in the EU. Information Economics and Policy, 28, 83–96. https://doi.org/10.1016/j.infoecopol.2014.05.002

INE. (2016). Survey on Equipment and Use of Information and Communication Technologies in Households. Retrieved July 19, 2017, from http://www.ine.es/dyngs/INEbase/en/operacion.htm?c=Estadistica_C&cid=1254736176741&menu=ultiDatos&idp=1254735976608

Manyika, J., Lund, S., Bughin, J., Woetzel, J., Stamenov, K., & Dhingra, D. (2016). Digital globalization: The new era of global flows. McKinsey Global Institute.

Marschak, J. (1959). Binary Choice Constraints on Random Utility Indicators (Cowles Foundation Discussion Papers No. 74). Cowles Foundation for Research in Economics, Yale University. Retrieved from https://ideas.repec.org/p/cwl/cwldpp/74.html

Martens, B. (2013). What Does Economic Research Tell Us About Cross-Border E-Commerce in the EU Digital Single Market? (SSRN Scholarly Paper No. ID 2265305). Rochester, NY: Social Science Research Network. Retrieved from https://papers.ssrn.com/abstract=2265305

McFadden, D. (2001). Economic Choices. American Economic Review, 91(3), 351–378. https://doi.org/10.1257/aer.91.3.351

ONTSI. (2016). Perfil sociodemográfico de los internautas (datos INE 2016). Retrieved November 19, 2017, from http://www.ontsi.red.es/ontsi/es/content/perfil-sociodemogr%C3%A1fico-de-los-internautas-datos-ine-2016

Park, D.-H., Lee, J., & Han, I. (2007). The Effect of On-Line Consumer Reviews on Consumer Purchasing Intention: The Moderating Role of Involvement. International Journal of Electronic Commerce, 11(4), 125–148. https://doi.org/10.2753/JEC1086-4415110405

PayPal. (2016). PayPal Cross-Border Consumer Research 2016.

Pfeffermann, D. (1993). The Role of Sampling Weights When Modelling Survey Data. International Statistical Review / Revue Internationale de Statistique, 61(2), 317–337. https://doi.org/10.2307/1403631

Stata Corporation. (2017). Stata Manual.

The Economist. (2017). E-commerce. The New Bazaar (pp. 1–14). Retrieved from https://www.economist.com/news/special-report/21730546-e-commerce-transforming-business-and-daily-life-mostly-better-says-charlotte

CHAPTER 3:
Adoption of e-commerce by individuals in Spain using pooled data 2008-2017[11]

[11] Book in Honor to Gary Madden. Manuscript submitted for publication (2019); with López, R., & Pérez-Amaral, T.

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, analysing 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 to being male, having higher levels of education, income, 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

Acknowledgments: Paper presented at the 29th ITS European Conference, Trento-Italy, 2018. This research has been 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.

Adopción del comercio electrónico por parte de individuos en España. Pool 2018-2017[12]

Resumen

Las tasas de penetración del comercio electrónico son distantes entre los grupos de personas con niveles más bajos y altos de adopción del comercio electrónico. Este es un indicador de brecha digital, que tienen efectos negativos en términos de oportunidades no explotadas por parte de personas, empresas y para el conjunto de la economía. En este trabajo se exploran los determinantes socioeconómicos y demográficos claves en la adopción del comercio electrónico, analizando una base de datos compuesta por 174.776 observaciones pare el período 2008-2017 en España. El análisis empírico se fundamenta en el marco neoclásico estándar de maximización de la utilidad. Se han empleado, el Modelo Lineal de Probabilidad, Regresión Logística y el modelo de Corrección de Selección de Heckman. Los resultados sugieren que la adopción del comercio electrónico está relacionada positivamente con ser hombre, contar con niveles altos de educación, ingresos y habilidades digitales, ser español y estar empleado; mientras que ser mujer, mayor edad y pertenecer a un hogar de dos o más miembros tiene efectos negativos. En el modelo se introdujo una interacción entre habilidades digitales y edad, donde las habilidades digitales altas parecen tener una influencia positiva, contrarrestando en parte los efectos negativos de algunos grupos de edad. Para fomentar la adopción del comercio electrónico se sugieren recomendaciones de políticas relacionadas con medidas de demanda y oferta.

[12] Book in Honor to Gary Madden. Manuscript submitted for publication (2019); with López, R., & Pérez-Amaral, T.

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 10th 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[13], Human Capital[14] (or digital skills), Use of Internet Services by citizens[15], Integration of Digital Technology[16] and Digital Public Services[17]. 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[18] average and several indicators suggest it is due to a weak demand side, especially on the private user side.

[13] Fixed broadband, mobile broadband and prices.

[14] Internet use, basic and advanced digital skills.

[15] Citizens' use of content, communication and online transactions.

[16] Business digitisation and e-commerce

[17] eGovernment and eHealth.

[18] EU-28: Member States of the European Union.

The modelling 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 are 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 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 of 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 of the individual decision of becoming an online buyer for private use, descriptive analysis and quantitative models are formulated.

The pooled 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 the 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.

Literature review

Models of ICT adoption

Adoption of technology, online consumer behavior, and the specific decision to adopt e-commerce have been studied based on several frameworks. By far, the most used approaches are applications, adaptations and/or unification 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.

Digital divide

Not long ago the main indicators of digital divide were related to the "haves" and "haves nots", however, NTIA[19] (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.

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 demand for 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

[19] National Telecommunications and Information Administration

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.

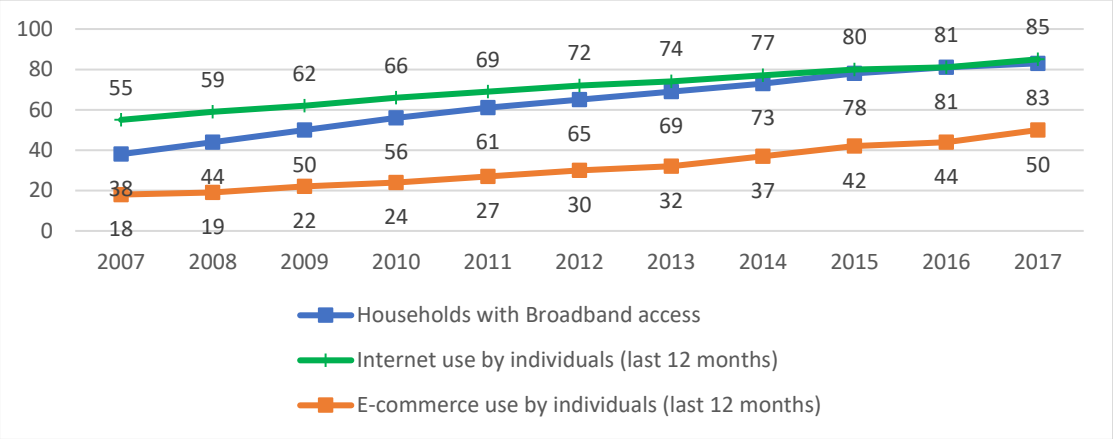
The above-cited empirical studies have mostly specified binary response models. The demographic and socioeconomic characteristics of households and individuals have been used and interpreted. Although most of the works that have dealt with INE's survey on ICT account for the explanatory factors, they do not analyze specific subcategories. 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[20]), 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).

[20] Having or not Spanish nationality.

Adoption of e-commerce in Spain

In 2017, 82.7% of Spanish households had a broadband connection and 85% of the Spanish population (between 16 and 74 years[21]) used Internet at least once in the last twelve months, although only 69.0% did 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 a percentage of people aged 16 to 74 (2007–2017)



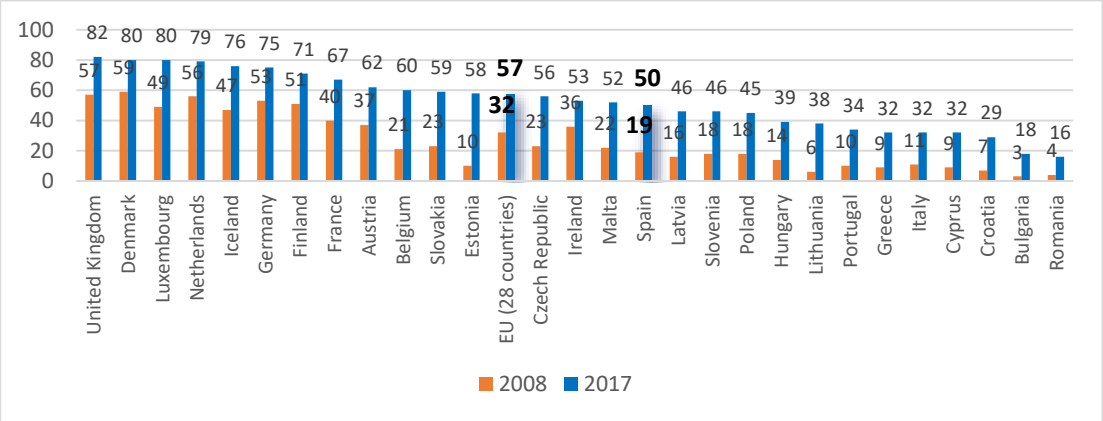
Source: Own elaboration based on data from INE, 2018. Survey on Equipment and Use of Information and Communication Technologies in Households.

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

[21] 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.

According to the country-specific profile report (European Commission, 2018a) Spain belongs to the medium performance group of countries[22] and ranks 10th among the 28 EU Member States in DESI 2018, improving two positions with respect to 2017[23]. 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 17th, 4 places behind the EU average (*Figure 2*).

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



Source: Own elaboration based on data from Eurostat, 2018. Penetration rates as the percentage of people aged 16 to 74.

Aggregate data seems to suggest that the penetration rates of e-commerce, like 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

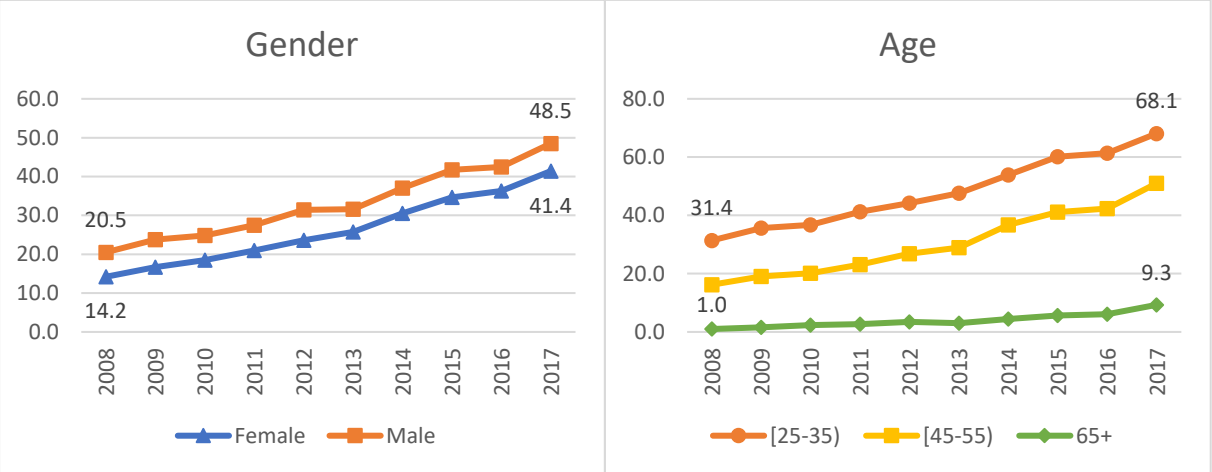
[22] DESI's medium-performing countries are Spain, Austria, Lithuania, Germany, Slovenia, Portugal, Czech Republic, France and Latvia.

[23] 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%.

are widening when the differences between groups with the lowest and largest penetration rates in 2008 and 2017 are measured in absolute values[24].

The distance between males and females 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 3:
Evolution of individual penetration rates of e-commerce in Spain (2008-2017): Gender and Age



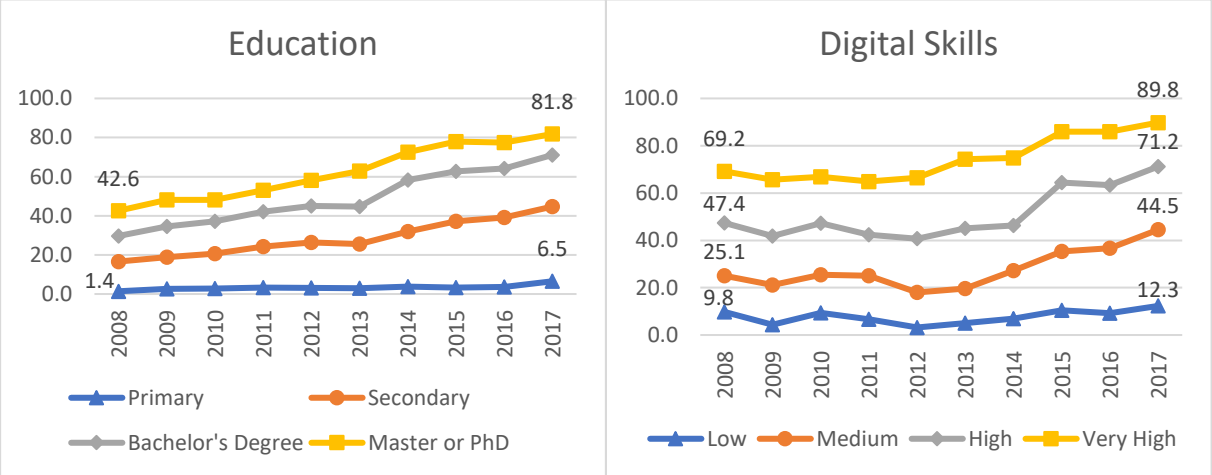
Source: Own elaboration based on data from INE, 2018. Survey on Equipment and Use of Information and Communication Technologies in Households.

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 Ph.D. went from 20.5% to 83.5%. Individuals with low digitals skills went from 9.8% in 2008 to

[24] 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.

12.3% in 2017, while those with a very high level went from 69.2% in 2008 to 89.8% in 2017.

Figure 4:
Evolution of individual penetration rates of e-commerce in Spain (2008-2017)

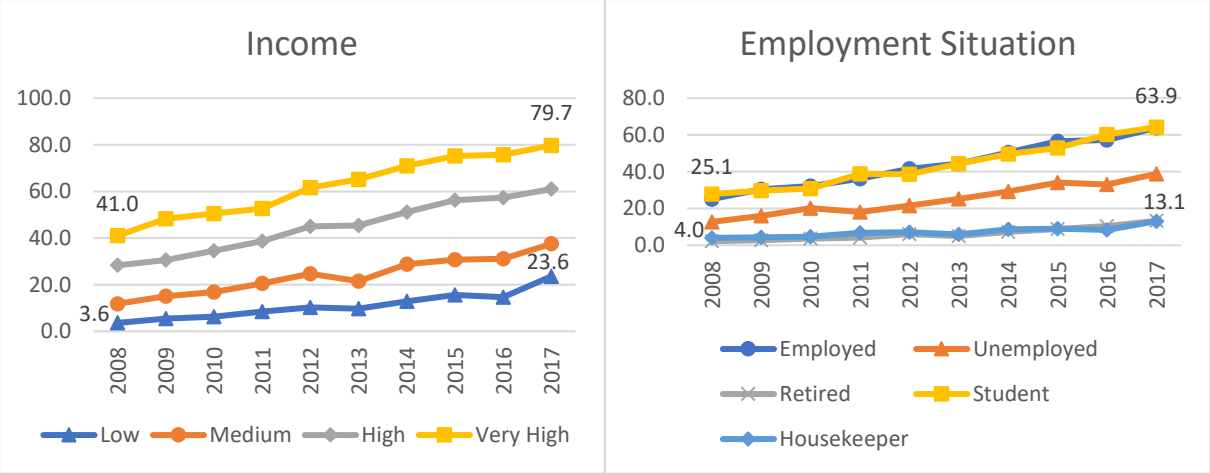


Source: Own elaboration based on data from INE, 2018. Survey on Equipment and Use of Information and Communication Technologies in Households.

