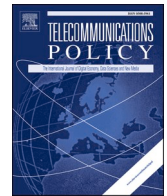




ELSEVIER

Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

## Telecommunications Policy

journal homepage: [www.elsevier.com/locate/telpol](http://www.elsevier.com/locate/telpol)

# Unleashing the potential of online learning in Spain: An econometric analysis

Rafael López, Ángel Valarezo<sup>\*</sup>, Teodosio Pérez-Amaral

*Instituto Complutense de Análisis Económico, Universidad Complutense de Madrid, Spain*

## ARTICLE INFO

### JEL classification:

C30  
L86  
L96

### Keywords:

E-learning adoption  
Online education  
Digital skills  
Linear probability model  
Logistic model  
Heckman model

## ABSTRACT

Online learning and training continue gaining momentum worldwide resulting in the reduction of the traditional form of face-to-face education with its temporal and spatial limitations. Online education improves access to education and training, as witnessed during the Covid-19 pandemic. This article focuses on online education adoption in Spain. A representative survey on ICT use in households conducted annually by the Spanish National Institute of Statistics is used to construct a panel database for the years 2008–2020. The first objective is to provide an econometric model for adopting online education using this panel data. Next is to measure the effects of relevant observable individual socioeconomic variables on adoption. A Heckman selection model allows for estimating the impact of gender, age, education, digital skills, habitat, and income. The article also measures the effects of Covid-19 in 2020 on different population groups. The drivers and impediments have the expected signs and plausible sizes. The paper concludes with policy recommendations and suggestions for further research.

## 1. Introduction

Online learning and training are gaining momentum worldwide, reducing the temporal and spatial limitations associated with the traditional form of face-to-face education. What's more, online education improves access to education and training, increases educational quality, reduces the cost of learning, and enhances education's cost-effectiveness (Bates, 1997).

Digital divide issues are not alien to online education adoption, as there are socioeconomic-related disparities in the use and adoption of e-learning among different groups of individuals (Reddy et al., 2021). Furthermore, as education is an essential element in promoting social mobility, even more critical in a knowledge-driven economy (Brown et al., 2003), gaps in the adoption of education provided by digital means may exacerbate current inequalities.

Online education is provided in asynchronous and synchronous environments along with its various blends, including partly online and partly face-to-face activities. Asynchronous online learning has the advantages of accessing material anytime and anywhere, reaching a wider audience, and achieving content uniformity. Synchronous learning is closer to the classroom experience while providing recordings for later use. Online learning allows teaching staff to team their expertise and distribute themselves efficiently geographically. Instructors can also successfully combine online activities with face-to-face teaching.

In order to make the terminology precise, we followed Money & Dean (2019) by considering online education, online delivery, online education design, and computer-based instruction as equivalent. According to Volery & Lord (2000), technologies that are

<sup>\*</sup> Corresponding author.

E-mail address: [angel.valarezo.unda@ucm.es](mailto:angel.valarezo.unda@ucm.es) (Á. Valarezo).

<https://doi.org/10.1016/j.telpol.2023.102544>

Received 29 March 2023; Accepted 30 March 2023

Available online 19 April 2023

0308-5961/© 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

computer-based, internet-based, or of other digital formats contain terms and phrases that are essentially interchangeable. This also includes a similarity in meaning for the words: online teaching and learning, e-learning, and distance learning. [Simonson et al. \(2011\)](#) define distance learning as an institution-based, formal education in which learners are physically separated, but connected with instructors, resources, and other learners by telecommunications systems. [Larreamendy-Joerns & Leinhardt \(2006\)](#) concluded that distance and online education have merged ([Walker & Kelly, 2007](#)). As Goodyear et al. suggest, online teaching and learning refer to the teaching and learning that occurs over a computer network in which interactions among people – including synchronous and asynchronous forms of interaction – remain an essential part of the learning process.

Despite several advantages of online learning, retaining students is a crucial challenge due to its high attrition rate ([Perna et al., 2014](#)). Online education also presents disadvantages relative to traditional face-to-face learning. The feedback between instructors and students is mainly lost, as is the relationship among students. The possibility of chatting before, after class, and during breaks is also limited. In synchronous learning, the student's physical presence in front of the computer is not assured. At times, instructors do not know to whom they are speaking.

Most existing literature uses technology adoption models to explain and understand decisions to adopt online educational technologies ([Panigrahi et al., 2018](#)).

The first objective of the current article is to provide an economic framework for adopting online education. A second objective is to model adoption using a large data set. Finally, a third objective is to measure the effects of relevant observable variables for adopting or not adopting online education. This framework may apply to online students and instructors. In addition, teaching institutions and national authorities may also benefit from the research output.

The rest of the article is as follows. Section 2 contains a literature review. Section 3 shows the theoretical framework. Section 4 includes the presentation and description of the data. The empirical model and estimation results are in section 5. Finally, section 6 compares 2020 and previous years, while section 7 includes the conclusions.

## 2. Literature review

Literature about online education and the link between the adoption of internet services and the digital divide is abundant, with various subtopics, emphasis, and methodologies. For example, [Panigrahi et al. \(2018\)](#) analyze the factors that influence online education's three critical elements: Adoption of technology, the continuation of technology use, and e-learning outcomes. Others, like [Andrews & Haythornthwaite \(2007\)](#), explore the specifics of the digital divide. Here we concentrate on two aspects: the specific adoption of online education and the related digital divide issue.

### 2.1. Adoption

The adoption of technology, online consumer behavior, and the specific decision of online learning adoption have been analyzed using several frameworks. The most used approaches are applications, adaptations, and unifications of individual acceptance and intention models. Among the soundest 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 \(1989\)](#); [Davis & Warshaw \(1989\)](#); 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 uses IDT, TRA, TAM, and TPB, including variables that affect an individual's motivation to accept new online learning, and explains the decision-making process. When the goal is to reach further than the initial acceptance, ECT is widely used. The contributions of this literature are precedents of the current study.

#### 2.1.1. Online learning adoption

[Panigrahi et al. \(2018\)](#) examine a bibliographical database, organizing the literature into three categories: online learning adoption, the continuation of online learning use, and learning outcomes. The antecedents and main extensions of the three categories are discussed extensively. In addition, the discussion focuses on personal and environmental factors for online education adoption, which is directly related to our research. A later survey by [Money & Dean \(2019\)](#) also adopts this strategy.

#### 2.1.2. Adoption, personal factors

Perceived usefulness (PU) and perceived ease of use (PEoU) are the main factors that, according to the Technology Acceptance Model ([Davis, 1989](#)), are the predominant antecedents of technology adoption. Other factors that affect technology acceptance are the perception of interaction, cognitive absorption, self-distraction, cognitive age, social network, national culture, and surrounding conditions (e.g. thermal climate and national wealth). Additional personal factors include perceived behaviorcontrol, performance and effort, expectancies, and user resistance.

#### 2.1.3. Adoption, environmental factors

The environmental factors for online learning adoption include perceived characteristics of innovation, subjective norms, facilitating conditions, technology inhibitors, and technology adoption in organizations. Some of these factors also influence ICT adoption by individual users ([Khalil Moghaddam & Khatoun-Abadi, 2013](#)).

Focusing on French universities, [Jacqmin \(2019\)](#) analyses the effect of massive open online courses (MOOCs) on new students'

enrollment in online and traditional programs and media coverage. The study found that all else being equal, offering MOOCs raised over 2% of student enrollments for universities hosting this kind of program. Furthermore, a double positive effect is identified due to its enhanced information and attention-grabbing quality.

Bryson & Andres (2020) analyze the effects of Covid-19 on adopting e-learning in a university. The pandemic forced rapid improvisation and the adoption of online teaching. The paper distinguishes between the development of distance learning programs and the rapid adoption of online learning. The article concludes that a shift to online education is not about substituting on-campus with online, but developing a new and transformational approach that will extend universities' reach and alter their fundamental essence. Part of this shift will reflect the emergence of new bimodal approaches to facilitating learning outcomes by accommodating all types of students.

The European Commission JRC's (Joint Research Centre) report pointed out the decrease in adult learning due to the economic impact of Covid-19 caused mainly by the disruption of on-the-job training due to the pandemic. However, using Eurostat's Community Survey on ICT usage in households, the study shows that Covid-19 is associated with more adults taking online courses in most EU countries. Furthermore, individuals aged 55 to 64 and less-educated adults have increased adoption.

## 2.2. Digital divide

Online education is considered one of the mechanisms to enable broader participation in education and lifelong learning among those who may be excluded or face a high risk of exclusion from the benefits of higher education attainment levels. Several studies (Andrews & Haythornthwaite, 2007; Hvorecký, 2004; Journell, 2007; Sims et al., 2008, among others) have established and analyzed the link between the adoption of internet services like online education and the digital divide.

In a report by OECD (2001, p. 5) digital divide was described as the "gap between individuals, households, business, and geographic areas at different socioeconomic levels with regard to both their opportunities to access information and communication technologies and their use of internet for a wide variety of activities."