Income and Employment situation also reflects 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, while employed went from 13.1% in 2008 to 63.9% in 2017 (*Figure 5*).

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 the 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*).

Figure 5:
Evolution of individual penetration rates of e-commerce in Spain (2008-2017)



Source: Own elaboration based on data from INE, 2018. Survey on Equipment and Use of Information and Communication Technologies in Households.

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[25] 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,

[25] 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.

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, use of e-government and ubiquitous connectivity. The data allows classifying households by region of residence, geographical location, population density, type of household and 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 Ph.D. degree. In the case of Digital Skills, the penetration is 89.8% for a very high level, compared with 12.3% for a low level. For low, medium, high and very high income, the penetrations rates were 23.6%, 37.5%, 61.0% and 79.7%, respectively.

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

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 corresponds 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.

Comparing regions, Madrid, Navarra, Baleares, Cantabria, and Cataluña are the top five autonomous communities with the highest penetration rates. País Vasco, that was 4th in 2008, dropped to the 9th position in 2017; and the 4th place now is occupied by Cantabria, which was 7th in 2008. It is remarkable how the autonomous city of Ceuta (in northern Africa) has moved from 17th to 7th, 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.

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 modelling the demand for bandwidth, and Demoussis & Giannakopoulos (2006) used to model the individual decision to adopt the Internet.

The adoption of this simple behavioral model assumes that the individual has an economically predictable behavior, allowing the analysis of choices made among a finite set of alternatives, which in this case are buying online or not.

The e-commerce adoption decision $\boldsymbol{y_i}$ is determined by the surplus that the decision-maker (individual $\boldsymbol{i}$) obtains from choosing whether to buy online for private use or not. This is when the costs of buying online are exceeded by the benefits the individual derives utility, $\boldsymbol{U_i}$, from using the Internet to buy goods and services.

$$\text{E-commerce adoption: } \begin{cases} \boldsymbol{y_i = 1} & \boldsymbol{U_i(buy\ online) > U_i(not\ buy\ online)} \\ \boldsymbol{y_i = 0} & \boldsymbol{otherwise} \end{cases} \quad \textbf{(1)}$$

The model of individual adoption based on Utility Maximization Framework:

$$\boldsymbol{U_i = f(\beta\ z_i) + \varepsilon_i} \quad \textbf{(2)}$$

$$z_i = z\,(x_i, s_i) \qquad \textbf{(3)}$$

- β is the vector of coefficients of the observed explanatory variables.
- z_i depicts the vector of observed variables related to person i (where the attributes of the alternatives interact with the specific characteristics of the decision-maker).
- ε_i is the influence of all unobserved factors.
- x_i attributes of the alternatives.
- s_i specific characteristics of the person i.

The decision of buying online for private use is represented by the binary dependent variable defined below:

E-commerce	= 1, if the individual bought online for private use in the last 12 months.
	= 0, otherwise.

Explanatory variables are grouped as follows: Sociodemographic, Individual Skills, Economic and Geographic, and Time variables:

Sociodemographic	Gender:	2 groups: 1 if male, 0 if female
	Age:	6 groups
	Habitat:	4 groups
	Household Members:	5 groups
	Nationality:	2 groups: 1 if Spanish, 0 if Foreign
Individual skills	Education:	4 levels of study
	Digital Skills:	4 levels

Economic	Employment Situation:	6 groups
	Income:	4 groups, monthly net income
Geographic and Time	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[26] 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.

[26] 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.

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
Gender	Female						
	Male	**0.04**	**10.33**	**1.29**	**10.40**	**0.04**	**10.2**
Age	16-24						
	25-34	-0.03	-1.89	0.82	-0.98	-0.03	-1.5
	35-44	**-0.09**	**-5.05**	**0.52**	**-3.49**	**-0.08**	**-4.64**
	45-54	**-0.11**	**-6.39**	**0.46**	**-4.13**	**-0.11**	**-5.86**
	55-64	**-0.16**	**-9.19**	**0.31**	**-6.00**	**-0.15**	**-8.35**
	> 65	**-0.19**	**-9.39**	**0.24**	**-6.38**	**-0.18**	**-8.37**
Education	Primary or less						
	Secondary	**0.04**	**5.77**	**1.45**	**6.36**	**0.04**	**5.13**
	Bachelor	**0.09**	**11.13**	**1.90**	**10.25**	**0.09**	**10.15**
	Master/PhD	**0.13**	**14.88**	**2.34**	**13.45**	**0.12**	**13.52**
Digital Skills	Low						
	Medium	**0.11**	**5.51**	**2.48**	**4.78**	**0.11**	**5.21**
	High	**0.26**	**12.65**	**4.88**	**8.71**	**0.25**	**12.2**
	Very high	**0.47**	**23.89**	**12.52**	**13.84**	**0.47**	**23.19**
Habitat	<20000						
	20000-100000	-0.01	-1.44	0.95	-1.57	-0.01	-1.78
	100000-500000	**-0.04**	**-5.07**	**0.78**	**-5.24**	**-0.04**	**-5.38**
	>500000	0.00	0.86	1.02	0.73	0	0.57
Household Members	One						
	Two	**-0.01**	**-2.10**	0.93	-1.85	**-0.02**	**-2.35**
	Three	**-0.04**	**-6.72**	**0.78**	**-6.36**	**-0.05**	**-6.86**
	Four	**-0.05**	**-7.00**	**0.76**	**-6.71**	**-0.05**	**-7.26**
	Five or more	**-0.08**	**-8.60**	**0.62**	**-8.34**	**-0.08**	**-8.72**
Nationality	Foreigner						
	Spanish	**0.03**	**3.30**	**1.16**	**2.82**	**0.03**	**3.09**
Employment Situation	Employed						
	Unemployed	**-0.05**	**-8.19**	**0.74**	**-7.80**	**-0.05**	**-7.96**
	Retired	**-0.04**	**-3.18**	**0.82**	**-2.72**	**-0.03**	**-2.49**
	Student	**-0.06**	**-4.71**	**0.73**	**-5.00**	**-0.06**	**-5.05**
	Housekeeper	**-0.02**	**-2.58**	**0.84**	**-2.65**	**-0.02**	**-1.91**
	Other	**-0.02**	**-1.60**	0.90	-1.36	**-0.02**	**-1.16**
Income	Low						
	Medium	**0.04**	**6.69**	**1.32**	**6.97**	**0.04**	**5.8**
	Medium-high	**0.11**	**15.78**	**1.91**	**15.14**	**0.1**	**14.45**
	High	**0.17**	**20.28**	**2.71**	**19.45**	**0.16**	**18.92**
Digital Skills × Age[27]	High × 55-64	**0.14**	**5.64**	**2.40**	**4.15**	**0.14**	**5.44**
	High × 65 or more	**0.12**	**3.83**	**2.40**	**3.61**	**0.11**	**3.59**
	Very high × 55-64	**0.08**	**3.50**	**2.04**	**3.30**	**0.08**	**3.26**
	Very high × 65 or more	**0.14**	**4.54**	**3.07**	**4.27**	**0.13**	**4.19**

[27] 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**
Year	2008						
	2009	-0.01	-1.58	**0.87**	**-2.41**	-0.02	-1.82
	2010	**0.03**	**2.64**	**1.14**	**2.32**	**0.02**	**2.32**
	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	**0.04**	**3.81**	**1.23**	**3.72**	**0.03**	**3.32**
	2014	**0.09**	**9.29**	**1.61**	**8.43**	**0.08**	**8.46**
	2015	**0.18**	**19.19**	**2.89**	**18.68**	**0.17**	**18.14**
	2016	**0.18**	**19.21**	**2.83**	**18.56**	**0.17**	**18.15**
	2017	**0.24**	**26.31**	**4.17**	**25.81**	**0.23**	**24.71**
Autonomous	Andalucía						
Community	Aragón	0.01	1.43	1.10	1.56	0.01	-1.82
	Asturias	**0.03**	**3.44**	**1.22**	**3.65**	**0.03**	**2.32**
	Baleares	**0.09**	**7.46**	**1.68**	**7.49**	0.08	1.15
	Canarias	**-0.05**	**-4.46**	**0.75**	**-4.31**	-0.05	0.74
	Cantabria	**0.05**	**5.26**	**1.40**	**5.42**	**0.05**	**3.32**
	Castilla la Mancha	0.02	1.79	1.11	1.68	**0.02**	**8.46**
	Castilla León	0.00	-0.05	1.01	0.19	**0**	**18.14**
	Cataluña	**0.05**	**5.68**	**1.32**	**5.68**	**0.05**	**18.15**
	Extremadura	0.01	1.07	1.06	0.98	**0.01**	**24.71**
	Galicia	0.00	0.00	1.00	0.09	0	-1.82
	La Rioja	**0.03**	**2.59**	**1.21**	**2.72**	**0.03**	**2.32**
	Madrid	**0.03**	**3.07**	**1.17**	**3.20**	0.02	1.15
	Navarra	**0.05**	**5.47**	**1.35**	**5.47**	0.05	0.74
	País Vasco	**0.05**	**5.82**	**1.39**	**5.86**	**0.05**	**3.32**
	Murcia	**-0.02**	**-2.52**	**0.87**	**-2.33**	**-0.02**	**8.46**
	Valencia	**-0.02**	**-1.99**	0.91	-1.81	**-0.02**	**18.14**
	Ceuta	-0.04	-1.55	0.78	-1.64	**-0.04**	**18.15**
	Melilla	-0.02	-0.58	0.88	-0.73	**-0.01**	**24.71**
Constant		0.01	0.44	0.05	-15.36	0.04	1.65
N. observations		**75,960**		**75,960**		**77,362**	
F		**556.10 DF: 70**					
Wald χ^2				**10581.81 DF: 70**		**29332.2 DF: 70**	
R^2		**0.3280**					
Pseudo R^2				**0.2759**			
Correctly classified				**75.95%**			
Wald χ^2: H_0: independent equations						**54.96 DF: 1**	

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^2 and the pseudo R^2 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 χ^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)[28]

Source: Own elaboration based on data from estimation results, *Table 2 (1)*

Since the dependent variable is dichotomous, a Logistic Regression Model (LRM), *Table 2 (2)*, has been estimated[29]. And to correct a possible sample selection bias, a Heckman Selection Model (HSM), *Table 2 (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[30].

Linear probability model

A linear probability model is a first approach to modelling 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 straightforward interpretation of the estimated coefficients in terms of differences in 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 educations are related to higher probabilities of adoption, going up to 0.09 for Bachelor and 0.13 for Master or Ph.D. A Spanish person has a probability of 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.

[28] Notes: point estimates are represented by dots; confidence intervals are represented by solid lines.

[29] All three models are estimated using Stata 15.

[30] 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 × Age' uses the same base categories already fixed for interacted explanatory variables.

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 Baleares 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 e-commerce adoption is moderated by the level of Digital Skills[31]. High Digital Skills seems to have a positive influence, partly counteracting the negative effects on the 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 analysing 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

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[32] estimated from the LRM also allow establishing 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 to the adoption of e-commerce.

[31] 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.

[32] 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.

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 than the odds of buying online have increased substantially over 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 to the odds of shopping online are Canarias and Murcia, with odds ratios of 0.75 and 0.87, respectively.

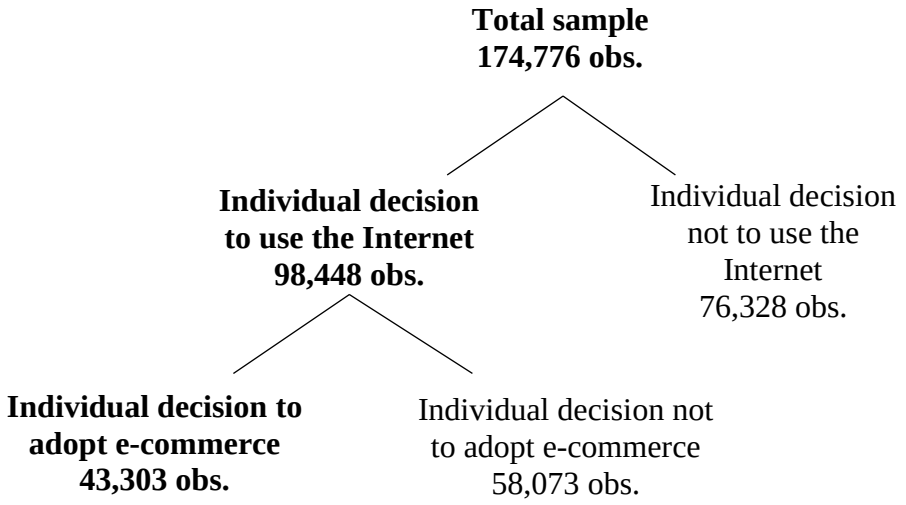
The magnitude of the coefficient estimates in the LPM and those of the LRM are not directly comparable. However, the signs are comparable and also the magnitudes of t and z statistics, which show a similar pattern in models (1) and (2).

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, Heckman's Selection Model (HSM) is estimated[33], as Cerno & Pérez-Amaral (2006a), Lera-López et al.(2011) and Pérez-Hernández & Sánchez-Mangas (2011), among others, did at modelling similar individual decision processes.

[33] 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.).

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



Source: Own elaboration based on processed data from INE, 2018.

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 (Baleares, 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[34] 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

[34] Estimation results of first stage of HSM: Appendix, *Table A-3*.

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.

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

Conclusions

General conclusions

When the proportion of e-buyers in Spain is compared with the 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) [35]; and, many of them are statistically significant at explaining the individual decision to perform cross-border e-commerce[36]. 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[37].

[35] Garín Muñoz, et. al., (2018)

[36] Valarezo, et. al., (2018)

[37] 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.

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

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 policymakers 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. The 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[38], 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[39].

Caveats

Some of the limitations found in this study are related to 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.

Further research

[38] A pure play e-commerce is an e-commerce business that only sell through the Internet.

[39] 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.

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.

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

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.bbvaresearch.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

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

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.bbvaresearch.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

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

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 Model (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
Digital Skills × Age	Medium × 16-24						
	Medium × 25-34	0.02	0.79	1.09	0.40	0.02	0.65
	Medium × 35-44	**0.07**	**2.90**	**1.65**	**2.43**	**0.07**	**2.85**
	Medium × 45-54	**0.06**	**2.57**	**1.57**	**2.16**	**0.06**	**2.57**
	Medium × 55-64	**0.08**	**3.23**	**1.89**	**2.94**	**0.08**	**3.17**
	Medium × 65 or more	0.04	1.63	**1.73**	**2.23**	0.04	1.57
	High × 16-24						
	High × 25-34	**0.09**	**3.88**	1.46	1.79	**0.09**	**3.57**
	High × 35-44	**0.13**	**6.00**	**2.12**	**3.79**	**0.13**	**5.83**
	High × 45-54	**0.12**	**5.38**	**2.01**	**3.47**	**0.12**	**5.26**
	High × 55-64	**0.14**	**5.64**	**2.40**	**4.15**	**0.14**	**5.44**
	High × 65 or more	**0.12**	**3.83**	**2.40**	**3.61**	**0.11**	**3.59**
	Very high 16 × 24						
	Very high × 25-34	**0.08**	**3.58**	**1.59**	**2.19**	**0.07**	**3.21**
	Very high × 35-44	**0.11**	**5.06**	**2.21**	**3.98**	**0.1**	**4.84**
	Very high × 45-54	**0.11**	**4.80**	**2.16**	**3.76**	**0.1**	**4.61**
	Very high × 55-64	**0.08**	**3.50**	**2.04**	**3.30**	**0.08**	**3.26**
	Very high × 65 or more	**0.14**	**4.54**	**3.07**	**4.27**	**0.13**	**4.19**

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's first stage estimates. Probit Regression Model.