The study of the digital divide has been traditionally related to the "haves" and "have nots," but the evolution of technology and its growing adoption by individuals called for a broader perspective (NTIA, 1995). Besides analyzing the determinants of access to equipment and an internet connection, there is an extensive body of literature that expands the research to factors like digital skills, technological evolution, internet usage, and the characterization of the individuals (DiMaggio & Hargittai, 2001; Hargittai, 2001; Norris, 2001; OECD, 2001).

Several authors have identified and analyzed three levels of the digital divide (Kraemer et al., 2005; Riggins & Dewan, 2005; Wei et al., 2010, among others). The first level of the digital divide points to the study of inequalities in access, while the second level introduces digital skills as a critical factor in performing internet-related tasks (Hargittai, 2001; van Deursen et al., 2014). As a third level, Helsper & van Deursen (2015) identified the differences in the achievement of tangible outcomes of internet use as the "third level of the digital divide." These three categories were addressed several times throughout the literature. Below we give an account of some of the most relevant for our analysis, where socioeconomic and demographic characteristics have been introduced in empirical approaches to model the individual adoption of ICTs and to measure digital divides.

Lera-Lopez et al. (2011) developed explanatory models to study the determinants of internet use in Spain, exploring both use and intensity of use. Controlling for different socioeconomic and demographic variables, the authors identified differences in the determinants of use and its extent of use. However, the lack of data about relevant variables is pointed out as the main limitation of the paper.

A study on how educational inequalities affect internet use in developing countries (Billon et al., 2021) analyses countries' educational system structures' relative influence on internet use. The main finding shows that within-country educational disparities negatively influence internet use.

Garín-Muñoz et al. (2019) identified the presence of a digital divide in the use of internet services in Spain and modelled the adoption of e-banking, e-government, and e-commerce. Age, education, income, and digital skills are the main factors the study found relevant to explain an individual's adoption patterns.

Using panel data for the 2007–2019 period, Pérez-Amaral et al. (2021) define and analyze the gaps in the use of internet services by households in Spain. Twelve different services are considered. The identified gaps show the evolution of digitization and, in some cases, the digital exclusion for specific groups.

The evolution of internet gender gaps in Spain and the effects of the COVID-19 pandemic is discussed by Garín-Muñoz et al. (2022). The authors conclude that the internet gap is a phenomenon with a specific gender component, such that the gender variable is always significant in explaining the likelihood of being a user of certain digital services.

## 2.3. Caveats

While previous literature has been extensive and valuable in many ways, it may be helpful to define additional concepts with precision. There is a need to clarify whether the arguments exposed as antecedents of e-learning adoption purposes were considered as private or work-related from an individual consumer's perspective. The distinction between the motivations of a personal user and an online education supplier would require further clarification.

The factors for the adoption of online technology considered in the literature are mostly psychological. However, adoption includes a bundle of technology and specific content, which may be relevant to embracing online education. This paper focuses on the economic factors pertinent to choosing courses that happen to be taught online. Therefore, there is an essential difference between adopting e-

learning technology versus engaging in specific online courses.

Part of the reviewed research shows little emphasis on the heterogeneity of the students: for example, their different fields of knowledge, ages, interests, gender, educational background, income levels, geographical locations, and socioeconomic characteristics all seem to be neglected. This characteristic may suggest adopting an economic framework. The adoption decision is a function of the costs and benefits of online education engagement based on observable individual and environmental variables. The following section contains an alternative model that accounts for as much heterogeneity as the data can support.

### 3. Theory

The present study follows an economic perspective using the neoclassical utility maximization approach (Varian, 2002). The demand for access to the internet is determined by the consumer surplus associated with usage and the cost of access. Regarding access to online learning, the relevant theory is that of the telecommunications demand framework of Artle & Averous (1973), Squire (1973), von Rabenau & Stahl (1974), Rohlfis (1974), Taylor (1994), Kridel et al. (1999), and Rappoport et al. (2003).

In telecommunications, using a specific service is conditional on access to this service (Taylor 1994). The current approach assumes that internet access is a prerequisite for adopting online education. A consumer could obtain internet access through various technical solutions: dial-up, cable, ADSL, broadband, narrowband, or mobile phones, tablets, and portable computers.

Often, access to the internet is not an entirely individual decision, but a circumstance governed by the carriers that incorporate internet access without knowledge of the consumer. Demoussis & Giannakopoulos (2006) use a similar argument for the European case using 2002–2003 data.

In this context, an individual user of the internet 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 & Giannakopoulos (2006), Fairlie (2004); Vicente & López (2008); Lera-Lopez et al. (2011), and Valarezo et al. (2018, 2020), the last four refer to the case of Spain.

From a standard neoclassical utility optimization approach, the maximization of the utility (U) of an individual obtained from online education (Y<sub>i</sub>), will be a function of the benefits B(Y<sub>i</sub>|x) of doing so and the costs C(Y<sub>i</sub>|x), where x is a set of conditioning variables associated with it. The conditional probability of an individual adopting online education is, at a given point in time, t:

$$P(Y_i|x_{it}) = P[B(Y_{it}) - C(Y_{it}) > 0|x_{it}] \quad i = 1, \dots, N, \text{ individuals.} \tag{1}$$

$$t = 1, \dots, T, \text{ years.}$$

The individual consumers considered in equation (1) are internet users. The economic model incorporates rational choice under uncertainty and is compatible with a theory of the decision of investment in human capital. The following section presents the data set used in this paper to estimate models derived from equation (1).

### 4. The data

This work uses annual panel data on individuals from the Households survey on Equipment and Use of Information and

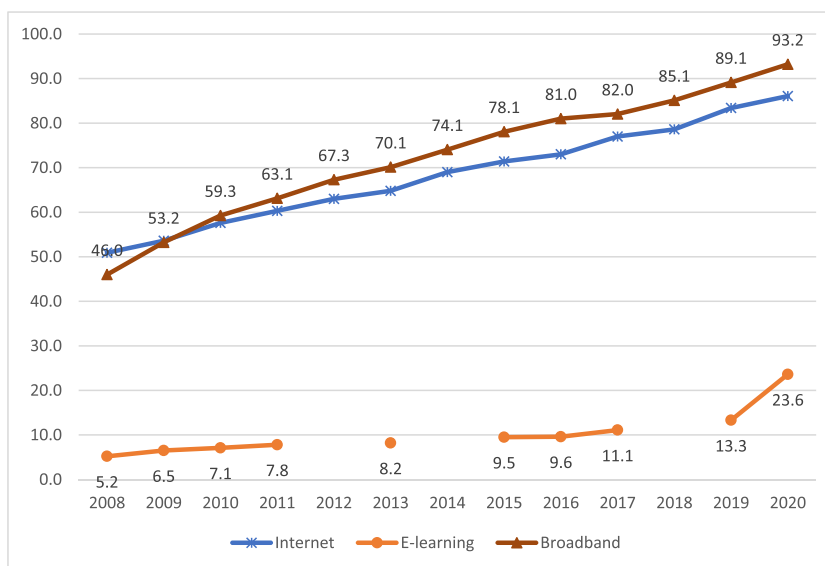


Fig. 1. Percentage access to the internet, broadband, and online education 2008–2020. Relative to the adult population 16 and above. Source: Self-elaboration based on INE, 2020.

Communications Technologies from 2008 to 2020. The survey followed Eurostat guidelines and was performed by the Spanish National Statistical Institute (INE, 2020). It is representative at regional and national levels and includes an elevation factor. The main focus is the adoption of ICT technology and services by households and individuals. Eurostat coordinates and makes the survey available in all European countries annually.

Fig. 1 shows the evolution of penetration rates for internet, broadband and online education in Spain for individuals aged 16 or more.

Fig. 1 shows that internet and broadband access increased steadily over the 13 years of the sample, reaching 83.4% and 89.1%, respectively. In addition, online learning grew faster, reaching a penetration of 13.3% of the population 16 and above in 2019 and 23.6 in 2020.

In the first 12 years of the sample, up to 2019, internet access increased by 63.8%, while online learning increased by 155.7%, which more than doubles the rise of internet access and shows the rapid increase of online education in Spain.

Fig. 2 shows a comparison with other European countries. Spain is in third place in penetration after Finland and Sweden in 2019 and second in 2020 after Finland. This position may be due to the existence of UNED, a large national distance education university, with 250,000 students enrolled in 2019. Other institutions at a distance include UOC, UNIR, Isabel de Castilla, and several smaller universities.

The primary data for 2008–2020 are publicly available at INE (2020) - together with additional information concerning the survey’s design and implementation. The data is obtained from a rotating survey that includes 18,000–21,000 dwellings annually. The same residence is interviewed for a maximum of four consecutive years with around a 30% replacement yearly. Approximately 60% of the interviews were conducted by phone and 40% in person.

However, we are interested in identifying individuals rather than dwellings (a dwelling that has participated in the survey for several years may have different respondents across years).

All the annual raw data were first subjected to a meticulous homogenization process of variables to obtain a pool. Then, the pool was thoroughly filtered to identify the same individuals across years (thus obtaining a panel). Descriptions of these processes are found in Pérez-Amaral et al. (2021).