Internet use. Probit Regression Model - Heckman first stage estimates

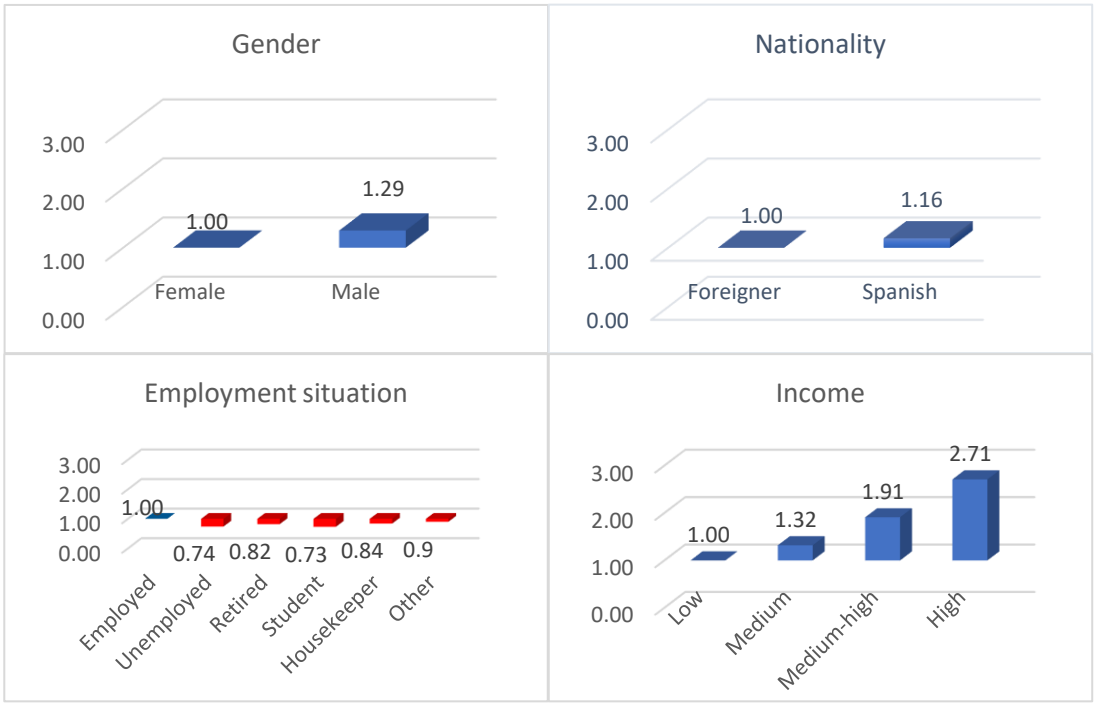
		Coef.	z
Gender	Female		
	Male	0.04	1.1
Age	16-24		
	25-34	**-0.18**	**-2.25**
	35-44	**-0.36**	**-4.54**
	45-54	**-0.56**	**-6.97**
	55-64	**-0.7**	**-8.44**
	> 65	**-0.87**	**-9.02**
Education	Primary or less		
	Secondary	**0.11**	**2.51**
	Bachelor	**0.33**	**5.62**
	Master/PhD	**0.59**	**9.09**
Habitat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

		Coef.	z
Computer availability	No		
	Yes	**1.04**	**29.34**
Year	2008		
	2009	**0.17**	**2.53**
	2010	**0.2**	**3.18**
	2011	**0.26**	**3.79**
	2012	**0.45**	**6.35**
	2013	**0.3**	**4.4**
	2014	**0.55**	**7.84**
	2015	**0.57**	**7.58**
	2016	**0.59**	**8.67**
	2017	**0.94**	**12.84**
Autonomous Community	Andalucía		
	Aragón	0.07	0.92
	Asturias	-0.03	-0.45
	Baleares	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	**-0.15**	**-2.09**
	Valencia	0.05	0.8
	Ceuta	-0.04	-0.2
	Melilla	-0.1	-0.62
Constant		0.86	7.48

N. observations	**77,362**
Wald χ^2	**2631.86 DF: 53**
Pseudo R^2	**0.2363**
Correctly classified	**97.08%**
Wald χ^2: H_0: independent equations	**54.96 DF: 1**

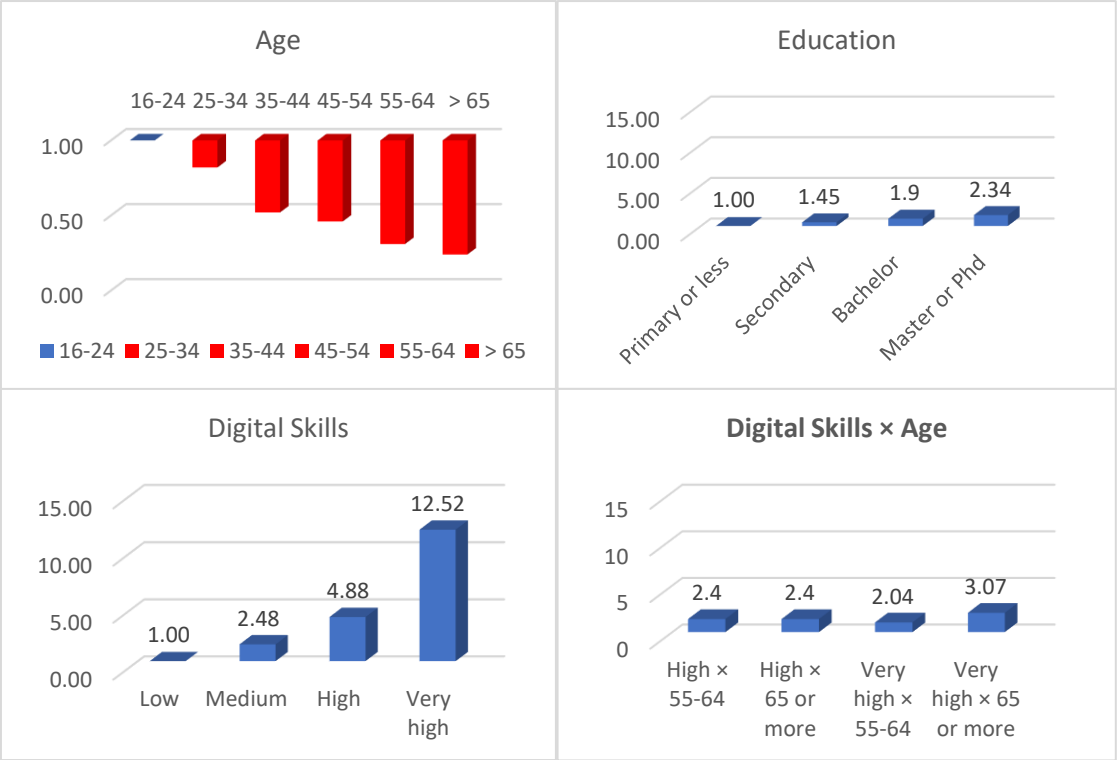
Notes: Coefficients and statistics (t and z) significant at the 5% are represented in bold. Heckman Selection's first step is estimated. Robust estimates. Pseudo R^2 is a measure of the in-sample goodness of fit. The relevant statistic for testing the significance of the estimated coefficients is the χ^2 test of joint significance of all slope coefficients equal to zero, which is high and has a 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)



Source: Own elaboration based on estimation results. *Table 2 (2)*.

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



Source: Own elaboration based on estimation results. *Table 2 (2)*.

CHAPTER 4:
E-commerce and digital divide in Spain using individual panel data 2008-2016[40]

[40] Manuscript submitted for publication (2019); with Pérez-Amaral, T., López, R., Garín-Muñoz, T., & Herguera, I.

Abstract

Digital technologies sometimes create digital divides. One of the remedies for certain divides in Europe is the creation of the Digital Single Market, of which e-commerce is one of the main elements. The focus of this work is e-commerce in Spain. The current study improves substantially on existing international literature by using a large and representative panel data set on individuals, with 133,420 observations for the period 2008-2016. Moreover, it uses economic models that incorporate previously omitted variables and employs a variety of panel techniques. This paper starts by measuring digital divides and their evolution along time. Next, a model that incorporates previously neglected explanatory variables, such as income and digital skills, is formulated. Individual demand models are estimated using panel logistic regression techniques. This allows quantifying the impact of each of the socioeconomic and geographic characteristics on the adoption of the service. The resulting models have high explanatory power. Newly incorporated variables like income and digital skills are highly significant. Age, education, gender, and geographical variables are also significant. The results also allow novel regional comparisons. Policy recommendations are derived, suggesting effective and affordable measures targeted at specific socio-demographic groups.

Keywords: e-commerce; digital divide; TIC-H survey; individual annual panel data; logistic regression; dynamic model; Heckman selection model.

JEL classification: C25; D11; O33

Acknowledgments: Presented at the ITS European Conference, 2018, Trento. Research 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.

Comercio electrónico y brecha digital en España, utilizando datos de panel de individuos 2008-2016[41]

Resumen

Las tecnologías digitales en ocasiones crean brechas digitales. Una de las soluciones para ciertas brechas en Europa es la creación del Mercado Único Digital, en el cual el comercio electrónico es uno de sus elementos esenciales. Este trabajo está enfocado en España. El presente estudio mejora sustancialmente los trabajos existentes a nivel internacional, utilizando una extensa y representativa base de datos de panel, con 133.420 observaciones para el período 2008-2016. Además, se utilizan modelos económicos que incorporan variables previamente omitidas y se emplean técnicas de datos de panel. Este artículo empieza midiendo las diferentes brechas digitales y su evolución a lo largo del tiempo. A continuación, se formula un modelo incluyendo variables explicativas que anteriormente se habían ignorado, como nivel de ingresos y las habilidades digitales. Los modelos de demanda individuales se estiman utilizando técnicas de regresión logística de datos de panel. Esto permite cuantificar el impacto de cada una de las características socioeconómicas y geográficas en la adopción del servicio. Los modelos resultantes tienen un alto poder explicativo. Las variables incorporadas como ingresos y habilidades digitales son altamente significativas. Edad, educación, género y variables geográficas también son significativas. Los resultados permiten novedosas comparaciones regionales. Se proponen recomendaciones de política económica, sugiriendo medidas efectivas y asequibles, dirigidas a grupos sociodemográficos específicos.

[41] Manuscript submitted for publication (2019); with Pérez-Amaral, T., López, R., Garín-Muñoz, T., & Herguera, I.

Introduction

The Internet is one of the most powerful agents of change in recent history. Since the world-wide-web opened up the Internet for everyone in 1991[42], its diffusion has been so rapid around the globe and across new services that it has given rise to major changes in everyday life. According to recent data provided by the International Telecommunications Union (ITU, 2016), almost half of the world's population is using the Internet. However, there are important differences according to countries and groups of individuals which are called digital divides. The definition of digital divide has evolved considerably over time. The National Telecommunications and Information Administration, NTIA, (1995) focused first its studies mainly on infrastructure access. Later, on the access to technology and services. There are also other facets of interest, such as the use of the Internet and the quality of use. The relevant definition for this paper is the one that considers the digital divide as the unequal use of the Internet between distinct groups of individuals (Hilbert, 2011).

E-commerce[43] is an important service of Internet and a European Union priority because it has social benefits, promotes competition, and advances towards the objective of a digital single market while avoiding digital exclusion[44]. Another reason for focusing on e-

[42] There is broad consensus in situating the birth of the World Wide Web in 1991 when Tim Berners Lee created it, combining three existing resources (HTML ─Hypertext Markup Language, HTTP ─Hypertext Transfer Protocol, URL ─Uniform Resource Locator) with a new program called Web Browser.

[43] Following Eurostat (2017) and INE (2017), the standard definition of e-commece is the placing of orders of goods and services via the Internet, excluding orders via manually typed e-mails or text messages. Only purchases made for personal reasons are considered.

[44] It is worth pointing out that digital exclusion refers to a variety of Internet services, not specifically to e-commerce.

commerce is the relatively low penetration in Spain versus selected European Union countries.

Specifically, e-commerce has had important economic and social implications. Customers are changing their consumption habits, and businesses need to adapt to the new situation modifying their business and marketing strategies. This situation is evolving rapidly, and it could be considered a sort of social revolution. At this point, some authors wonder whether conventional shopping centers will survive or are doomed to disappear. In the early days of e-commerce, many consumers searched for information and products on the Internet, but they ultimately went to the physical store to shop. Now in some cases, the situation is the opposite: some individuals use the physical store as a catalog where to feel and try the products that may then be bought online[45].

The main objectives of this paper are twofold: first to measure the e-commerce digital gaps in Spain and categorize them by sociodemographic and regional groups, and second to identify and estimate the effects of the determinants of the e-commerce adoption by individual consumers as well as formulate policy recommendations to narrow the relevant gaps.

This paper analyzes e-commerce using microdata on individual consumers from the "Survey on Equipment and Use of ICT in Households" (INE, 2018) which is similar to those carried out in the European Union and allows interesting comparisons across the 28 countries. Dynamic logistic models of adoption of e-commerce in Spain are estimated for the first time using individual panel data.

The current paper improves substantially on existing literature by using a 9-year-individual panel, providing disaggregated measures of digital divide, and using an explicit and richer economic model for individual panel data. Additional relevant variables, such as individual-level income, individual digital skills, and geographical variables, are included

[45] Nevertheless, in Spain there is a high density of shops that are close to the final user, so this may be a partial explanation for the lower e-commerce figures.

as well as a variety of panel estimation techniques. The focus here is on specific individual consumer policy recommendations. The results reported may be amenable to further exploitation by other researchers.

The rest of the paper is organized as follows: Section 2 contains a literature review. In Section 3, the model is presented. Section 4 introduces the data, highlighting the construction of the panel database on individuals from the original data on dwellings. Section 5 presents the estimation results and discussion. The main conclusions, policy recommendations, caveats, and further research are shown in section 6.

Literature review

E-commerce has been an important subject of study as well as a political priority in the European Union and other international institutions; see, for instance, OECD (2001). There is a great amount of literature on the topic, which is classified below as three subsections: international literature on digital divide, international studies on e-commerce and studies on e-commerce and the Internet for Spain. In this section, we include some documents that may look somewhat old but are considered seminal papers in the area and have significant value as references for the current work. Producing a comprehensive survey of this literature is beyond the scope of this research.

International literature on digital divide

A seminal report by the National Telecommunications and Information Administration (1999) discusses the digital divides[46] (telephone, computer, and Internet use), and their

[46] "The gap between individuals, household, business and geographic areas at different socio-economic levels with regard to their opportunities to access information and communications technologies and to their use for a wide variety of activities" (OECD, 2001, pp. 8-9).

evolution in the US using data for the previous 15 years. They define digital divide[47] (absolute and relative), and they relate it with its evolution along time. The way of interpreting the level and evolution of the digital divide is similar to the one adopted in the current study.