Thus obtained Panel 2008–2019 consists of 188,683 observations, corresponding to 88,345 different individuals: 16,174 of which were interviewed four times (years); 15,315 three times; 21,186 two times, and 35,670 just once.

Table 1. Dimensions, variables, and categories related to e-learning. Source: Self-elaboration based on the ICT-H survey of INE (2020).

In the table, the first column contains the dimensions used in this study: sociodemographic, individual skills, and economic characteristics. The second column specifies the variables in this panel data set for approximating each dimension. Finally, the third column describes each variable’s different levels or categories.

Almost all of these variables are well-established in the literature. Their levels or categories were mainly obtained directly from the surveys. For example, for the variable Internet Trust, individuals are directly asked to indicate their level of trust by choosing among three levels: Low (little or nothing), Medium (relatively), or High (a lot).

Individuals chose among intervals for the variable income to indicate the approximate total household income level. However, these intervals are not constant throughout the years 2008–2020. The four groups we constructed are as follows:

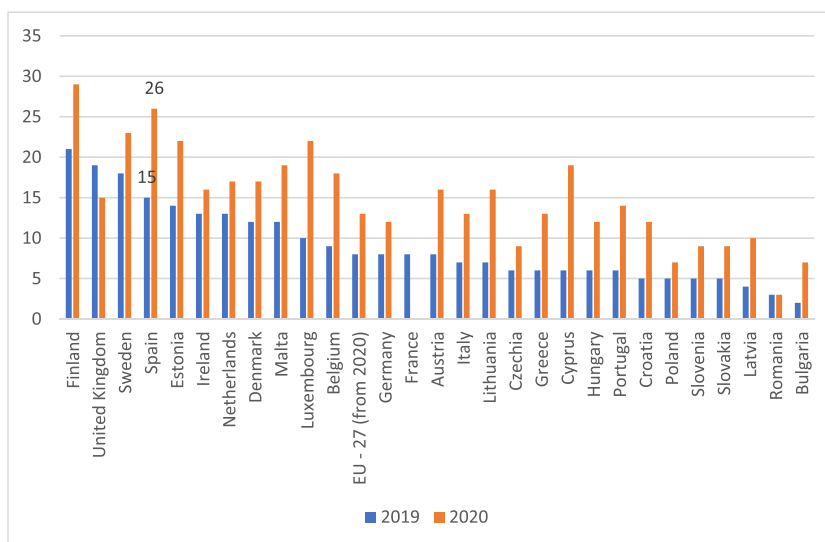


Fig. 2. Percentage of the population (people aged 16–74 years) taking an online course. European countries 2019–2020. 2020 data missing for France.

Source: Self-elaboration based on Eurostat, 2021.

**Table 1**

Contains a list and description of the main variables from the panel used in this article.

Dimensions	Characteristics/Variables	Categories or levels within each variable
Services	Internet access	2 groups: 1 if access; 0 otherwise.
	Internet user	2 groups: 1 if internet user; 0 otherwise.
	(in the last 3 months)	2 groups: 1 if e-learning user; 0 otherwise
Sociodemographic	E-learning	
	Gender	2 groups: 1 if male, 0 if female
	Age	7 groups: 16–25, 25–35, 35–45, 45–55, 55–65, 65–75, 75 or more.
	Household members	5 groups: One, Two, Three, Four, Five, or more.
	Population size (Habitat)	
Individual	Nationality	5 groups: less than 10,000; 10,000–20,000; 20,000–50,000; 50,000–100,000; and 100,000 or more and province capitals.
	Education	2 groups: 1 if Foreigner; 0 if Spanish. 4 levels of study: None or Primary, Secondary, Bachelor's Degree, and Master's or Ph.D.
	Digital Skills	4 levels: Low, Medium, High, and Very High.
	Internet trust	3 groups: Low, Medium, and High.
	Economic	Income (household monthly net income)

2008–2012: Low (less than 1100€), Medium (between 1100€ and 1800€), High (between 1800€ and 2700€), Very High (more than 2700€).

2013–2020: Low (less than 900€), Medium (between 900€ and 1600€), High (between 1600€ and 3000€), Very High (more than 3000€).

The variable Digital Skills, which, according to the [European Commission \(2019, 2020\)](#), is crucial to measure and promote digital inclusion, was constructed as a synthetic index based on the former European Commission's Digital Skills Indicator ([European Commission, 2016](#)). It covers the total number and the complexity of several activities involving the adoption and use of digital equipment and internet services. A detailed description of how this variable was constructed is found in [Pérez-Amaral et al. \(2021\)](#).

#### 4.1. Data description

[Table 2](#) shows the penetration rates of online education according to the different variables from 2008 to 2020. The data are missing for the years 2012, 2014, and 2018. Internet Trust is not available before 2015.

The penetrations for 2019 and 2020 reveal no gender gap; the ages with the highest penetration are 25–34 years, and the number of household members is also available. The population size has an increasing relationship with online education, and Spanish nationals are more inclined to engage in online education than foreigners. Foreigners represented 5% of the population in 2021.

Online education has an increasing relationship with formal education and with Digital Skills. There is also a positive relationship between the adoption of online education with internet trust and income.

[Table 2](#) shows that the penetrations of different services vary across categories and time. The variables related to more considerable variations in penetration are age, education, digital skills, and income. This may suggest the importance of each explanatory variable on the variations of online education adoption.

[Fig. 3](#) shows the gaps in the penetrations of online education according to selected variables from 2008 to 2019. Each gap is computed as the difference between the penetration of the category with the highest penetration and the penetration of a lower category divided by the penetration of the highest category ([Pérez-Amaral et al., 2021](#)). The gap measures the relative distance in penetrations between two categories. The gaps for 2020 are not shown until more reliable data, that includes the whole year, are available. See the Appendix, [Table A2](#), for data availability in 2020.

[Fig. 3](#) shows that the gender gap is almost zero in 2019, while the gaps by age are significant and do not decrease over time.

The gaps by population size are present with a 30% for sizes of less than 10,000 inhabitants compared to populations of more than 100,000 and provincial capitals.

The gaps in online learning by education are significant and do not decrease over time. In particular, the gap for none or primary studies is close to 100% compared to master's or Ph.D. degrees, and the gap for the group with secondary studies is 68%.

The gaps related to digital skills are also large and stable. They go from nearly 100% for low digital skills to around 60% for high digital skills (compared to very high digital skills).

The gaps by income are also large and stable, ranging from 80% for low to 30% for high (compared with very high).

According to their main explanatory variables, many gaps are large and do not show a clear decreasing pattern. This persistence leads us to the need to model these relationships in a synthetic and interpretable manner.

[Table 2](#) and [Fig. 3](#) indicate which variables may be relevant for adopting online learning in Spain. Still, the relationships are only indicative.

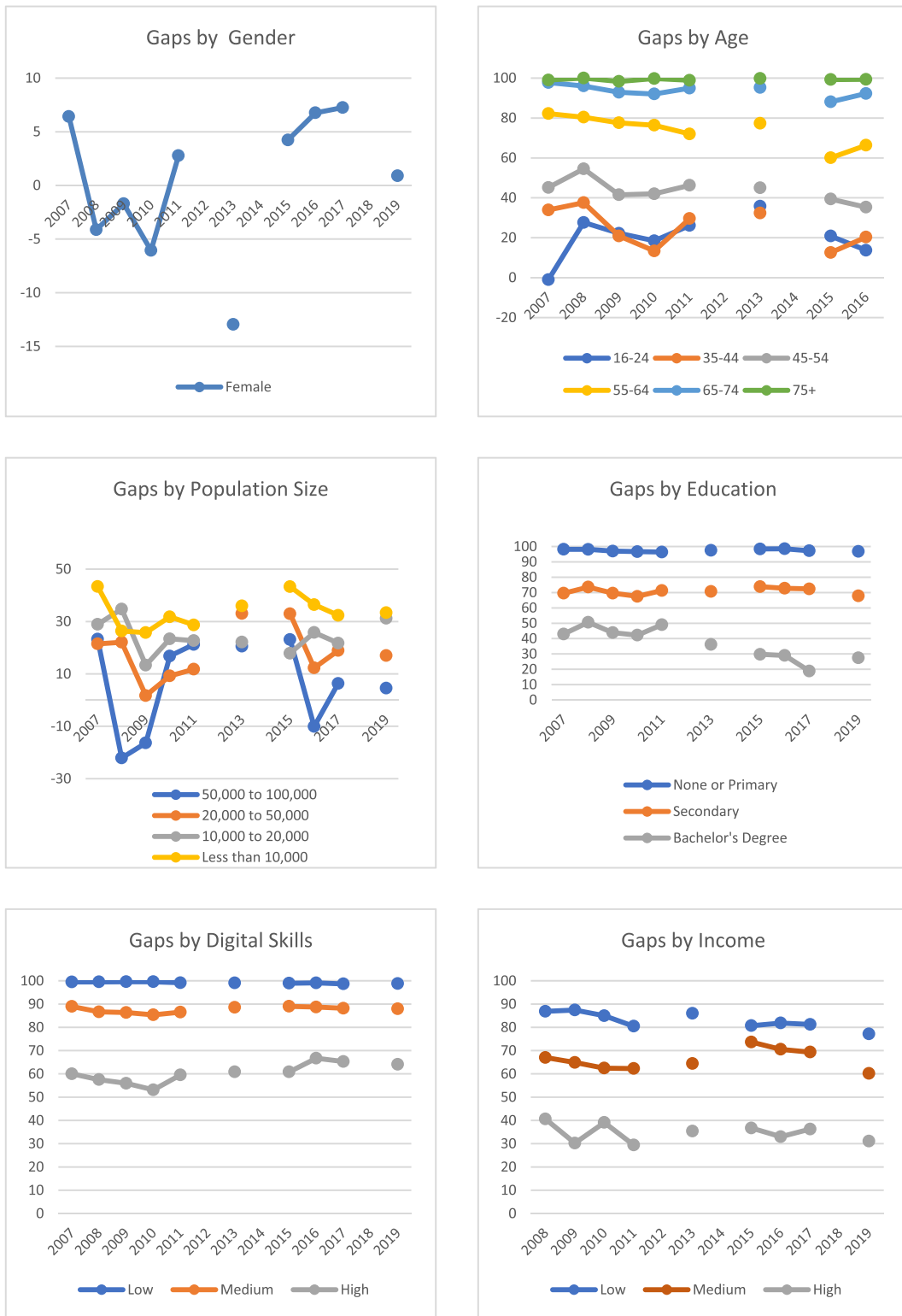
A high correlation between explanatory variables may be problematic in terms of the efficiency of the estimators and hypothesis testing. [Table 3](#) contains the polychoric ([Drasgow, 1988](#); [Kolenikov, 2016](#)) correlation matrix between the variables. Polychoric correlations measure the pairwise correlations between a set of variables allowing for discrete dummy variables. The general

**Table 2**

Online education penetration rates by variables 2008–2020, Blanks mean not available.

Source: Self-elaboration based on the TICH survey of [INE \(2020\)](#).

Variables	Categories	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Gender	Female	5.3	6.5	7.4	7.7		8.7		9.3	9.2	10.7		13.2	23.5
	Male	5.1	6.4	6.9	7.9		7.7		9.7	9.9	11.6		13.3	23.7
Age	16–24	7.3	8.7	9.8	10.6		10.4		12.6	14.4	16.6		16.7	49.1
	25–34	10.1	11.2	12	14.4		16.2		16.0	16.7	19.7		23.7	40.4
	35–44	6.3	8.9	10.4	10.1		11.0		14.0	13.3	15.7		19.1	30
	45–54	4.6	6.5	7.0	7.7		8.9		9.7	10.8	13.2		16.2	24.6
	55–64	2.0	2.5	2.8	4.0		3.7		6.4	5.6	6.1		8.4	13.5
	65–74	0.4	0.8	1	0.7		0.8		1.9	1.3	1.5		2.3	4.6
	75+	0	0.2	0	0.2		0		0.1	0.1	0.1		0.4	0.8
Household Members	One	2.7	5.0	5.0	5.7		6.0		6.3	7.0	7.8		11.3	16.2
	Two	4.0	5.3	5.1	5.8		5.8		7.4	7.6	8.2		8.8	14.5
	Three	6.0	7.4	8.1	9.0		10.3		10.5	9.9	13.0		14.9	26.5
	Four	6.8	7.3	9.3	9.7		9.3		11.9	12.5	13.7		18.7	33.6
	Five or more	4.0	6.1	5.6	7.3		8.7		9.7	9.3	11.8		12.6	30.7
Population Size	100.000+, Capitals	5.8	6.9	8.2	9		10.0		11.7	10.7	12.7		15.3	27.7
	50.000 to 100.000	7.1	8.0	6.8	7.1		8.0		9.0	11.8	11.9		14.6	22.9
	20.000 to 50.000	4.6	6.8	7.4	7.9		6.7		7.8	9.4	10.3		12.7	22.1
	10.000 to 20.000	3.8	6.0	6.3	7.0		7.8		9.6	8.0	10.0		10.5	19.3
	Less than 10.000	4.3	5.1	5.6	6.4		6.4		6.6	6.8	8.6		10.2	18.8
Nationality	Spanish	5.2	6.8	7.2	8		8.5		9.7	9.8	11.4		13.9	24.1
	Foreigner	4.8	4.6	6.6	6.2.0		6.3		7.3	7.3	8.9		8.0	19.4
Education	None or Primary	0.3	0.5	0.6	0.8		0.5		0.4	0.4	0.8		1.0	1.6
	Secondary	4.2	5.5	6.1	6.3		6.2		6.8	7.0	7.6		9.9	20.8
	Bachelor's Degree	7.8	10.1	10.9	11.2		13.5		18.2	18.1	22.4		22.5	36.3
	Master's or PhD	15.7	18	18.9	22		21.2		26	25.5	27.7		31.0	48.3
Digital Skills	Low	0.3	0.2	0.2	0.4		0.4		0.5	0.4	0.6		0.6	0.7
	Medium	7.4	5.8	7.1	5.7		5.1		4.9	5.1	5.7		6.0	8.9
	High	23.4	18.8	22.6	17		17.4		17.3	15.2	16.7		17.9	30.3
Internet Trust	Very High	55	42.7	48.2	42.1		44.5		44.3	45.6	48.2		49.9	69.5
	Low								8.9	10.5	9.0		9.4	22
	Medium								14.3	13.9	16.5		17.8	31.3
	High								20.6	16.2	20.8		28.3	30.4
Income	Low	1.6	1.8	2.6	3.4		2.8		4.6	3.9	4.7		6.1	12.3
	Medium	4.1	5.1	6.4	6.6		7.0		6.2	6.3	7.7		10.6	16.8
	High	7.4	10.2	10.4	12.4		12.7		15	14.3	16.0		18.3	29.8
	Very High	12.5	14.6	17	17.5		19.6		23.7	21.4	25.1		26.6	46.5



**Fig. 3.** Gaps in adopting online education by socioeconomic variablesSource: Authors' elaboration from INE (2020). Notes: Data were not available for online education (2012, 2014, and 2018). Reference categories: gender: male; age: 25–34; population size, 100,000+ and provincial capitals; education: Master's or Ph.D.; digital skills: very high digital skills; Income: very high income. The gaps are computed as the difference in penetration between the reference category and each of the other categories divided by the penetration of the reference category.

conclusion is that there are low correlations between the variables considered (except for the column on online learning, as expected). Therefore, values in the range  $\pm .29$  are considered low, while those above 0.7 in absolute value are considered high.

The following section contains a multivariate model incorporating the possible explanatory variables discussed. We intend to disentangle the effect that the explanatory variables may have and look for the partial effect of each variable.

## 5. The model and estimation

This section presents an estimable model of the decision to adopt online education. The model includes 2008–2019 and leaves out 2020 for heterogeneity reasons and comparison purposes.

Fig. 4 illustrates the decisions individuals make when confronted with the choices of using the internet and e-learning.

The decision to adopt online education for personal reasons is represented by the binary dependent variable defined below:

$$\begin{aligned} \text{Online learning} &= 1, \text{ if the individual has been engaged in online learning in the last 3 months.} \\ &= 0, \text{ otherwise.} \end{aligned}$$

Explanatory variables are grouped as follows: Sociodemographic, Individual Skills, and Economic, as shown in Table 1:

The interpretation of the variables and their effects goes as follows:

- Gender allows for a gap in the Adoption of online learning between males and females.
- Age affects the costs and benefits of Adoption, depending on the age range.
- Habitat increases or reduces the benefits of adopting online education, depending on the size of the population to which the individual belongs.
- Household Members may affect the benefits and the costs of online education.
- Nationality might signal effects dependent on costs (e.g., different language) and possible benefits (e.g., access to a broader range of services, education in your language).
- Education, Internet Trust and Digital Skills<sup>1</sup> are expected to diminish the costs of using online education and signal higher potential benefits.
- Economic: Income is supposed to increase the benefits of online education.

The following is the basis for an estimable model of the decision to adopt online education:

$$P(\text{Online\_education}_{it}) = f(\text{Gender}_{it}, \text{Age}_{it}, \text{Education}_{it}, \text{Digital Skills}_{it}, \text{Income}_{it}, \text{Habitat}_{it}, \text{Household\_members}_{it}, \text{Internet Trust}_{it}, \text{Nationality}_{it}, \text{uit}) \quad (2)$$

The subindex  $i$  stands for individuals, while  $t$  stands for time.  $P(\text{Online\_education}_{it})$  is the probability of an individual  $i$  adopting online education at time  $t$ . When  $f$  is linear, we are in the linear probability model, LPM. In the logistic case,  $f$  is nonlinear.