Hilbert (2011) focuses on digital divides and their definitions. The author considers the adoption of a new service as a contagious disease instead of an economic decision that would depend mostly on income, price, and other factors. This paper suggests that there are heterogeneous digital divides that cannot be added up in a significant way since they are difficult to synthesize in a single index. This paper poses relevant questions that would need more specific answers.

Demoussis and Giannakopoulos (2006) deal with the determination and extent of Internet use as a facet of the digital divide in Europe. They use 2002-2003 individual cross-section data from a variety of European countries and focus on the decision to use or not use the Internet as well as how much to use it. This paper contributes with very useful discussions on the right way to account for the availability of the Internet and the relevant population groups concerned. This insightful study, however, is missing a measure of digital skills, so the estimates of the coefficients of the rest of the variables may have omitted variable biases.

A recent paper by Răileanu (2018) deals with the regional digital divide in the European Union. It uses panel data specific models aggregated at the levels of regions and countries, but no formal theoretical model is used. Variables such as digital skills are absent, and the level of education turns out to be insignificant in some models. The endogeneity of other explanatory variables may be an issue. The interpretation of the signs of the coefficients of several variables seems to be problematic in several cases.

[47] A digital divide can be regarded as a situation in which the demand for access or use of a given digital service by a specific group is considered as insufficient (by policy makers or researchers), compared to that of other group of reference.

The levels of the digital divide in firms have been studied in Bach et al. (2013). They review several papers that can be classified into three groups depending on the phenomena used as a measure: ICT usage, adoption of Internet and broadband adoption, and, ICT usage for specific business purposes. External and internal factors are identified as determinants of the digital divide among firms.

Concerning the indices used to measure the digital divide among individual consumers, Barzilai-Nahon (2006) criticizes policymakers who rely on simplistic measures. Besides presenting a conceptual definition of the digital divide, the researcher supports the use of comprehensive indices rather than monotopic ones. The article highlights the context as the most important framework for both the conceptualization of the digital divide and the construction process of an index to measure it. The author emphasizes that policymakers should embrace the analysis of the purpose of the tool, the level of observation, and the methodological approach to the data.

To understand a complex subject such as digital divide, Vehovar et al., (2006) argue that standard methodological approaches are not enough, and the measures must be considered in a multivariate setting. Three advanced methodological approaches are proposed to measure digital divides. 1) Multivariate log-linear modelling allows addressing interactions among variables. 2) Compound indices that integrate several variables into a single indicator. And, 3) Time-distances methodology, to analyze changes in digital divide across time. The authors conclude that implementing these sorts of analysis yields often different conclusions compared with the usual bivariate comparisons.

Hacker & Mason (2003) argue that if researchers avoid the problem of ethical indifference concerning the analysis of data and pay more attention to the links of ethics with their findings related to digital divides, a better quality of research and policy considerations are likely. The paper claims that most analytical works on the digital divide tend to neglect ethical discussions, which “contaminates” studies and reports with ideological filtering.

Ruiz-Rodríguez et al. (2018) conduct a comparative analysis on the adoption and use of ICTs in enterprises and on the digital divide between them at a regional scale in Europe.

Salemink et al. (2017) address the relationship between digital developments and rural development. This work is a systematic literature review of 157 papers, focusing on the general conclusions to grasp the potential impact of the coming Next Generation Access revolution. The paper reports that connectivity and inclusion are the two major aspects. A relevant remark of the paper is that rural communities are less connected, and they are the most in need of better digital connectivity to compensate for problems associated with their remoteness.

Zoroja (2011) studies the digital divide between European developed and post-communist countries. Internet usage, e-commerce, and e-government are at the center of the analysis. The study reports the average penetration rates for the three areas considered, aggregating the measures for the developed European countries and Post-communist countries. Based on the comparison among groups of countries, the article concludes that the group of developed countries have significantly better access and shows higher intensity in use than the group of post-communist countries. The work also discusses digital divides and link them to socioeconomic factors, especially education.

International papers on e-commerce

Limayem et al. (2000) study e-commerce empirically. They use the Theory of Planned Behavior with a longitudinal sample of 705 consumers from the US. Their methodology is not directly comparable with that of the present work. Theirs is based mostly on unobservable variables, while ours deals with observable variables.

Hackl et al. (2014) focus on the interaction between market structure and market performance of e-commerce centered on supply-side considerations of the retail sector for the case of Austria, using cross-section data.

The European Commission (2018) shows that Spain has an average level of digital development within Europe. The main barriers for development are in the demand side, in particular, the shortcomings in human capital related to ICT as consumers. The document contains a useful account of recent and current programs for the advancement of the Information Society in Spain, both from the supply and the demand sides.

Based on internationally comparable micro-aggregated data for the period 2002-2010, Biagi & Falk (2017) present empirical evidence against the hypothesis that the utilization, by firms, of complex and sophisticated forms of ICT and internet applications is leading to labor substitution overall.

A multivariate perspective is adopted by Billon et al. (2016) to study the links and determinants of ICT use by households and firms in the EU. Using canonical correlations analysis, the authors investigate the presence of regional patterns that combine ICT use by firms and by households in the EU. The research also explores the drivers of the patterns detected, identifying synergies and regional factors explaining the use of ICT both by households and by firms in the EU.

Bose and Luo (2011) go a step further at analyzing information technologies adoption by firms, proposing a different perspective from an environmental point of view. The main goal of this work is to contribute to the literature with a model that incorporates technological, organizational, and environmental factors linked to the firm's potential to undertaking Green IT initiatives via process virtualization.

Valarezo et al. (2018) identify two dimensions of trust among the factors driving the adoption of cross-border e-commerce by individuals in Spain: trust in the product and supplier, and trust in the channel. Oliveira et al. (2017) conduct more specific research on the topic of consumer trust, modelling, and assessing three dimensions: competence, integrity, and benevolence. Based on a structural equation model, using a sample of 365 individuals, this last paper discusses how consumer and firm characteristics, and website infrastructure and interactions with consumers are sources of trust. Both works coincide that the higher the consumer trust, the higher the intention to purchase online.

Papers on e-commerce and Internet use in Spain:

Previous studies on e-commerce or closely related topics that use cross-section data of the ICT-H survey for Spain are:

Cerno and Pérez-Amaral (2006, 2009) deal with Internet access and use as well as e-commerce use in Spain using cross-section data on individual consumers. Garín-Muñoz

and Pérez-Amaral (2011) concentrate on the factors affecting e-commerce use in Spain using cross-section data on individual consumers from the ICT-H survey.

Garín-Muñoz et al. (2019) specify and estimate binary response models for individual consumer adoption of e-commerce, e-banking, and e-government in Spain, using cross-section data obtained from the ICT-H survey of INE for 2016. This paper discusses the signs and effects of demographic and socioeconomic individual characteristics on the probability of adoption of the mentioned digital services.

Using data from the ICT-H survey of INE for the year 2016, Valarezo et al. (2018) study the drivers and barriers of Spanish individual consumer adoption of cross-border e-commerce for private use. In addition to the socioeconomic and demographic variables, this work includes the variable "how often the consumer sees other customer reviews before buying online", to account for the effect of trust in the product and vendor.

Pérez-Hernández, J. and R. Sánchez-Mangas (2011) study online shopping jointly with having Internet at home using the ICT-H survey of INE in Spain for the period 2004-2009 employing pooled individual data. In this paper, they cannot control for unobserved individual heterogeneity since they use pooled instead of panel data. Likewise, they omit relevant variables like individual digital skills, geographic variables, and individual income (which are available only from 2008 onwards). This omission can cause inconsistency in the estimates as recognized by the authors in p. 221.

Robles and Torres-Albero (2012) analyze individual cross-section data from INE for the year 2009 to calculate penetrations of the use of the Internet. They conclude that "…between the most advanced communities and the communities with the lowest percentage of users these differences, far from decreasing, have remained stagnant or even increased slightly over the last five years". They also model the decision to use the Internet. Their logistic model does not control for individual heterogeneity and fails to include two variables such as digital skills and individual income that were available at the time. This omission leads to inconsistent estimates and an upward bias in the estimation of the effect of "level of education," which is positively correlated with the two omitted variables. These flaws undermine some key conclusions of the paper.

A report by Correa et al. (2015) of BBVA Research deals with previous waves of the same TIC-H survey of INE for Spain during the period 2008-2014. They study the adoption of broadband and e-commerce. They use pool data, not a panel, so they cannot control for unobserved individual heterogeneity. They do not use an explicit economic model, but they do use time dummies and geographical variables. Their conclusions are subject to caveats due to the type of sample and model they use and the importance imposed on e-banking.

A recent report by Fundación BBVA (2018) found that the digital divide in Spain disappeared in 2017. However, they only analyze, at an aggregate level, the digital divide in access and use of the Internet but not e-commerce, like the present study does, which is more relevant for this paper. They use descriptive statistics as well as intuition, but they do not use economic or econometric models. The effect of mobile broadband access is not explicitly considered.

Another study by Burgos et al. (2018) focuses on e-commerce in Spain. It uses a highly mathematical model based on epidemiology but fails to consider the economic nature of the decision to adopt e-commerce as well as the findings contained in previous literature. The treatment of the data seems perfunctory. The multiple factors that may affect the adoption of e-commerce are largely ignored.

The above papers are defective in one way or another. Some use aggregate data only, while others do not use economic models. Some others use cross-section or pool data, instead of panel data, while others omit relevant variables such as income and digital skills. Due to these limitations, the policy recommendations are limited or subject to caveats. In the next sections, we present results that fill these gaps in the existing literature.

The model

The present study follows an economic perspective using the neoclassical utility maximization approach (Varian, 2002). The demand for access is determined by the size of consumer surplus associated with Internet usage and the cost of access. Regarding access

to e-commerce, the relevant theory is that of the telephone demand framework of Artle and Averous (1973), Squire (1973), Von Rabenau and Stahl (1974), Rohlfs (1974), Taylor (1994), Kridel et al. (1999) and Rappoport et al. (2002).

In telecommunications, the use of a specific service is conditional on access to this service, (Taylor (1994)). The current approach assumes that Internet access is a prerequisite for adopting e-commerce. In any case, Internet access could be obtained in the period considered through a variety of channels and places: buses, trains, airports, ships, work, home, school, university, hotels, restaurants, public Wi-Fi zones, community access centers, libraries, post offices, internet parlors, as well as using a variety of technical solutions: dial-up, cable, ADSL, broadband, narrowband, or through mobile phones, tablets and portable computers. Summing up, Internet access has been ubiquitous during the years of the sample 2008-2016.

In many cases, access to the Internet is not an explicit decision, but rather a circumstance governed by the commercial policies of carriers that incorporate Internet to a traditional service, even without explicit knowledge by the consumer. A similar argument is sustained by Demoussis and Giannakopoulos (2006) for the European case using 2002-2003 data. The argument is compelling in the sample considered. Nevertheless, in the Appendix, *Table A1*, a model incorporating a Heckman-style mechanism is estimated and presented, suggesting that the selection mechanism is not necessary in this case, and model 1 is an adequate approximation.

When access to the Internet is widespread, the decision to use the Internet no longer needs to be modeled. The hypothesis is that consumers decide to use e-commerce, given that they have access to the Internet[48].

In this context, an individual derives utility (U) from adopting a particular Internet service (Y), if the benefits from using that service B(Y) exceed its costs C(Y). Empirical works based on this approach are Demoussis and Giannakopoulos (2006); Fairlie (2004); Vicente

[48] Effective use by 100% of individuals across the population cannot be expected, since there are people who are severely ill, physically or mentally handicapped, very old, very young, and minorities for whom Internet may not be attractive.

and López, (2008); Lera-López et al., (2011) and Valarezo et al. (2018, 2019), the last four referred to the case of Spain.

From a standard neoclassical utility optimization approach, the maximization of the utility (U) of an individual obtained from e-commerce (Y_i), will be a function of the benefits $B(Y_i|x)$ of doing so and the costs, $C(Y_i|x)$, where x is a set of conditioning variables associated with it. The conditional probability of using e-commerce is:

$$\boldsymbol{P(Y_i|x_i) = P[B(Y_i) - C(Y_i) > 0|x_i] \quad i = 1, \dots, N,\ individuals.} \qquad \textbf{(1)}$$

The individual consumers considered are those e-commerce users who purchased online goods or services for private use in the last year with respect to the total population.

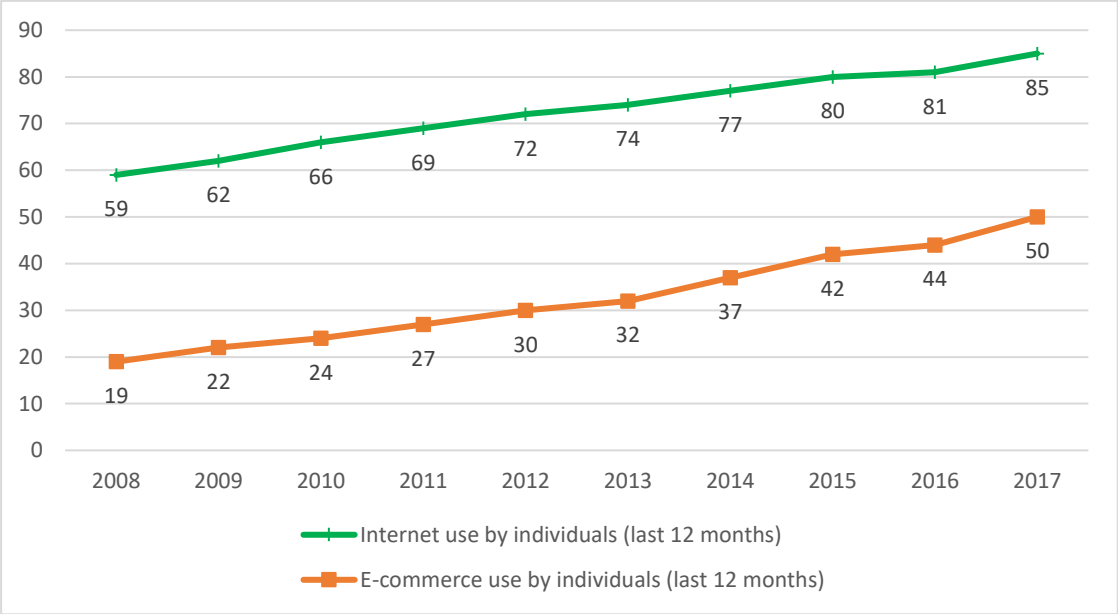
The model used in this paper follows Valarezo et al. (2018) and Garín-Muñoz et al. (2019), where binary models of e-commerce adoption are estimated.

Data

Figure 1 shows the evolution of the aggregate use of the Internet and e-commerce in Spain for the years 2008 to 2017, obtained from Eurostat (2017). The penetration of both services grows over time. The percentage of Internet users increases from 59% to 85% of the population aged 16 to 74. Likewise, the percentage of the population that uses e-commerce grew from 19% to 50% of those aged 16 to 74[49], which suggests that there has been a strong increase, but that there is still room for improvement.

[49] The data of our survey, INE (2018), contains information on people of all ages, including those over 74 years of age. However, the data in this section contains information only on people up to the age of 74, for compatibility with the corresponding data of Eurostat.