### 5.1. Linear probability model (1)

The linear probability model (LPM) constitutes a convenient approach to modeling a binary variable and its determinants. However, it may suffer the limitation that a predicted probability could lie outside the  $[0, 1]$  interval and present heteroskedasticity. These problems can be solved by truncating to 0 or 1 predictions below 0 or above 1. On the other hand, heteroskedasticity can be treated using a consistent variance-covariance matrix. Theoretical properties and applications of the LPM can be found in Wooldridge (2015), p. 224–273.

In contrast, the LPM shows reliable estimates of signs and the significance of the coefficients. Moreover, it provides a straightforward interpretation of the estimated coefficients in terms of differences in predicted probabilities (Aldrich & Nelson, 1984; Horrace & Oaxaca, 2006).

Each significant coefficient of the independent variables in the LPM represents changes in the probability that a person adopts online education for private use; everything else held constant.

### 5.2. Logistic model (2)

The logistic model is an alternative to the LPM. The logistic produces predicted probabilities in the interval  $[0, 1]$  by employing a nonlinear logistic functional form widely used in the literature (Maddala, 1986; Manski & McFadden, 1981).

The signs of the coefficients and  $z$  statistics are compatible in sign and significance with those of the LPM, although they are not directly comparable.

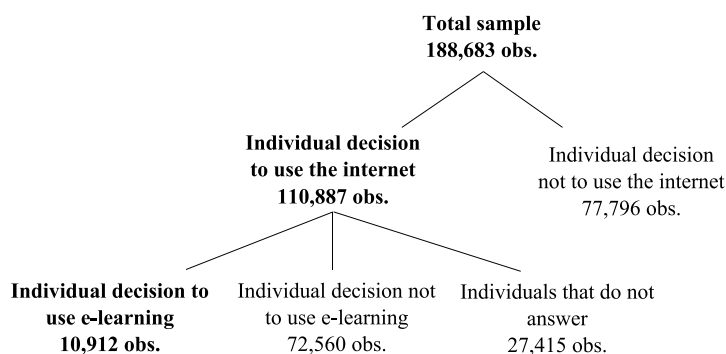
<sup>1</sup> Digital Skills is a self-elaborated index, based on the answers where the respondent declares whether he or she used specific internet services and/or performed specific computer and internet-related tasks.

**Table 3**

Polychoric correlation matrix

Source: Self-elaboration based on INE (2020).

	Elearn.	Int. Access	Int. User	Gender	Age	Hh. Memb.	Pop. Size	Nation	Educ	Dig. Skills	Int. Trust
Elearn.	1										
Int. Access	0.30	1									
Int. User	0.90	0.48	1								
Gender	-0.02	0.00	-0.02	1							
Age	-0.19	0.05	-0.18	0.02	1						
Hh.Memb.	0.03	0.14	0.07	0.02	-0.31	1					
Pop. Size	-0.08	-0.12	-0.10	0.03	-0.05	0.07	1				
Nation	-0.15	-0.24	-0.03	-0.03	-0.16	0.06	-0.01	1			
Educ	0.40	0.25	0.33	-0.05	-0.06	-0.01	-0.16	-0.14	1		
Dig. Skills	0.65	0.40	0.77	0.08	-0.37	0.08	-0.12	-0.16	0.52	1	
Int. Trust	0.19	0.18	0.30	0.07	-0.14	0.05	-0.05	-0.07	0.22	0.36	1
Income	0.24	0.41	0.29	0.07	0.06	0.15	-0.12	-0.29	0.43	0.37	0.19

**Fig. 4.** The decision process of the internet and online education adoption 2008–2019.

### 5.3. Heckman selection model (3)

The Heckman selection model (Heckman, 1977), shown in Table 3, under (3), uses a two-stage approach to the modelling. First, it uses a selection equation to model the decision to be an internet user, as shown in Table 1A of the Appendix.<sup>2</sup> Second, it uses another equation to explain the decision to adopt online education, given that an individual is an internet user.

### 5.4. Estimation

The estimation results are presented in Table 4 and Fig. 5. The second column of Table 4 includes the categories of each variable. The following columns marked with (1) show the estimates and z statistics of the LPM of adoption. The columns marked with (2) contain the results for the logistic model of adoption, and the last two columns marked with (3), include the estimates and z statistics for the adoption equation of the Heckman model. The selection equation of the Heckman model is deferred to the Appendix to facilitate the comparison between the three adoption equations (1-3). The estimations employ panel data techniques, using Stata 17 on unbalanced panel data samples.

Fig. 5 is essentially a robustness check. Both models agree on the sign and significance. The coefficients in the Heckman model tend to be higher in absolute value than those of the LPM. We can conclude that the Heckman model is somewhat robust to this specification change, except for the sizes of the coefficients.

### 5.5. Specifics on estimation. The LPM

The F test of homogeneity of individual coefficients rejects the null, with a computed value of 1.35 and a p-value of 0.0000, so it is more appropriate to use panel data methods compared to pooled regression.

The LPM is applied to the subsample of internet users. It models the decision of whether to adopt e-learning or not. The results shown in Table 4 suggest that being male decreases the probability of adopting e-learning by 3.3%. Belonging to age groups 25–34,

<sup>2</sup> The selection equation models the probability of an individual to adopt the use of the internet at a given point in time. This probability depends on the individual characteristics shown in equation A1 of the Appendix. The excluded variable in the Heckman model is population size.

**Table 4**  
Models of Adoption of online education by individual internet users. Panel data (2008–2019).

Dep. Variable:	Online Education	(1) LPM		(2) Logistic		(3) Heckman	
		Coef.	z	Coef.	z	Coef.	z
Gender	Male	<b>-0.033</b>	<b>-13.30</b>	<b>-0.442</b>	<b>-12.22</b>	<b>-0.030</b>	<b>-11.96</b>
Age	25–34	<b>0.022</b>	<b>4.380</b>	<b>0.225</b>	<b>3.280</b>	0.007	1.010
	35–44	<b>0.016</b>	<b>3.320</b>	<b>0.150</b>	<b>2.280</b>	-0.005	-0.710
	45–54	<b>0.026</b>	<b>5.280</b>	<b>0.303</b>	<b>4.440</b>	0.000	-0.040
	55–64	0.003	0.620	-0.004	-0.060	<b>-0.031</b>	<b>-4.240</b>
	65–74	<b>-0.019</b>	<b>-2.980</b>	<b>-0.622</b>	<b>-5.300</b>	<b>-0.062</b>	<b>-7.450</b>
	75+	-0.015	-1.590	<b>-0.853</b>	<b>-3.460</b>	<b>-0.143</b>	<b>-10.43</b>
Education	Secondary	-0.004	-0.940	<b>0.227</b>	<b>2.240</b>	<b>0.042</b>	<b>10.260</b>
	Bachelor's Degree	<b>0.030</b>	<b>5.670</b>	<b>0.708</b>	<b>6.690</b>	<b>0.092</b>	<b>17.180</b>
	Master's or PhD	<b>0.066</b>	<b>12.350</b>	<b>1.039</b>	<b>9.830</b>	<b>0.127</b>	<b>23.120</b>
Digital Skills	Medium	<b>0.039</b>	<b>11.760</b>	<b>1.754</b>	<b>20.430</b>	<b>0.184</b>	<b>33.230</b>
	High	<b>0.154</b>	<b>42.400</b>	<b>3.183</b>	<b>35.820</b>	<b>0.289</b>	<b>47.840</b>
	Very High	<b>0.439</b>	<b>90.280</b>	<b>5.097</b>	<b>49.560</b>	<b>0.565</b>	<b>71.090</b>
Income	Medium	0.000	0.130	0.008	0.140	<b>0.017</b>	<b>5.170</b>
	High	0.000	-0.100	0.006	0.110	<b>0.023</b>	<b>6.110</b>
	Very High	0.009	1.910	0.098	1.530	<b>0.032</b>	<b>6.520</b>
	Constant	0.013	1.970	<b>-5.832</b>	<b>-40.05</b>	-0.195	-21.09
Number of Observations		70,616		70,616		138,879	
Selected						70,616	
Non-selected						68,263	
Groups		45,742		45,742		78,224	
Wald $\chi^2$		14,074.34; DF: 16		4,136.24; DF: 16		9,067.62; DF: 16	

F of individual effects = 0: F(45741, 24858) = 1.35 Prob > F = 0.0000.

Notes: All estimations are performed using panel data with Stata 17. Coefficients and z statistics (significant at 5%) are represented in bold due to the lack of space for using stars (\*). Equation (1) is a random-effects linear probability model, equation (2) is a logistic model, and equation 3 incorporates the adoption equation of Heckman's selection model, using panel data techniques in all cases. The selection equation is reported in the Appendix. Base categories are female, age 16–24, primary or no education, low digital skills, and low income. The test of correlation between the errors of the equations is 0.95. Internet Trust turned out to be insignificant.

35–44, and 45–54 increases the probability by 2.2%, 1.6%, and 2.6%, respectively, while being in the 65–74 age group reduces the probability by 1.9%, everything else held constant.

Higher levels of education are related to higher probabilities of adoption, going up by 3% for Bachelor and 6.6% for a Master's or Ph.D. Higher digital skills increase the probability of adopting e-learning by 3.9%, 15.4%, and 43.9% for medium, high, and very high. The different levels of income do not appear significant.

Being male is associated with a lower probability of adopting online education. In addition, belonging to age groups 25–34, 35–44, and 45–54 increases the probability, while being in the 65–74 and 75+ age groups reduces the probability; everything else held constant.

The higher the level of education, the higher the likelihood of Adoption. Higher levels of digital skills increase the probability of using e-learning. Differences in income also appear insignificant in this specification. We retain this basic specification for robustness checks and comparison with other models. It answers the question of the specific determinants of those consumers who are already Internet users and have to decide whether to adopt online education. The models were also estimated, including time dummies for each year. As a result, the point estimates and t-statistics of the coefficients of interest are very similar to those presented here.

### 5.6. The logistic model

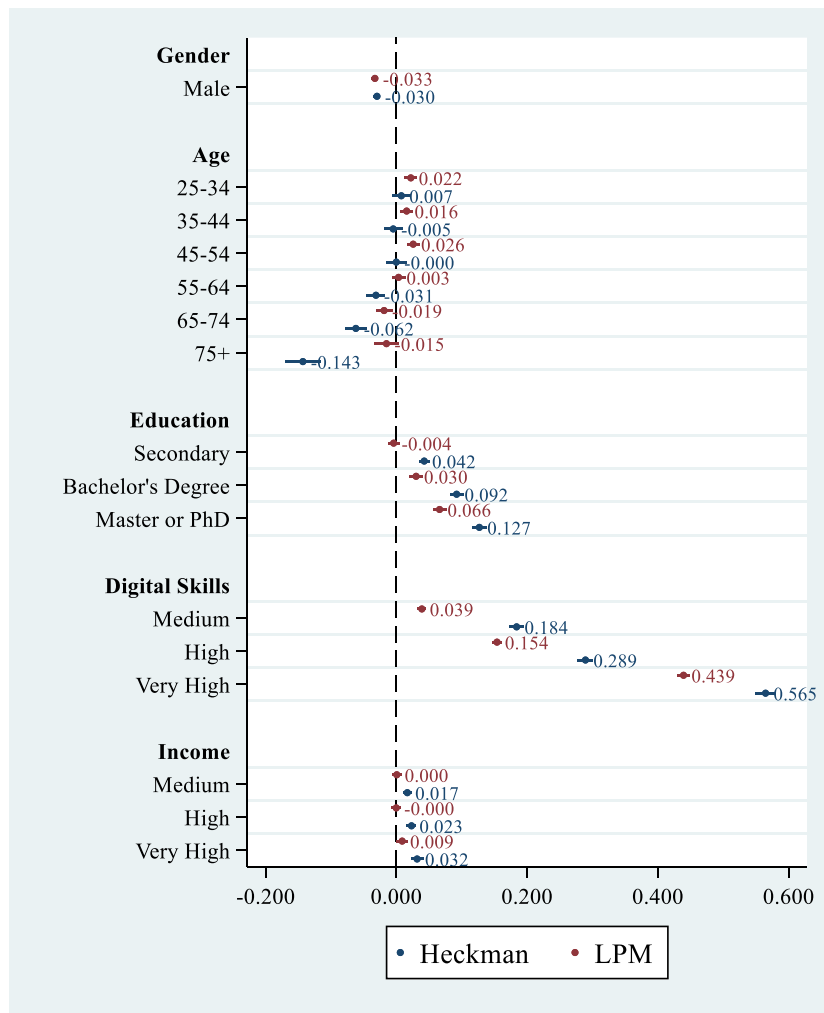
The logistic model is applied to the subsample of internet users. It models the decision of adopting e-learning or not for internet users. Table 4 shows the results of estimating this model 2, the logistic. It can be observed that the signs and z-statistics are in agreement with the other two models, while the point estimates are not directly comparable with the LPM or the Heckman model. This model was estimated again using panel data techniques using Stata 17. All results are available upon request. We will concentrate on LPM and Heckman. The software treated missing values adequately, ignoring the observations that had missing values. Attrition was also treated adequately by Stata 17.

The first equation of this model applies to the whole sample and models those who use and do not use the internet. The second equation models the adoption of e-learning from the sample of internet users.

In model (3) being male decreases the probability by 3%. Belonging to age groups of 55–64, 65–74, and more than 75 reduces the probability of Adoption by 3.1%, 6.2%, and 14.3%, respectively, and for the groups 25–34, 35–44, and 45–54, there is no evidence in this model and sample that the differences are significantly distinct from zero.

Higher education categories are positively related to online education, increasing the probability of Adoption by 4.2% for secondary education, 9.2% for bachelor's degree, and 19.7% for master's and Ph.D.

Higher digital skills increase the probability of adopting e-learning by 18.4%, 28.9%, and 56.9% for medium, high, and very high



**Fig. 5.** Models of Adoption of online learning by individual internet users. Coefficients and confidence intervals of the linear probability and Heckman models. Panel 2008–2019. Notes: Coefficients are depicted by circles and 95% confidence intervals by attached lines. Base categories are female, age 16–24, primary or no education, low digital skills, and low income. LPM in red. Heckman in blue. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

levels. Income levels are significant, but with minor effects and increase the probability of using e-learning by 1.7%, 2.3%, and 3.2%, respectively.

### 5.7. The LPM and the Heckman models

Fig. 4 compares the point estimates and the confidence intervals for the LPM and the Heckman models. In many cases, it is apparent that Heckman’s estimates are higher in absolute values. In particular, the effects of education, digital skills, and income are more substantial than the LPM.

For gender, the effect is similar in both cases, around –3%. However, the results are insignificant only in the Heckman model for age groups between 25 and 54. In contrast, they are large, negative, and significant for older age groups in both models.

The estimated values for education are higher, positive, and significant for the Heckman model compared to the LPM. Digital skills also have a large and significant effect on the adoption of online education, and once again, the Heckman model has larger effects than the LPM. In the case of income, the Heckman model can detect positive and significant effects while the LPM cannot.

The previous discussion highlights that all three models give results that are similar in spirit. However, the Heckman model is preferred because it answers the relevant question of the determinants of online education access for the general population, avoiding the sample selection bias. In addition, it incorporates more structure and information and measures the effects of potentially relevant variables. Moreover, using the three models shows that the estimation results are not peculiar to the specification adopted. Therefore, the results are robust concerning the various estimation techniques used.

## 6. Comparison with 2020 after the first wave of Covid-19

This section compares the outcomes observed in 2020 with 2008–2019. The data for 2020 were collected right after the first period of confinement; therefore, the questions tackled here differ from those in the previous sections. A direct comparison is performed due to a possible structural change that may affect some or all the parameters. In that context, a dummy variable for 2020 would reflect a difference in the constant, but not necessarily a structural change in the different coefficients. The 2020 data only reflect part of the year, not the entire period.

Another consideration is that the decision to participate in online learning during the Covid period is due to individual optimization and confinement rules. These rules forced many students to move their learning activities online.

Table 5 shows the penetrations of e-learning by selected variables and categories. The third and fourth columns contain the penetrations in 2019 and 2020. The fifth column shows the average annual growth rate for 2008–2019. Finally, the last column shows the acceleration factor, defined as the ratio between the growth rate in 2020 over the average annual rate of increase from 2008 to 2019. This is meant to measure the change in the speed in 2020 compared with the average speed in 2008–2019.

The last column of Table 5 shows that the average acceleration is around 8.8, similar for both genders. Some categories stand out, like age 16–24, which increases by 24.8 times, and household members five or more, which increases by 13.4 times.

Other significant increases occurred for the categories of population size less than 10,000, which increased by 10.3, and foreigners, which increased by 30. In addition, secondary education increased by a factor of 13.6, and very high income increased by a factor of 10.5.

These figures illustrate the impact of the first wave of Covid-19 and its associated policy measures. Covid-19 increased the adoption of online education by an amount equivalent to the previous nine years. However, this average does not represent the whole picture since heterogeneity exists across the different groups. Table 5 shows how these increases differ by variables and categories.

## 7. Conclusions and recommendations

This section contains four subsections: general conclusions, policy recommendations, caveats, and further research.

**Table 5**

Acceleration: growth rates of adoption. 2020 over 2008–2019 Source: Self-elaboration based on the TICH survey of INE (2020).