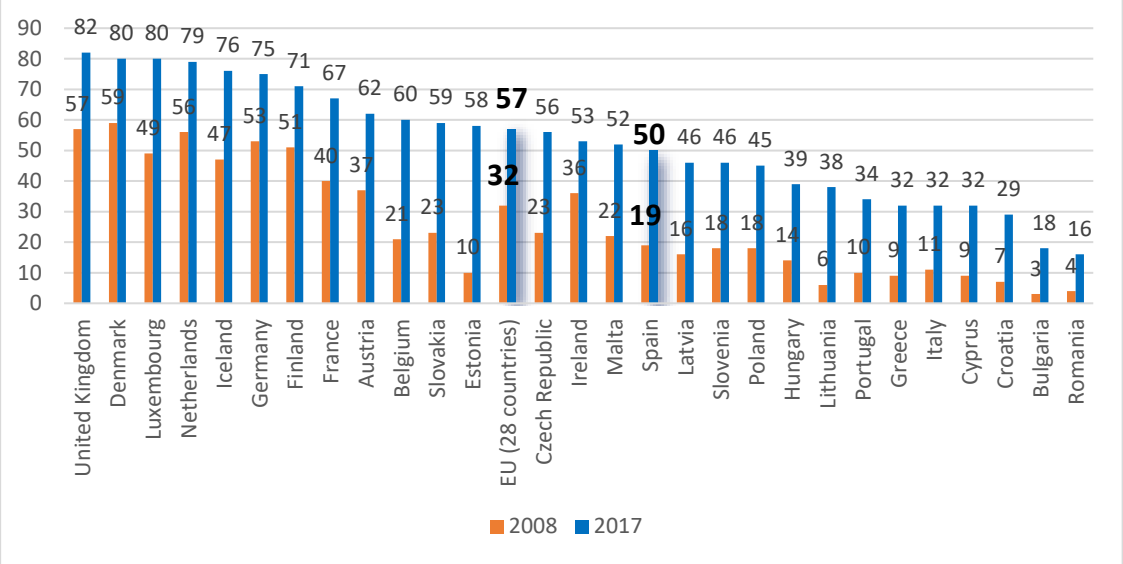
Figure 1:
Internet and e-commerce use as a percentage of people aged 16 to 74 in Spain (2008-2017)



Source: Own elaboration based on data from INE, 2018. Survey on Equipment and Use of Information and Communication Technologies in Households.

Figure 2 shows an international comparison of the penetration rates of e-commerce for European Union countries and Spain in 2008 and 2017. The European Union average increases from 32% to 57% in the period, while Spain increases from 19% to 50%. The increase in penetration is 31pp in Spain and 25 pp for the European Union between 2008 and 2017. The gap is closing significantly between Spain and the European average by 6pp in the period considered 2008-2017. However, there is still a significant gap for the first seven countries on the left of *Figure 2*, which are those that could be the reference for the policy objectives of Spain. This gap is still around 28 pp in 2017, when comparing it with the average of the seven countries.

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



Source: Own elaboration based on data from Eurostat, 2018. Penetration rates as the percentage of people aged 16 to 74.

This work starts by using annual data on individual dwellings from the survey on Equipment and Use of Information and Communication Technologies in Households (ICT-H Survey) from 2008 to 2017.

The survey is conducted by the Spanish National Statistical Institute (INE, 2018) using the methodology of Eurostat. It includes an elevation factor, and it is representative at both national and regional levels. The raw data are available at INE (2018), as well as a variety of documents related to the design, implementation, and statistics (except for the individual identifier). It is a rotating survey that includes 15,000-20,000 dwellings each year. The same dwelling is interviewed a maximum of four (consecutive) years, and around 30% of the dwellings are replaced every year. 60% of the interviews are conducted by phone (CATI) and 40% in person (CAPI). Following Eurostat's guidelines, similar surveys have also been performed in the rest of the EU countries.

During the writing of this paper, Pool data were available for 2008-2017, but panel data could only be used for the period 2008-2016 due to the lack of a dwelling identifier for

2017. The raw panel data underwent meticulous analysis and filtering to extract and homogenize the information about individuals uniquely identified throughout the observation period. The process goes as follows:

Pool data

Read the information on dwellings for each year (publicly available from INE, see INE 2017). For each year, there is a zip file containing a raw data file with the sociodemographic information of all the members of each dwelling, a raw data file with the information of the dwelling's responses to the survey, and an Excel file with the full description of both raw data files.

Homogenization of variables across years: the survey's questionnaire varies each year according to the evolving situation of ICT. This implies that some new variables may emerge, and old ones may disappear. Besides, each year, the names of most variables change (in the publicly available microdata, most variables are named after their ordinal order in the questionnaire –which usually changes). To homogenize the variables, we developed a script in R that scraps the description excel files in order to identify, across years, the same variables with different names and different variables that were assigned the same name.

Panel data

Incorporate the dwelling identifier supplied by INE to the previous pool database. However, we are interested in identifying individuals rather than dwellings (a dwelling participating in the survey several years may have different respondents across years). Using the sociodemographic information available for each member of a dwelling, we applied filters (like gender and date of birth) to identify whether the respondent of a particular dwelling was always the same individual or not.

In the end, we obtained a Panel 2008-2016 database consisting of 133,420 observations (corresponding to 59,252 different individuals) and around 700 different variables.

This was the final micro panel database that was used in all of our models.

The explanatory variables used are classified as follows:

Sociodemographic	
Gender:	2 groups: 1 if male, 0 if female
Age:	6 groups
Habitat:	4 groups
Household Members:	5 groups
Nationality:	2 groups: Spanish, Foreign
Individual skills	
Education:	4 levels of study
Digital Skills:	4 levels
Economic	
Employment Situation:	6 groups
Income:	4 groups, monthly net income
Time and Geographic	
Yearly Dummies:	1 for each year
Regional Dummies:	17 Autonomous Communities + 2 Autonomous cities

The variables were categorized as shown in *Table 1*, which also contains the penetrations of e-commerce on the different groups, measured in the years 2008 and 2017. The table suggests a general increase in the penetrations in the period considered.

Digital skills are among the key elements to foster digital inclusion (European Commission, 2019b). The digital skills variable used in this study is a synthetic index based on the former[50] European Commission's Digital Skills Indicator, which accounts for the number and complexity of activities involving the adoption and use of digital equipment and

[50] The later indicator includes information about ICT specialist employment and ICT graduates. This data is not available for the sample of this study.

internet services (European Commission, 2019a). The index covers four areas of competence: information and data literacy, communication and collaboration, problem-solving, and software skills for content manipulation.

Answering specific questions in each area, the respondent reveals his or her competences. Information skills area accounts for individuals who have copied or moved a file or folder, used Internet storage space, search for information online, search for information about goods and services, search for information about health-related topics. Communication and collaboration areas are approximated by the use of email, social networks, telephone calls through the Internet, and web sites for sharing own content. The problem-solving area of competence regards to transfer files between devices, installing software and apps, change the settings of the software, selling of goods or services through the Internet, taking online courses, using online education material, and carrying out online banking activities. The software skills for content manipulation area includes using text processing software, spreadsheet software, software to edit audiovisual content, creating documents that integrate different files, writing a computer program using specialized programming language, and using spreadsheet's advanced functions.

In this work, four categories of digital skills are used: low, medium, high, and very high. These correspond to the quartiles of the percentage of tasks that an individual consumer declares to be able to perform.

Table 1 reports the e-commerce penetration rates for different groups, which represent, among others, the digital skills divide. For the considered period substantial improvements are seen for groups with medium, high and very high levels of digital skills, but the same is not true for those with low levels of digital skills. This is one of the main concerns of the European Commission (2019b). The DESI report for Spain identifies that the human capital dimension needs to be improved in order to bridge the gap respect the EU average.

Table 1:
Penetration rates of e-commerce in Spain (2008 and 2017)

Characteristics	Categories	2008	2017
Gender	Female	14.2	41.4
	Male	20.5	48.5
Age	<25	25.1	64.5
	[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	1.4	6.5
	Secondary	16.6	44.7
	Bachelor's Degree	29.7	71.1
	Master´s or PhD	42.6	81.8
Digital Skills	Low	9.8	12.3
	Medium	25.1	44.5
	High	47.4	71.2
	Very High	69.2	89.8
Habitat	500,000+	22.2	51.2
	100,000-500,000	18.2	42.2
	20,000-100,000	16.3	45.4
	<20,000	13.0	38.7
Household Members	1	12.4	31.1
	2	14.6	32.5
	3	18.1	51.2
	4	21.5	58.7
	5+	14.9	43.0
Employment Situation	Employed	25.1	63.9
	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

Table 1 (continued)
Penetration rates of e-commerce in Spain (2008 and 2017)

Characteristics	Categories	2008	2017
Nationality	Foreigner	14.5	41.0
	Spanish	17.7	45.3
Income	Low	3.6	23.6
	Medium	11.8	37.5
	High	28.3	61.0
	Very High	41.0	79.7
Autonomous Community or Autonomous City	Andalucía	13.1	41.0
	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	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
	Ceuta	12.6	48.6
	Melilla	17.9	47.8
TOTAL		17.3	44.9

Note: percentage of individuals between 16 and up that have used ecommerce for private purposes at least once in the previous 12 months.

The following map can be used as a reference for the geographic location of the above Autonomous territories.

Figure 3:
Autonomous Communities of Spain.

Source: Adapted from Wikipedia, 2019.

Graphs of Figure 4 shows selected entries of *Table 1* and the values of the relative digital divides. It shows the evolutions of the penetration of e-commerce by selected categories, also the absolute divide in perpendicular. It is worth pointing out that the absolute digital divides for gender, age, education, and income increase along time. They are measured in terms of the absolute difference between the two rates of penetration in 2008 and 2017[51].

[51] An alternative measure of the relative divide consists in dividing the difference between the two categories at a given point in time over the highest of the two rates of penetration. This measure can give different results from the absolute measure. An additional measure of divide can be the ratio of penetration rates. That is, the ratio between two penetration rates of two different categories at a given point in time, using the highest as

The absolute age divide increases by 31.1 pp, while the gender divide increases by 0.8 pp, the education divide increases by 34.1 pp, the income divide increases by 18.7 pp, the employment divide increases by 29.7 pp, and the digital skills divide by 18.1 pp along the 10 years.

However, the relative divides measures shown (gender, age, education, income, and employment) decreased by 16.1, 10.4, 4.7, 20.8, and 4.6% respectively.

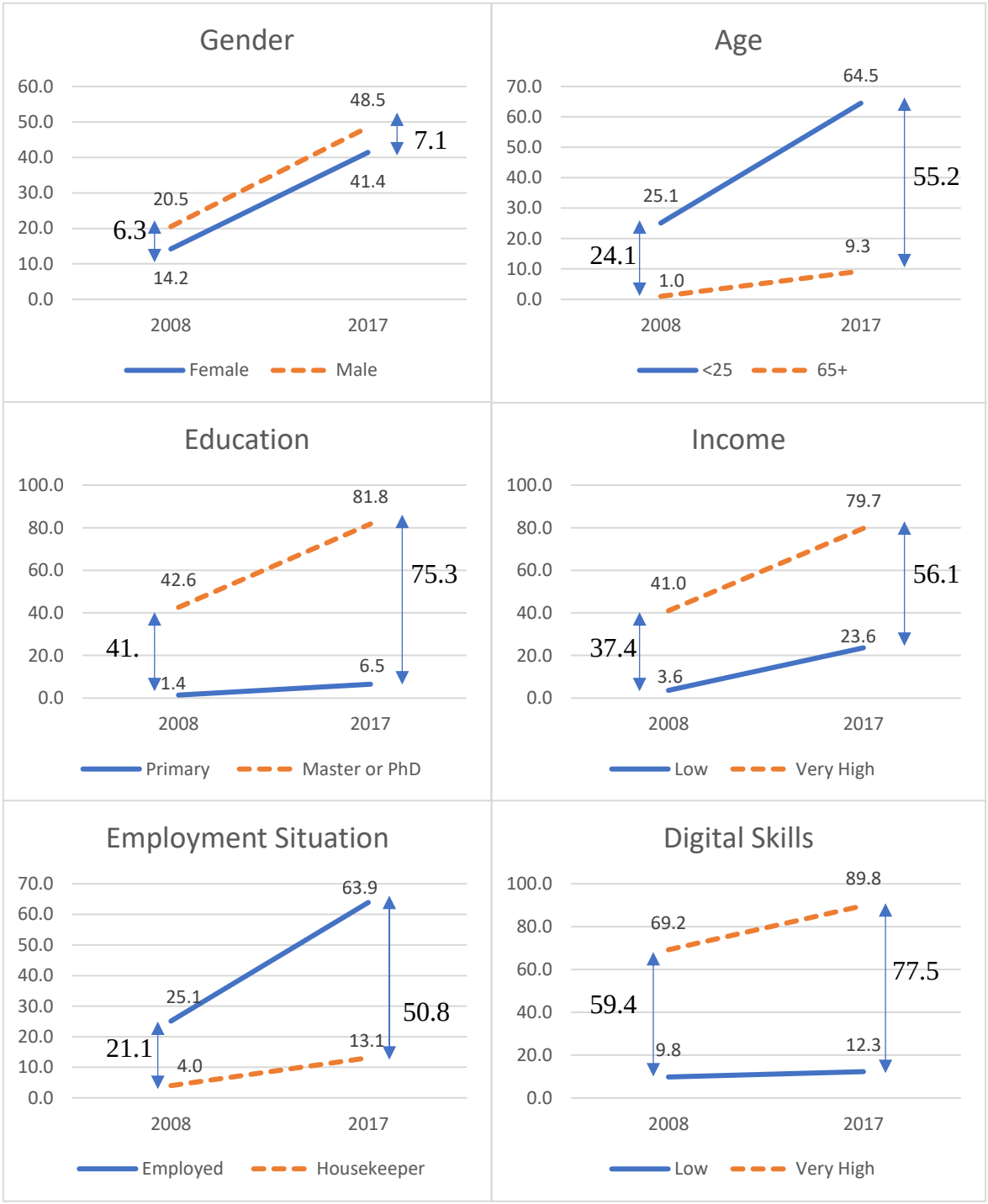
Explanations of the variations in inequality along time and across sociodemographic and geographical groups may be interesting by themselves.

In particular, OECD (2001) deals specifically with the digital divides, although the report does not provide a precise definition of digital divide, or a method to compare digital divides at two different points in time, or its temporal evolution. An additional reference for the specific case of Spain is Robles and Torres (2012). The study of variations on inequality is deferred to further research.

A similar situation occurs when the regional differences are analysed, *Figure 5*. The digital divides between selected regions are shown in *Figure 4*, where the digital gaps between richer and poorer territories seem to be increasing by 4.7, 5.2, 0.7 and 2.9 pp respectively in the specific regions shown. Meanwhile, the measures of the relative gap decrease along time. Notice that the choice of cases shown is not random. The cases that display less convergence are chosen here to underscore the overall convergence of the different groups and regions along time.

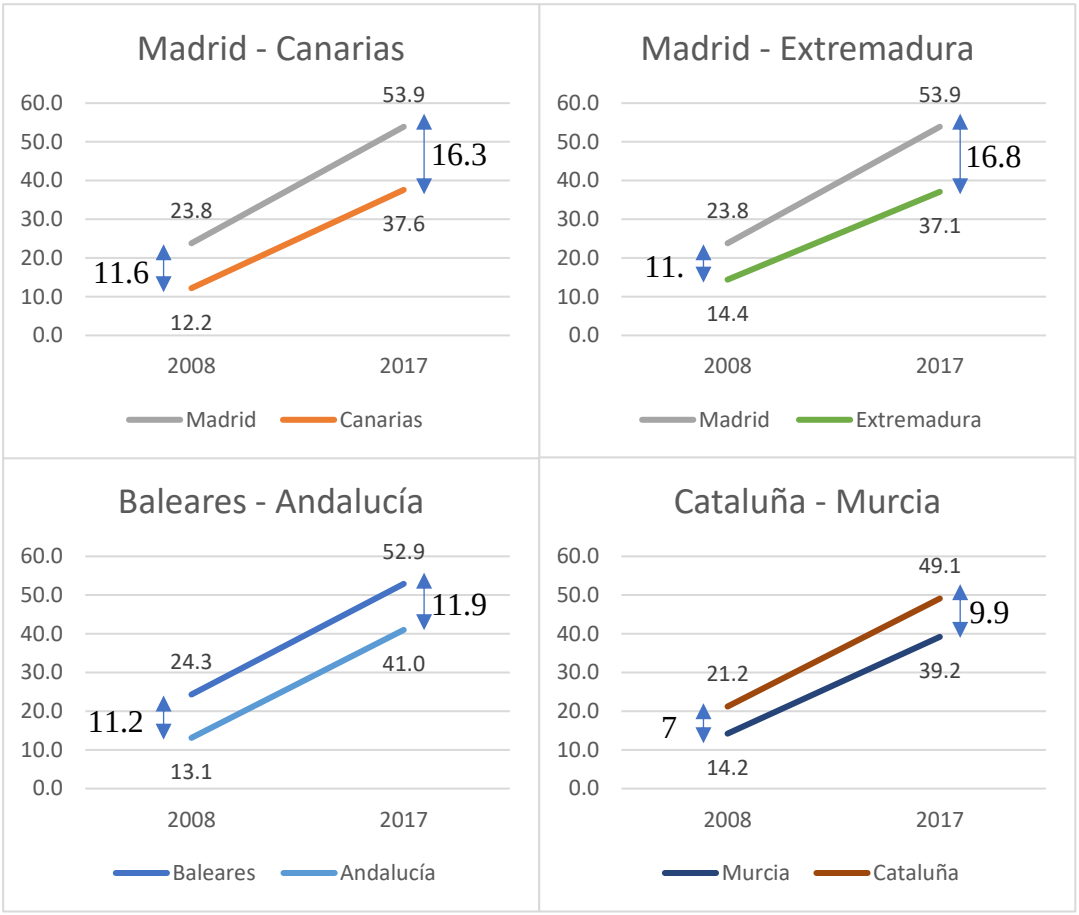
the denominator. Different measures may generate apparent contradictions and different impressions for the readers. Graphs like those in *Figure 3* may be useful to help the researcher or policy maker decide for him/herself.

Figure 4:
Digital divides: E-commerce penetration rates by Gender, Age, Education and Income in Spain 2008-2017



Source: Own elaboration based on data from INE, 2018. Survey on Equipment and Use of Information and Communication Technologies in Households.

Figure 5:
Evolution of selected penetrations of e-commerce by regions in Spain 2008-2017



Source: Own elaboration based on data from INE, 2018. Survey on Equipment and Use of Information and Communication Technologies in Households.

Estimation results and discussion

This section contains the main results obtained from the empirical models. They are displayed in tables for easiness of reading. The main explanatory variables of the models have been introduced in the previous sections. Lagged values of e-banking and e-government, which are meant to capture the dynamics and the cross effects on e-commerce are also introduced as explanatory variables in the second model.

Table 2 contains the two logistic specifications of the model, static and dynamic, including the point estimates of the odds ratios and the corresponding z statistics[52].

Both models are similar in several respects. The signs and sizes of most coefficients and z statistics are similar and have similar interpretations. This can be regarded as a sign of the robustness of the models.

The two lagged variables, e-banking, and e-government are significant with positive effects and with odds ratios of 2.56 and 1.31, respectively. However, there seems to be little gain in the dynamic[53] model, which loses almost half of the observations due to the lags, so we concentrate on the results for model 1.

The estimates are presented in the form of odds ratios; estimated odds ratios with values smaller than one are associated with variables that have negative effects on the adoption of e-commerce relative to the reference group. Odds ratios above one are associated with explanatory variables that have a positive effect on the dependent variable. The individual significance of each coefficient is tested using the z (standard normal) statistic. Corresponding z statistics are calculated from the null of the odds ratio being equal to one (no effect). Odds ratios below one have corresponding negative values of z, while odds ratios above one have corresponding z with positive signs.

The interpretation of the estimation results of the static model is as follows:

Males have higher odds than females of using e-commerce 1.45, which is highly significant. This is revealing of a gender gap that may be disappearing in a relatively short time as women incorporate to e-buying.

[52] All estimations are performed using Stata 15.

[53] The model is dynamic in the sense that the information set at time t incorporates the values of some variables at time t-1.

Table 2:
Models of adoption of e-commerce by individual Internet users. Panel data (2008–2016)

	Variables		(1) E-commerce static logistic model		(2) E-commerce dynamic logistic model	
			Odds ratios	z	Odds ratios	z
Lagged variables	ebanking_{t-1} egovernment_{t-1}				**2.56** **1.31**	**21.37** **6.20**
Sociodemographic	Gender	Female				
		Male	**1.45**	**11.90**	**1.52**	**9.51**
	Age	16-24				
		25-34	0.85	-0.81	0.94	-0.16
		35-44	**0.56**	**-3.12**	0.49	-1.82
		45-54	**0.38**	**-5.03**	**0.38**	**-2.49**
		55-64	**0.23**	**-7.19**	**0.26**	**-3.30**
		65<	**0.14**	**-8.37**	**0.17**	**-4.00**
	Education	Primary or less				
		Secondary	**1.58**	**7.12**	**1.51**	**3.94**
		Bachelor	**2.34**	**11.95**	**1.98**	**5.99**
		Master/Phd	**3.36**	**16.86**	**2.68**	**8.64**
	Digital Skills	Low				
		Medium	**2.56**	**4.90**	**2.74**	**2.55**
		High	**7.95**	**11.22**	**7.16**	**5.17**
		Very high	**25.32**	**17.36**	**24.29**	**8.32**
	Habitat	<20,000				
		20,000-100,000	0.94	-1.49	1.00	-0.03
		100,000-500,000	**0.78**	**-3.82**	**0.73**	**-3.52**
		500,000<	1.00	-0.05	1.03	0.63
	Household Members	One				
		Two	**0.87**	**-2.69**	0.87	-1.78
		Three	**0.73**	**-6.02**	**0.75**	**-3.74**
		Four	**0.71**	**-6.42**	**0.76**	**-3.57**
		Five or more	**0.60**	**-7.15**	**0.64**	**-4.35**
	Nationality	Foreigner				
		Spanish	**1.26**	**3.67**	1.00	-0.02
Economic	Employment Situation	Employed				
		Unemployed	**0.73**	**-6.94**	**0.76**	**-4.21**
		Retired	**0.84**	**-2.16**	0.84	-1.52
		Student	**0.70**	**-4.78**	**0.77**	**-2.38**
		Housekeeper	**0.81**	**-2.79**	0.83	-1.66
		Other	0.83	-1.89	0.93	-0.51
	Income	Low				
		Medium	**1.60**	**10.43**	**1.52**	**6.01**
		Medium-high	**2.58**	**19.04**	**2.42**	**11.74**
		High	**4.00**	**23.43**	**3.67**	**14.73**

Table 2 (Continued)

	Variables		(1) E-commerce static logistic model		(2) E-commerce dynamic logistic model	
			Odds ratios	z	Odds ratios	z
Interaction	Digital Skills × Age	High × 55-64	**2.44**	**4.02**	1.62	1.15
		High × 65 or more	**2.33**	**3.24**	1.53	0.92
		Very high × 55-64	**1.80**	**2.53**	1.21	0.44
		Very high × 65<	**3.09**	**3.90**	1.67	1.05
Time effects	Year	2008				
		2009	**0.84**	**-3.28**	0.87	-1.85
		2010	**1.17**	**2.88**	1.11	1.28
		2011	1.05	0.76	0.91	-1.10
		2012	**1.13**	**2.03**	1.16	1.77
		2013	**1.23**	**3.57**	**1.20**	**2.22**
		2014	**2.09**	**12.57**	**2.01**	**8.25**
		2015	**4.21**	**24.33**	**3.90**	**15.94**
		2016	**4.09**	**23.94**	**3.90**	**16.04**
Sociogeographic effects	Autonomous Community	Andalucía				
		Aragón	1.16	1.81	1.25	1.96
		Asturias	**1.26**	**3.01**	**1.33**	**2.65**
		Baleares	**2.35**	**9.39**	**2.07**	**5.37**
		Canarias	**0.82**	**-2.24**	**0.69**	**-2.86**
		Cantabria	**1.60**	**5.36**	**1.67**	**4.11**
		Castilla la Mancha	**1.26**	**2.71**	**1.29**	**2.09**
		Castilla León	1.04	0.45	1.15	1.28
		Cataluña	**1.55**	**6.58**	**1.40**	**3.60**
		Extremadura	1.12	1.29	1.26	1.79
		Galicia	1.04	0.46	0.97	-0.32
		La Rioja	**1.27**	**2.58**	**1.36**	**2.35**
		Madrid	**1.42**	**5.16**	**1.32**	**2.93**
		Navarra	**1.66**	**6.92**	**1.85**	**5.83**
		País Vasco	**1.79**	**7.27**	**1.72**	**4.83**
		Murcia	0.94	-0.75	0.98	-0.18
		Valencia	0.91	-1.34	0.86	-1.57
		Ceuta	0.72	-1.72	0.73	-1.10
		Melilla	0.98	-0.11	0.87	-0.41
	Constant		0.05	-15.36	0.02	-10.25
	N. observations		66,169		34,032	
	Wald χ^2		7982.03, DF: 69		3704.33, DF: 71	

Notes: Odd ratios and z statistics significant at the 5% are represented in bold. Random-effects logistic model (Equation 1). Random effects dynamic logistic model (Equation 2). Heteroskedasticity robust estimates. Estimations refer to all Internet users that have purchased (or not) online in the last year. The first category of each variable is the reference, and it is omitted. The samples reflect the fact that some variables are missing for some observations, which reduces the usable sample size. The dynamic model loses almost half of the sample due to the lagged variables. All estimations have been calculated using Stata 15.

Meanwhile, age has a negative impact on e-commerce. As the age of the individual increases from one age range to the next, the odds ratios decrease monotonically from 1 for ages 16-24 to 0.14 for the ages of 65 and above. This is also compatible with an age divide, which may be diminishing as new cohorts of young people reach more mature ages. Age, like most other variables, cannot be manipulated by policy, so it cannot be considered a policy instrument in this context. However, identifying age groups with a low likelihood of e-commerce adoption allows more precise targeting of policies.

Education has an important and positive effect measured by odds ratios that go from 1 for primary or less to 3.36 for the case of Master/Ph.D. This, as the previous ones is a pure education effect (not incorporating income) and shows that individual consumers with higher levels of formal education tend to engage in e-commerce with higher likelihood than consumers with lower education.

Next, digital skills have important positive effects, measured by odds ratios that go from 1 for the category of low digital skills to 25.32 for the category of very high digital skills. This is a variable that can be partially manipulated in the short run, and thus, it is the center of intense debate on whether and how to do it more efficiently. This variable can be manipulated through education on digital skills as well as information and technical support. The European Commission (2018) points out that one of the main barriers to the development of e-commerce in Spain is in the demand side, in particular, the shortcomings in human capital related to ICT as consumers.

The effect of habitat is essentially negligible for most of the categories, say below 20,000, between 20,000 and 100,000 and also for above 500,000. However, it is negative for middle size populations between 100,000 and 500,000 inhabitants with an odds ratio of 0.78, which is significantly below 1. This may suggest that inhabitants of middle size populations are less prone to engaging in e-commerce than the rest of the population. It would be very interesting to have more disaggregate information regarding habitats below 20,000 inhabitants to be able to assess the effect of rural and remote areas which may have important differences with the rest of the population.

The number of household members has a decreasing effect, with an odds ratio of 1 for the families of size one, to a value of 0.60 for the families of 5 or more individuals. This is

measuring the effect of the size for a given amount of the other variables, in particular, it would mean lower income per individual for a given amount of income.

Being Spanish has a positive effect reflected in an odds ratio of 1.26. The effect of nationality is not obvious since foreigners could be more prone to cross border e-commerce than nationals. However, the measure of e-commerce used here does not distinguish for origin, intensity, or expenditure on e-commerce just if there has accessed to e-commerce or not.

Those individuals currently employed tend to do more e-commerce (odds ratio of 1) than the rest of the categories, specifically the unemployed, with an odds ratio of 0.73, the retired, with an odds ratio of 0.84, students with an odds ratio of 0.70, housekeeper (0.81) and other (0.83). These results seem intuitive, while the exact estimates and significance cannot be anticipated by previous intuition.

Income has a positive and significant effect. The estimated odds ratios vary from 1 to 1.60, 2.58, and 4.00, respectively, when income increases. The inclusion of income is important in this article. While several previous papers ignored its importance and omitted its effect (but not others like Valarezo et al., 2019, and Garín et al., 2019), it is the typical variable that intervenes in most economic decisions.

Notice that if we are interested in the possibility of doing policy to promote e-commerce, there are few variables amenable to manipulation by policymakers. The variables gender, age, education, habitat, household members, nationality, employment situation, and income cannot be manipulated in the short term. Only digital skills can be altered in the short run for policy purposes. This coincides with the emphasis placed by the European Commission on the deficiency of these specific skills in Spain and the suggestions of putting in place new programs in this area.

The time effects, shown in *Figure 8* of the Appendix are dummies that can be thought of as capturing other demand-side variables, like time-varying effects and the economic crisis, as well as supply-side effects not included in the explicit explanatory variables. The time effects are significant and mostly growing over time.

The first two variables that belong only to the dynamic model correspond to the first lags of e-banking and e-government. They have significant effects, signaling that the adoption of e-commerce may be anticipated by the adoption of e-government and e-banking by a given individual.

The dummy variables corresponding to the 17 autonomous communities (and two autonomous cities) capture specific effects not accounted for in the rest of the model. They would be essentially autonomous community individual effects.

A list of the autonomous communities together with its individual specific odds ratios in descending order are, first, the group of high and significant values: Baleares (2.35), País Vasco (1.79), Navarra (1.66), Cantabria (1.60), Cataluña (1.55) and Madrid (1.42).

Second, the group of middle and significant values of individual-specific effects consists of la Rioja (1.27), Asturias (1.26), and Castilla La Mancha (1.26).

The third group consists of those territories that have individual effects that are not significantly different from the reference (which is Andalucía). These are Aragón (1.16), Extremadura (1.12), Castilla León (1.04), Galicia (1.04), Melilla (0.98), Murcia (0.94), Valencia (0.91), and Ceuta (0.72). The archipelago of Canarias (0.82) is an outlier with a negative and statistically significant effect.

Interestingly enough, the first and last positions are occupied by the two archipelagos of Spain: Baleares and Canarias. This suggests that the effects are mostly due to income and wealth, and not so much by geography, in this particular case.

These effects seem to be capturing a divide which is based on an income and wealth (private and public) divide. Additionally, in some cases, like the Canarias archipelago, the logistic costs may be so high as to substantially limit the supply of certain products that are readily available in continental Spain.

Additional diagnostic of the empirical model

Multicollinearity may be a concern in this model. However, due to the panel structure of the data together with the type of categorical data which are orthogonal within each

variable, the issue is not especially relevant. In any case, polychoric correlations among the relevant explanatory variables have been computed to obtain formal tests, as shown in *Table 3*.