Variables	Categories	Penetrations		Average growth rate	Growth rate	Acceleration factor
		2019	2020	2008–2019	2019–2020	2019–2020
Gender	Female	13.2	23.5	8.6	78.0	9.0
	Male	13.3	23.7	9.1	78.2	8.6
Age	16–24	16.7	49.1	7.8	194.0	24.8
	25–34	23.7	40.4	8.1	70.5	8.7
	35–44	19.1	30	10.6	57.1	5.4
	45–54	16.2	24.6	12.1	51.9	4.3
	55–64	8.4	13.5	13.9	60.7	4.4
	65–74	2.3	4.6	17.2	100.0	5.8
	75+	0.4	0.8		100.0	
Household Members	One	11.3	16.2	13.9	43.4	3.1
	Two	8.8	14.5	7.4	64.8	8.7
	Three	14.9	26.5	8.6	77.9	9.0
	Four	18.7	33.6	9.6	79.7	8.3
	Five or more	12.6	30.7	11.0	143.7	13.1
Population Size	100,000+, Capitals	15.3	27.7	9.2	81.0	8.8
	50,000 to 100,000	14.6	22.9	6.8	56.8	8.4
	20,000 to 50,000	12.7	22.1	9.7	74.0	7.7
	10,000 to 20,000	10.5	19.3	9.7	83.8	8.7
	Less than 10,000	10.2	18.8	8.2	84.3	10.3
Nationality	Spanish	13.9	24.1	9.4	73.4	7.8
	Foreigner	8.0	19.4	4.8	142.5	30.0
Education	None or Primary	1.0	1.6	11.6	60.0	5.2
	Secondary	9.9	20.8	8.1	110.1	13.6
	Bachelor's Degree	22.5	36.3	10.1	61.3	6.1
	Master's or PhD	31.0	48.3	6.4	55.8	8.7
Digital Skills	Low	0.6	0.7	6.5	16.7	2.6
	Medium	6	8.9	−1.9	48.3	−25.6
	High	17.9	30.3	−2.4	69.3	−28.8
	Very High	49.9	69.5	−0.9	39.3	−44.6
Internet Trust	Low	9.4	22		134.0	
	Medium	17.8	31.3		75.8	
	High	28.3	30.4		7.4	
Income	Low	6.1	12.3	12.9	101.6	7.9
	Medium	10.6	16.8	9.0	58.5	6.5
	High	18.3	29.8	8.6	62.8	7.3
	Very High	26.6	46.5	7.1	74.8	10.5

### 7.1. General conclusions

This article focuses on identifying and measuring the effects of the drivers and impediments to adopting online learning in Spain from 2008 to 2020. It employs the ICT-H survey of the National Institute of Statistics, which allows the construction of a panel database of 210,370 individuals (including 2020).

We use a well-established economic model for online education adoption to estimate the econometric models. The Heckman model is the preferred specification. The estimations reveal the importance of gender, age, education, digital skills, and income in adopting online education.

Using the Heckman model, the results reveal that being male (−3.0%) and older (up to −14.3%) are impediments to adopting online education (relative to the basis category). Moreover, being more educated (up to +12.7%), having better digital skills (up to +56.5%), and having higher income (up to +3.2%) are drivers for adopting online education.

A general conclusion is that the adoption of online education depends on individual factors that can help explain and understand the behavior related to its Adoption. Of all the variables, the one that makes the most difference is Digital Skills. This is apparent in the sizes of their estimated coefficients across the three equations.

The first wave of Covid-19 increased the speed of adopting online education by a factor of 9 on average; however, the increases were heterogeneous across different groups of consumers.

The general conclusions of this article are in the spirit of previous literature; however, the contribution of this article is centered on identifying the relevant equations and variables, the point estimation of their coefficients and standard errors, and the significance and quality of the joint models. All this has been done using a large and representative panel data set, using panel techniques. This methodology ensures consistent and efficient estimates that support results that constitute a contribution of this paper, in particular for the case of Spain. Moreover, the results are likely similar to those of other European countries.

### 7.2. Policy recommendations

Digital skills cause a sizeable divide in the adoption of online education, as apparent in [Tables 2 and 4](#). Therefore, this variable has the most impact on the probability of adoption. Moreover, it is the only policy that can influence in the short or medium run. Therefore, the Information and Communications policy must focus on this variable.

The European Commission often cites this as the main reason for low adoption levels and has pointed out its importance in many services in Spain ([European Commission, 2019](#)). Various initiatives for improving Digital Skills are in place by the European Commission, the Central government, and Autonomous governments ([Garín-Muñoz et al., 2019](#); [Valarezo et al., 2020](#)).

Other initiatives often relate to a specific digital divide, like disfavored children, addressed in [Garín-Muñoz et al. \(2020\)](#) and [Red.es \(2021\)](#). They propose supplying computers, tablets, and internet connections to disfavored children to enable them to participate in online education, which was mandatory in many territories for several weeks or months during 2020. This initiative is currently being implemented in several autonomous communities by Red.es (2021) and is expected to cover all geographical areas.

### 7.3. Caveats

The database for this paper is extensive and representative of the whole population of Spain and its autonomous communities; however, it is declarative. This problem may introduce recollection bias in the estimates.

The data for year 2020 mainly capture the surge in online education from March to August (See [table A2](#) in the Appendix). The survey refers to the respondent person (an individual aged 16 or more). At the same time, children and higher education students were obliged to adopt online learning overnight. Data on the continuation and outcome of online education is not available in this survey.

### 7.4. Further research

The assessment of Covid-19's effect on online education in Spain is a promising area of research. The 2020 data for the ICT-H survey is partially helpful since it was collected from March to September 2020. This database covers the period right after the first pandemic wave, but before the successive waves.

### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### Acknowledgments

Presented at the 23rd ITS Biennial Conference (Online Conference, Gothenburg, June 2021) and ICAE's Research Seminar (Madrid, June 2021).

The collaboration of INE is gratefully acknowledged. Research funded by the Autonomous Community of Madrid, Spain (Project: H2019/HUM-5859, MIMA-CM, Mercados, Innovación y Medio Ambiente), and National R&D Plan, Project: PID2019-107161GB-C32, 2020–2023.

## Appendix

**Table A1**  
Internet user selection equation of Heckman's Model. Panel data (2008–2019).

		(3) Heckman	
Dep. Variable	Internet User	Coef.	z
Gender	Male	0.016	0.810
Age	25–34	<b>−1.124</b>	<b>−7.51</b>
	35–44	<b>−1.458</b>	<b>−11.41</b>
	45–54	<b>−1.772</b>	<b>−19.79</b>
	55–64	<b>−2.299</b>	<b>−44.87</b>
	65–74	<b>−2.848</b>	<b>−70.00</b>
	75+	−3.735	.
Education	Secondary	<b>0.849</b>	<b>16.88</b>
	Bachelor's Degree	<b>1.666</b>	<b>13.870</b>
	Master or Ph.D.	<b>1.774</b>	<b>13.340</b>
Digital Skills	Medium	<b>4.350</b>	<b>10.530</b>
	High	<b>2.859</b>	<b>19.130</b>
	Very High	<b>3.990</b>	<b>22.020</b>
Income	Medium	<b>0.535</b>	<b>13.740</b>
	High	<b>0.979</b>	<b>13.270</b>
	Very High	<b>1.024</b>	<b>11.380</b>
Population size	50,000 to 100,000	<b>0.100</b>	<b>3.500</b>
	20,000 to 50,000	<b>−0.061</b>	<b>−2.520</b>
	10,000 to 20,000	<b>−0.257</b>	<b>−7.520</b>
	Less than 10,000	<b>−0.428</b>	<b>−9.720</b>
	Constant	0.065	0.340
Observations		138,879	
Selected		70,616	
Non-selected		68,263	
Groups		78,224	
Wald $\chi^2$		9,067.62; DF: 16	

Notes: Coefficients and z statistics significant at 5% are represented in bold. The z statistic is not reported for age 75+. Internet user in the last 3 months is the selection equation in Heckman's sample selection procedure (Equation 3). The adoption equation is reported in Table 3. Base categories are female, age 16–24, primary or no education, low digital skills, low income, population size of 100,000+ and province capitals. The equations are estimated by random effects using the command xheckman in Stata 17. The correlation between the errors of the equations is 0.95.

**Table A2**  
Dates interviews 2020.  
Source: Self elaboration based on the ICT-H survey of INE (2020).

Month	Interviews	Percentage
February	1	0%
March	997	6%
April	382	2%
May	4891	32%
June	4174	27%
July	3156	21%
August	1275	8%
September	466	3%
TOTAL	15,342	100%

## References

- Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50(2), 179–211. [https://doi.org/10.1016/0749-5978\(91\)90020-T](https://doi.org/10.1016/0749-5978(91)90020-T)
- Ajzen, I., & 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, I., & Fishbein, M. (1980). *Understanding attitude and predicting social behavior*. NJ: Prentice-Hall. Prentice-Hall.
- Aldrich, J. H., & Nelson, F. D. (1984). *Linear probability, logit, and probit models*. SAGE.
- Andrews, R., & Haythornthwaite, C. (2007). *The SAGE handbook of E-learning research*. SAGE.
- Artle, R., & Averous, C. (1973). The telephone system as a public good: Static and dynamic aspects. *Bell Journal of Economics and Management Science*, 4(1), 89–100. <https://doi.org/10.2307/3003140>. JSTOR.