Table 3:
Polychoric correlations among selected independent variables

	Gender	Age	Education	Digital Skills	Household Members	Income
Gender	1.00					
Age	0.02	1.00				
Education	-0.08	0.00	1.00			
Digital Skills	0.07	-0.29	0.43	1.00		
Household Members	0.02	-0.31	-0.02	0.06	1.00	
Income	0.07	0.06	0.46	0.30	0.24	1.00

The values of all off-diagonal elements that lie below .5 suggest that there is a limited amount of multicollinearity between the data of the independent variables.

Conclusions

This paper contributes to the literature on e-commerce. First, by improving on some of the limitations of the models used in previous papers. This is done by using a large and representative panel database of individual consumers and a variety of economic models. Variables such as income and digital skills at the individual level are included here, allowing for a natural interpretation as an income effect and a cost effect. By employing appropriate panel data econometric models, individual heterogeneity can be controlled, and estimates that are consistent and efficient are obtained.

The second contribution is measuring the importance of digital skills, which gives quantitative support to the current EC policy of promoting the training of specific sociodemographic groups. High digital skills have a positive influence on the adoption of e-commerce; insufficient digital skills seem to be partly responsible for the digital divide, as it is highlighted by the European Commission (2018).

The third contribution is to characterize penetrations across different groups of individuals: it can be noted that several digital divides appear for gender, age, education, digital skills, occupation, and income, as seen in *Table 1*. Some absolute digital divides do not tend to close over time while others do, however relative digital divides generally tend to decrease over time.

Based on these conclusions some policy recommendations can be formulated as follows: when promoting e-commerce is a priority, several measures can be implemented at national and regional levels with the restriction of representing low costs on the part of the governments, firms, and citizens. The measures proposed here are low cost, specific, and targeted at individuals or groups, independently of their geographical location.

Focusing first on demand-side measures, for example, may be desirable to reinforce a training program on specific digital demand-side factors, like digital skills (pointed out by the EC as a critical bottleneck for demand) in order to bridge the digital divides. These measures could be focused on females, people over 55, those with low digital skills and homemakers. A complementary measure could be to provide technical support online, by phone or in person, to those groups that are more at a disadvantage. A training program

may be more effective when focused on those that could increase from low to medium level of digital skills.

Focusing on supply-side measures, the government has recently implemented some interesting general measures, like the “digital by default” program (Estrategia 2015-2020), aimed at digitalizing all interactions of public services. It also created Offices of assistance on the use of public services for serving citizens using e-government, see European Commission (2017). Another measure for 2018 is the Grant program that offers basic training in ICT to young digital professionals so that they can gain access to jobs in this sector, see European Commission (2018).

Additionally, central and regional governments could promote complementary services such as e-health and identify and re-edit supply-side programs that were successfully implemented, whether in their territory or others. The administrations could also reduce transaction costs, red tape, and trade barriers. Some territories, like the archipelago of Canarias, still do not have access to some e-commerce transactions that are available in the mainland, possibly due to its long-distance across the Atlantic Ocean and high logistic costs. Guidelines or incentives for efficient e-commerce platforms could also be established.

For the private sector, practical recommendations could be to implement easy-to-use platforms, facilitate legitimate customer reviews for their products and services, and assure more security for payment and transaction processes (which includes the handling of credit card fraud).

These results and conclusions present some caveats and limitations. The data set is large, but it is declarative, not observed data, which limits the quality of the data due to the difficulty of recollection. Moreover, the data are not specifically designed for this research on e-commerce, which also limits the applicability of the results.

An exercise similar to this one using data for services other than e-commerce and each European country is feasible. We leave this for further research.

The continuation of the analysis of the data set for topics like spending in e-commerce, broadband access, cloud services, trans-border e-commerce, and use by children constitute a rich research agenda. An in-depth study of digital divide using individual data and

employing additional indicators is another priority. Clearly stating its definition and formulas for its computation, as well as the calculation of its differences across groups and the interpretation of its evolution over time, is also in the research agenda. Testing the hypotheses that the evolution of a given digital divide is analog to a life cycle model along time is another avenue of research. On the other hand, the availability of data on actual behavior would be useful for defining new priorities of research.

References

Artle, R., & Averous, C. (1973). "The telephone system as a public good: Static and Dynamic Aspects", Bell Journal of Economics and Management Science, Vol. 4, No 1, Spring, 89-100.

Bach, M. P., Zoroja, J., & Vukšić, V. B. (2013). Determinants of Firms' Digital Divide: A Review of Recent Research. Procedia Technology, 9, 120–128. https://doi.org/10.1016/j.protcy.2013.12.013

Barzilai-Nahon, K. (2006). Gaps & Bits: Conceptualizing Measurements for Digital Divide/s. The Information Society, 22(5), 269–278. https://doi.org/10.1080/01972240600903953

Biagi, F., & Falk, M. (2017). The impact of ICT and e-commerce on employment in Europe. Journal of Policy Modelling, 39(1), 1–18. https://doi.org/10.1016/j.jpolmod.2016.12.004

Billon, M., Lera-Lopez, F., & Marco, R. (2016). ICT use by households and firms in the EU: Links and determinants from a multivariate perspective. Review of World Economics, 152(4), 629–654. https://doi.org/10.1007/s10290-016-0259-8

Bose, R., & Luo, X. (2011). Integrative framework for assessing firms' potential to undertake Green IT initiatives via virtualization – A theoretical perspective. The Journal of Strategic Information Systems, 20(1), 38–54. https://doi.org/10.1016/j.jsis.2011.01.003

Burgos, C, Cortés, J-C, Shaikhet, L, & Villanueva, R-J. (2018). "A nonlinear dynamic age-structured model of e-commerce in Spain: Stability analysis of the equilibrium by delay and stochastic perturbations." Commun Nonlinear Sci Numer Simulat 64, 149-158.

Calzada, J., García-Mariñoso, B., Ribé, J., Rubio-Campillo, R., & Suárez, D. (2017). "Fiber deployment in Spain". UB Economics Working Papers 2017/364.

Cerno, L., & Pérez-Amaral, T. (2006). "Demand for Internet Access and Use in Spain", in PREISSI, B. and MULLER, J. (eds.), Governance of Communication Networks, connecting societies and markets with IT, Physica Verlag, Heidelberg.

Cerno, L., & Pérez-Amaral, T. (2009). E-Commerce Use in Spain. In P. Curwen, J. Haucap, & B. Preissl (Eds.), Telecommunication Markets: Drivers and Impediments (157–172). Heidelberg: Physica-Verlag HD. https://doi.org/10.1007/978-3-7908-2082-9_9

Correa, M, García, JR, & Tabanera, A. (2015). "Comercio electrónico y hábitos de consumo en España: la importancia de la banca on-line" BBVA Research. Observatorio de la economía digital.

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, 235-246.

European Commission (2017) Digital Economy and Society Index Country profile (DESI) 2017: Profile of Spain, https://ec.europa.eu/digital-single-market/en/desi

European Commission (2018). Digital Economy and Society Index Country profile (DESI) 2018: Profile of Spain, https://ec.europa.eu/digital-single-market/en/desi

European Commission (2019a). Human Capital and Digital Skills. Retrieved July 10, 2019, from Digital Single Market - European Commission website: https://ec.europa.eu/digital-single-market/en/human-capital

European Commission (2019b). The Digital Economy and Society Index (DESI). Retrieved from https://ec.europa.eu/digital-single-market/en/desi

Eurostat (2017). ICT usage in households and by individuals from Eurostat http://ec.europa.eu/eurostat/cache/metadata/en/isoc_i_esms.htm

Fairlie, R.W. (2004). "Race and the digital divide, Contributions to Economic Analysis and Policy". The Berkeley Electronic Journals, vol. 3, 1-38.

Fundación BBVA (2018). "Uso de Internet en España y la Unión Europea" Esenciales IVIE 29/2018.

Garín-Muñoz, T., & Pérez-Amaral T., (2011). "Factores determinantes del comercio electrónico en España". Boletín económico de ICE, Información Comercial Española, vol. (3016), 51-65.

Garín-Muñoz T., López, R., Pérez-Amaral T., Herguera. I. & Valarezo, Á. (2019). "Models for individual adoption of eCommerce, eBanking and eGovernment in Spain", Telecommunications Policy Vol. 43, 100-111.

Hacker, K. L., & Mason, S. M. (2003). Ethical gaps in studies of the digital divide. Ethics and Information Technology, 5(2), 99–115. https://doi.org/10.1023/A:1024968602974

Hackl, F., Kummer, M.E., Winter-Ebmer, R., & Zulehner, C. (2014). "Market structure and market performance in E-commerce" European Economic Review, 68, 199-218.

Hilbert, M. (2011) "The end justifies the definition: The manifold outlooks on the digital divide and their practical usefulness for policy-making", Telecommunications Policy Vol. 35, 715-736.

INE (2018). "Survey on Equipment and Use of Information and Communication Technologies in Households". https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176741&menu=ultiDatos&idp=1254735976608

International Telecommunications Union, ITU (2016). ICT Facts and Figures.

Lera-López, F., Billon, M., & Gil, M. (2011). "Determinants of internet use in Spain". Economics of Innovation and New Technology, 20(2), 127–152.

Limayem, M., Khalifa, M., & Frini, A. (2000) "What makes Consumers Buy from Internet? A longitudinal Study of Online Shopping", IEEE Transactions on Systems, Man and Cybernetics 30, No. 4.

Littlechild, S. C. (1975) "Two-part tariffs and consumption externalities", Bell Journal of Economics and Management Science. Vol. 6, No 2, 661-670.

National Telecommunications and Information Administration, NTIA, (1999) Falling Through the Net: Defining the Digital Divide. Washington DC. https://www.ntia.doc.gov/report/1999/falling-through-net-defining-digital-divide

OECD (2001). "Understanding the Digital Divide", OECD Digital Economy Papers, No. 49, OECD Publishing, Paris. http://dx.doi.org/10.1787/236405667766.

Oliveira, T., Alhinho, M., Rita, P., & Dhillon, G. (2017). Modelling and testing consumer trust dimensions in e-commerce. Computers in Human Behavior, 71, 153–164. https://doi.org/10.1016/j.chb.2017.01.050

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, 213-226.

Von Rabenau, B., & Stahl, K. (1974). "Dynamic Aspects of Public Goods: A further Analysis of the Telephone System", Bell Journal of Economics and Management Science, Vol 5. No. 2, autumn, 651-669.

Răileanu, M. (2018) "New insights from a multilevel approach to the regional digital divide in the European Union", Telecommunications Policy 42, 452-463.

Robles, J.M., & Torres, C. (2012). "Digital Divide and the Information and Communication Society in Spain". Sociologija i Prostor, 194 (3), 291-307.

Rohlfs, J. (1974). "A Theory of Interdependent Demand for a Communications Service". Bell Journal of Economics and Management Science, vol. 5, 1, 16-37.

Ruiz-Rodríguez, F., Lucendo-Monedero, A. L., & González-Relaño, R. (2018). Measurement and characterization of the Digital Divide of Spanish regions at enterprise level. A comparative analysis with the European context. Telecommunications Policy, 42(3), 187–211. https://doi.org/10.1016/j.telpol.2017.11.007

Salemink, K., Strijker, D., & Bosworth, G. (2017). Rural development in the digital age: A systematic literature review on unequal ICT availability, adoption, and use in rural areas. Journal of Rural Studies, 54, 360–371. https://doi.org/10.1016/j.jrurstud.2015.09.001

Sharma, S. K. (2005). "Socio-Economic Impacts and Influences of E-Commerce in a Digital Economy" in Digital Economy: Impacts, Influences and Challenges, IGI Global, 1-20.

Squire, L. (1973). “Some aspects of optimal pricing for Telecommunications”, Bell Journal of Economics and Management Science, Vol. 4, No 2, autumn, 515-525.

Szeles, M. R. (2018). New insights from a multilevel approach to the regional digital divide in the European Union. Telecommunications Policy, 42(6), 452–463. https://doi.org/10.1016/j.telpol.2018.03.007

Taylor, L. D. (1994). Telecommunications Demand in Theory and Practice. Kluwer Academic Publishers.

Valarezo, Á., Pérez-Amaral, T., Garín-Muñoz, T., Herguera, I., & López, R. (2018). “Drivers and Barriers to Cross-Border E-Commerce: Evidence from Spanish Individual Behavior”, Telecommunications Policy, Vol. 42, Issue 6, 464-473.

Valarezo, Á., López, R., Pérez-Amaral, T., Garín-Muñoz, T., & Herguera, I. (2019). Adoption of e-commerce by individuals: Evidence from pooled data in Spain. Forthcoming, Book in Honor of Gary Madden.

Varian, H.R. (2002). “The Demand for Bandwidth: Evidence from the INDEX Experiment”. In Robert, C., & James, A. (Eds.). Broadband: Should We Regulate High-Speed Internet Access? Washington, D.C.: AEI-Brookings Joint Centre for Regulatory Studies.

Vehovar, V., Sicherl, P., Hüsing, T., & Dolnicar, V. (2006). Methodological Challenges of Digital Divide Measurements. The Information Society, 22(5), 279–290. https://doi.org/10.1080/01972240600904076

Vicente, M.R., & López, A.J. (2008). “Some empirical evidence on Internet diffusion in the New Members States and Candidate Countries of the European Union”. Applied Economic Letters, 15, 1015-1018.

Wikipedia, 2019. Autonomous communities of Spain no names. https://es.wikipedia.org/wiki/Archivo:Autonomous_communities_of_Spain_no_names.svg

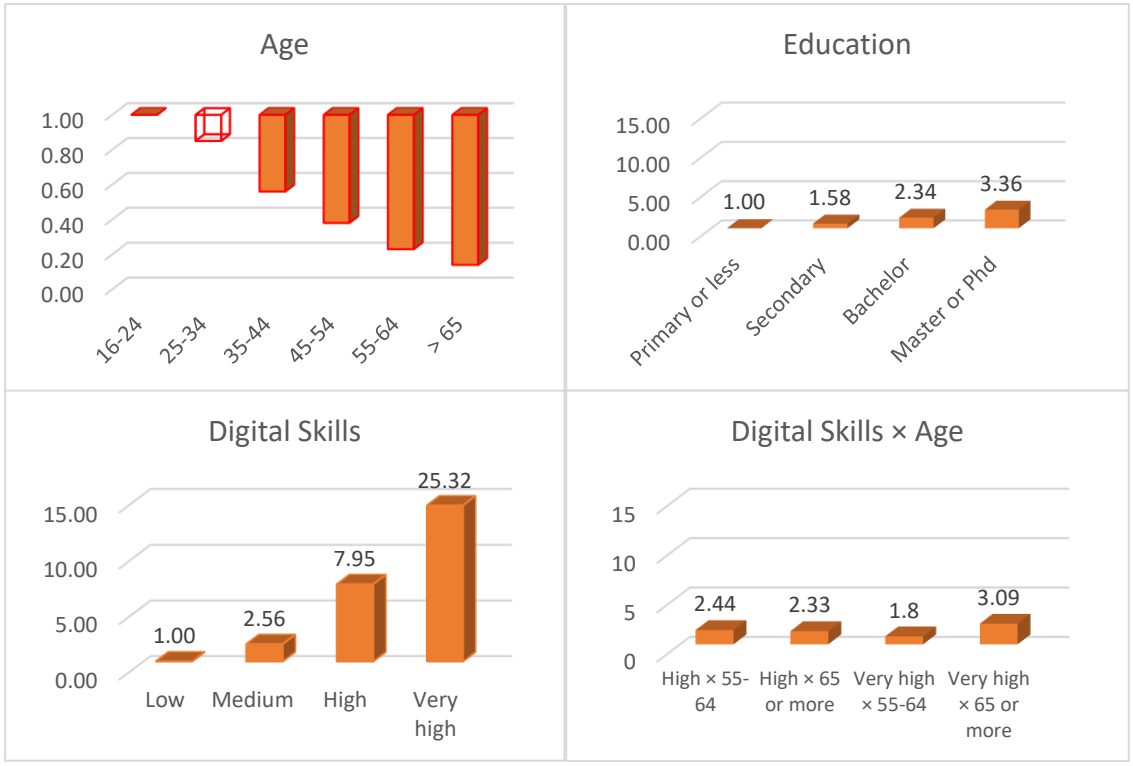
Zoroja, J. (2011). Internet, E-commerce and E-government: Measuring the Gap between European Developed and Post-Communist Countries. Interdisciplinary Description of Complex Systems, 9(2), 119–133.

Appendix

Figures 6-9 represent selected estimated coefficients of the equations in section 4.

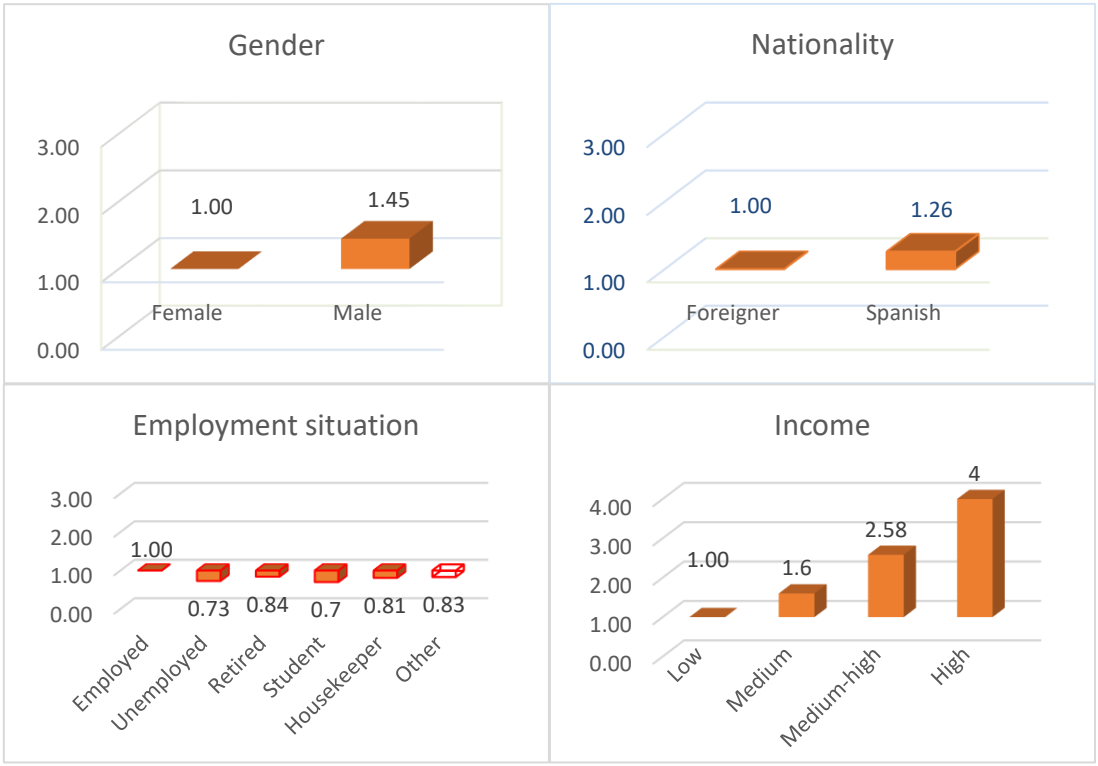
Figure 6

Odds of e-commerce adoption. Panel (2008-2016)



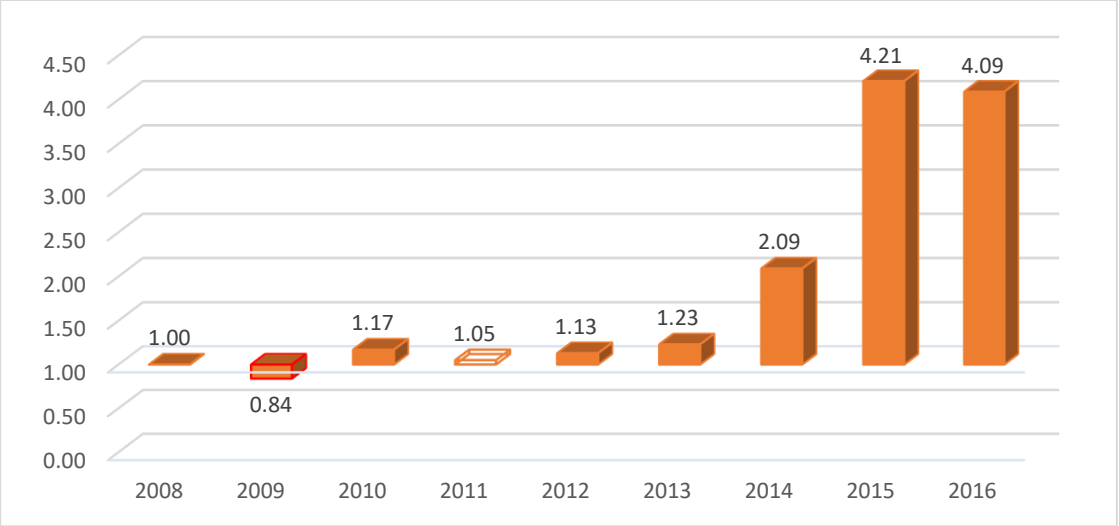
Source: Own elaboration based on estimation results. *Table 2 (1)*.

Figure 7:
Odds of e-commerce adoption. Panel (2008-2016)



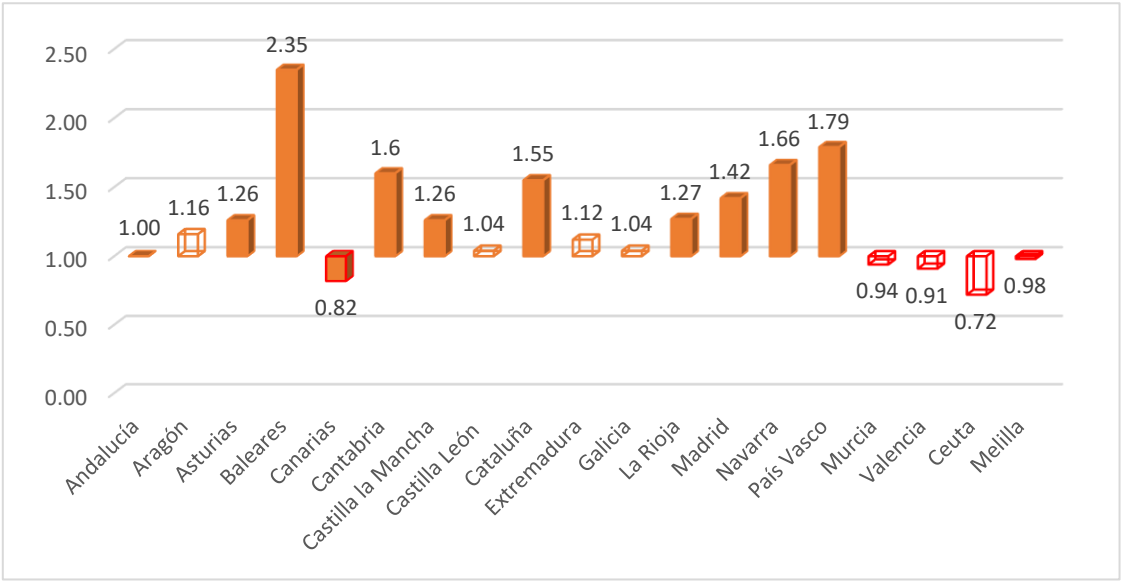
Source: Own elaboration based on estimation results. *Table 2 (1)*.

Figure 8:
Odds ratios of yearly dummies. Panel (2008-2016)



Source: Own elaboration based on estimation results. *Table 2 (1)*.

Figure 9:
Odds ratios of e-commerce adoption by regions. Panel (2008-2016)



Source: Own elaboration based on estimation results. *Table 2 (1)*.

Table A-1:
Heckman selection model of adoption of e-commerce by individual Internet users

	Variables		(3) Internet use. Selection equation. Probit		(4) E-commerce adoption. Probit	
			Coef.	z	Coef.	z
	Gender	Female				
		Male	0.05	1.08	**0.15**	**11.17**
Exclusion restriction	Broadband	No				
		Yes	**1.10**	**21.71**		
Sociodemographic	Age	16-24				
		25-34	0.01	0.08	-0.08	-1.00
		35-44	-0.01	-0.09	**-0.29**	**-3.77**
		45-54	-0.24	-1.69	**-0.43**	**-5.41**
		55-64	**-0.43**	**-2.91**	**-0.63**	**-7.47**
		65<	**-0.40**	**-2.48**	**-0.84**	**-8.67**
	Education	Primary or less				
		Secondary	-0.06	-1.15	**0.19**	**6.70**
		Bachelor	**0.03**	**0.38**	**0.33**	**10.37**
		Master´s/Phd	**0.01**	**0.13**	**0.47**	**14.97**
	Digital Skills	Low				
		Medium	**1.32**	**4.47**	**0.39**	**4.87**
		High	**5.53**	**40.47**	**0.89**	**11.52**
		Very high	**5.45**	**39.65**	**1.44**	**18.67**
	Habitat	<20,000				
		20,000-100,000	**0.12**	**1.98**	-0.03	-1.46
		100,000-500,000	0.03	0.27	**-0.12**	**-4.32**
		500,000<	0.00	-0.02	-0.01	-0.46
	Household Members	One				
		Two	-0.03	-0.51	**-0.07**	**-2.91**
		Three	-0.15	**-2.06**	**-0.15**	**-6.45**
		Four	-0.13	-1.65	**-0.16**	**-6.71**
		Five or more	-0.11	-1.05	**-0.22**	**-6.98**
	Nationality	Foreigner				
		Spanish	-0.10	-1.20	**0.08**	**2.88**
Economic	Employment Situation	Employed				
		Unemployed	**-0.15**	**-2.42**	**-0.14**	**-7.02**
		Retired	**-0.43**	**-5.22**	-0.06	-1.57
		Student	**1.15**	**2.81**	**-0.15**	**-4.64**
		Housekeeper	**-0.24**	**-3.10**	-0.04	-1.34
		Other	**-0.35**	**-3.07**	-0.06	-1.35
	Income	Low				
		Medium	0.08	1.52	**0.21**	**10.41**
		Medium-high	0.12	1.65	**0.43**	**19.67**
		High	0.10	0.99	**0.62**	**23.89**

Table A1 (Continued)

Variables			(3) Selection equation: Internet use. Probit		(4) E-commerce adoption. Probit	
			Coef.	z	Coef.	z
Interaction	Digital Skills × Age	Medium × 25-34	-0.19	-0.59	0.15	1.64
		Medium × 35-44	-0.16	-0.51	**0.34**	**3.85**
		Medium × 45-54	0.28	0.80	**0.33**	**3.65**
		Medium × 55-64	-0.10	-0.31	**0.34**	**3.63**
		Medium × 65 <	-0.13	-0.40	**0.37**	**3.49**
		High × 25-34	0.06	0.39	**0.24**	**2.70**
		High × 35-44	0.06	0.42	**0.37**	**4.36**
		High × 45-54	0.21	1.39	**0.37**	**4.24**
		High × 55-64	**0.44**	**2.80**	**0.41**	**4.40**
		High × 65 <	**0.46**	**2.70**	**0.40**	**3.65**
		Very high × 25-34	0.09	0.56	**0.27**	**3.02**
		Very high × 35-44	0.01	0.07	**0.43**	**4.97**
		Very high × 45-54	0.19	1.23	**0.42**	**4.72**
		Very high × 55-64	**0.45**	**2.54**	**0.33**	**3.34**
		Very high × 65 <	**0.62**	**2.93**	**0.54**	**4.50**
Time effects	Year	2008				
		2009	-5.90	-1.86	**-0.08**	**-3.67**
		2010	0.08	0.03	**0.07**	**2.79**
		2011	-5.97	-1.86	0.01	0.44
		2012	**-6.07**	**-2.01**	0.03	1.13
		2013	-0.10	-0.03	**0.07**	**2.92**
		2014	-5.87	-1.89	**0.29**	**11.53**
		2015	-0.08	-0.03	**0.60**	**23.96**
		2016	-0.10	-0.03	**0.58**	**23.24**
Sociogeographic effects	Autonomous Communities	Andalucía				
		Aragón	0.15	1.38	0.06	1.56
		Asturias	0.05	0.46	**0.11**	**3.16**
		Baleares	0.27	1.90	**0.35**	**8.78**
		Canarias	0.09	0.74	**-0.09**	**-2.33**
		Cantabria	0.22	1.70	**0.20**	**5.14**
		Castilla la Mancha	0.09	0.81	**0.09**	**2.37**
		Castilla León	**0.36**	**3.08**	0.02	0.52
		Cataluña	0.01	0.10	**0.18**	**6.35**
		Extremadura	0.22	1.92	0.05	1.24
		Galicia	0.17	1.64	0.01	0.41
		La Rioja	**0.27**	**2.02**	**0.10**	**2.42**
		Madrid	0.16	1.56	**0.13**	**4.53**
		Navarra	0.17	1.70	**0.22**	**6.72**
		País Vasco	0.11	0.91	**0.26**	**7.34**
		Murcia	0.01	0.12	-0.02	-0.56
		Valencia	**0.20**	**2.12**	-0.05	-1.51
	Autonomous Cities	Ceuta	0.47	1.39	-0.15	-1.77
		Melilla	0.05	0.15	-0.04	0.15
Constant			6.88	2.23	1.71	-20.26
N. observations			66,099		66,999	
Wald χ^2			4592.85, DF: 70		13666.89, DF: 70	
Pseudo R^2			0.5116			
Wald χ^2: H_0 = independent equations			1.12 DF:1 p-value = 0.2900			

Notes: Coefficients and z statistics significant at the 5% are represented in bold. Probit model, selection equation (3). Probit model, e-commerce adoption equation (4). Heckman probit clustered standard errors. Estimations equation (3) refer to all Internet users that have purchased (or not) online in the last year. The first category of each variable is the reference and is omitted. The samples reflect the fact that some variables are missing for some observations, which reduces the usable sample size. All estimations have been calculated using Stata 15.

Statistical significance and signs of coefficients of all explanatory variables coincide between the probit estimation results (second stage of Heckman selection model), *Table A1*, and estimations of the two logistic specifications, *Table 2*. On the other hand, a formal test of independence of two equations of the selection model (equations (3) and (4), with $\chi^2 = 1.12$, and p-value = 0.29) does not provide sufficient evidence to reject the hypothesis of independent equations, implying that, for the considered sample, the initial decision of using the Internet is unrelated to the decision of adopting e-commerce for private use. This may suggest that the modelling approach of using just one step delivers a useful model for e-commerce.

“The farther back you can look, the farther forward you are likely to see”

— Winston Churchill

“Following the law-and-order crowd’s lead and reducing the ocean of facts to simple numerical scores has drawbacks but it also has the one great merit of forcing everyone to confront the same evidence —with surprising results”

— Ian Morris