- Bates, A. W. (Tony) (1997). The impact of technological change on open and distance learning. *Distance Education*, 18(1), 93–109. <https://doi.org/10.1080/0158791970180108>
- Billon, M., Crespo, J., & Lera-Lopez, F. (2021). Do educational inequalities affect internet use? An analysis for developed and developing countries. *Telematics and Informatics*, 58, Article 101521. <https://doi.org/10.1016/j.tele.2020.101521>
- Brown, P., Hesketh, A., & Williams, S. (2003). Employability in a knowledge-driven economy. *Journal of Education and Work*, 16(2), 107–126. <https://doi.org/10.1080/1363908032000070648>
- Bryson, J. R., & Andres, L. (2020). Covid-19 and rapid adoption and improvisation of online teaching: Curating resources for extensive versus intensive online learning experiences. *Journal of Geography in Higher Education*, 44(4), 608–623. <https://doi.org/10.1080/03098265.2020.1807478>
- Davis, F. 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, F. D., & Warshaw, P. R. (1989). User acceptance of information technology: System characteristics, user perceptions, and behavioral impacts. *Management Science*, 35(8), 982–1003.
- 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>
- van Deursen, A. J. A. M., Helsper, E., & Eynon, R. (2014). *Measuring digital skills: From digital skills to tangible outcomes* [Monograph]. <http://www.oii.ox.ac.uk/>.
- DiMaggio, P., & Hargittai, E. (2001). *From the "digital divide" to "digital inequality": Studying internet use as penetration increases (No. 47; working papers)*. Princeton University, Woodrow Wilson School of Public and International Affairs, Center for Arts and Cultural Policy Studies. <https://ideas.repec.org/p/pri/cpanda/workpap15.html>.
- Drasgow, F. (1988). Polychoric and polyserial correlations. In *Encyclopedia of Statistical Sciences*. John Wiley & Sons, Inc. <https://doi.org/10.1002/0471667196.ess2014.pub2>.
- European Commission. (2016). *A new comprehensive digital skills indicator. Digital Single Market*. <https://ec.europa.eu/digital-single-market/en/news/new-comprehensive-digital-skills-indicator>.
- European Commission. (2019). *Digital economy and Society index (DESI). 2019 country report. Spain* [Text]. <https://ec.europa.eu/digital-single-market/en/scoreboard/spain>.
- Fairlie, R. W. (2004). Race and the digital divide. *Contributions to Economic Analysis and Policy*, 3(1). <https://doi.org/10.2202/1538-0645.1263>
- 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>
- Garín-Muñoz, T., Pérez-Amaral, T., & Garín-Muñoz, M. (2020). *Brecha digital infantil y el COVID-19. Nada es Gratis*. May 11 <https://nadaesgratis.es/admin/brecha-digital-infantil-y-el-covid-19>.
- Garín-Muñoz, T., Pérez-Amaral, T., & Valarezo, Á. (2022). Evolution of the internet gender gaps in Spain and effects of the Covid-19 pandemic. *Telecommunications Policy*, 46(8). <https://doi.org/10.1016/j.telpol.2022.102371>. Scopus.
- Hargittai, E. (2001). *Second-level digital divide: Mapping differences in people's online skills*. ArXiv:Cs/0109068 <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)*. National Bureau of Economic Research. <https://doi.org/10.3386/w0172>. Working Paper No. 172).
- 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* (pp. 29–52). Emerald Group Publishing Limited. <https://doi.org/10.1108/S2050-20602015000010002>, 10.
- Horrace, W. C., & Oaxaca, R. L. (2006). Results on the bias and inconsistency of ordinary least squares for the linear probability model. *Economics Letters*, 90(3), 321–327. <https://doi.org/10.1016/j.econlet.2005.08.024>
- Hvorecký, J. (2004). Can E-learning break the digital divide? *European Journal of Open, Distance and E-Learning*, 7(2). Article 2 <https://old.eurodl.org/?p=archives&year=2004&halfyear=2&article=143>.
- INE, I. N. de E. (2020). *Encuesta sobre equipamiento y uso de tecnologías de información y comunicación en los hogares/Últimos datos*. INE. [https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=estadistica\\_C&cid=1254736176741&menu=ultiDatos&idp=1254735976608](https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=estadistica_C&cid=1254736176741&menu=ultiDatos&idp=1254735976608).
- Jacqmín, J. (2019). Providing MOOCs: A FUN way to enroll students? *Information Economics and Policy*, 48, 32–39. <https://doi.org/10.1016/j.infoecopol.2018.10.002>
- Journell, W. (2007). The inequities of the digital divide: Is E-learning a solution? *E-learning and Digital Media*, 4(2), 138–149. <https://doi.org/10.2304/elea.2007.4.2.138>
- Khalil Moghaddam, B., & Khatoun-Abadi, A. (2013). Factors affecting ICT adoption among rural users: A case study of ICT center in Iran. *Telecommunications Policy*, 37(11), 1083–1094. <https://doi.org/10.1016/j.telpol.2013.02.005>
- Kolenikov, S. (2016). *Polychoric, by any other 'namelist'*. [https://www.stata.com/meeting/chicago16/slides/chicago16\\_kolenikov.pdf](https://www.stata.com/meeting/chicago16/slides/chicago16_kolenikov.pdf).
- Kraemer, K., Ganley, D., & Dewan, S. (2005). Across the digital divide: A cross-country multi-technology analysis of the determinants of IT penetration. *Journal of the Association for Information Systems*, 6(12). <https://doi.org/10.17705/1jais.00071>
- Kridel, D. J., Rappoport, P. N., & Taylor, L. D. (1999). An econometric study of the demand for access to the internet. In D. G. Loomis, & L. D. Taylor (Eds.), *The Future of the Telecommunications Industry: Forecasting and Demand Analysis* (pp. 21–41). Springer US. [https://doi.org/10.1007/978-1-4615-4643-6\\_3](https://doi.org/10.1007/978-1-4615-4643-6_3).
- Larreaendo-Joerns, J., & Leinhardt, G. (2006). Going the distance with online education. *Review of Educational Research*, 76(4), 567–605. <https://doi.org/10.3102/00346543076004567>
- Lera-Lopez, F., Billon, M., & Gil, M. (2011). Determinants of internet use in Spain. *Economics of Innovation and New Technology*, 20(2), 127–152.
- Maddala, G. S. (1986). *Limited-dependent and qualitative variables in econometrics*. Cambridge University Press.
- Manski, C. F., & McFadden, D. (1981). *Structural analysis of discrete data with econometric applications*. Cambridge, MA: MIT press.
- Money, W. H., & Dean, B. P. (2019). Incorporating student population differences for effective online education: A content-based review and integrative model. *Computers & Education*, 138, 57–82. <https://doi.org/10.1016/j.compedu.2019.03.013>
- National Telecommunications and Information Administration, NTIA. (1995). *Falling THROUGH THE net: A Survey of the "have nots" in Rural and Urban America*. <https://www.ntia.doc.gov/ntiahome/fallingthru.html>.
- Norris, P. (2001). *Digital divide: Civic engagement, information poverty, and the internet worldwide*. Cambridge University Press.
- OECD. (2001). *Understanding the digital divide (OECD digital economy paper No. 49)*. OECD Publishing. <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>
- Panigrahi, R., Srivastava, P. R., & Sharma, D. (2018). Online learning: Adoption, continuance, and learning outcome—a review of literature. *International Journal of Information Management*, 43, 1–14. <https://doi.org/10.1016/j.ijinfomgt.2018.05.005>
- Pérez-Amaral, T., Valarezo, A., López, R., & Garín-Muñoz, T. (2021). Digital divides across consumers of internet services in Spain using panel data 2007–2019. *Narrowing or not? Telecommunications Policy*, 45(2), Article 102093. <https://doi.org/10.1016/j.telpol.2020.102093>
- Perna, L. W., Ruby, A., Boruch, R. F., Wang, N., Scull, J., Ahmad, S., & Evans, C. (2014). Moving through MOOCs: Understanding the progression of users in massive open online courses. *Educational Researcher*, 43(9), 421–432. <https://doi.org/10.3102/0013189X14562423>. Scopus.
- 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*, 5(2), 651–669. <https://doi.org/10.2307/3003127>. JSTOR.
- Rappoport, P., Kridel, D. J., Taylor, L. D., & Duffy-Demo, K. T. (2003). Residential demand for access to the internet. *The International Handbook of Telecommunications Economics*, 1, 55–72.
- Reddy, A. B., Jose, S., & R. V. (2021). Of access and inclusivity digital divide in online education. In ArXiv e-prints. <https://ui.adsabs.harvard.edu/abs/2021arXiv210710723R>.
- Red.es. (2021). *Programa "Educa en Digital"*. <https://www.red.es/redes/gl/node/16592>.
- Riggins, F. J., & Dewan, S. (2005). The digital divide: Current and future research directions. *Journal of the Association for Information Systems*, 6(12), 4.

- Rogers, E. M. (2003). *Diffusion of innovations* (5th ed.). Simon and Schuster.
- Rohlf, J. (1974). A theory of interdependent demand for a communications service. *Bell Journal of Economics and Management Science*, 5(1), 16–37. <https://doi.org/10.2307/3003090>. JSTOR.
- Simonson, M., Schlosser, C., & Orellana, A. (2011). Distance education research: A review of the literature. *Journal of Computing in Higher Education*, 23(2), 124. <https://doi.org/10.1007/s12528-011-9045-8>
- Sims, J., Vidgen, R., & Powell, P. (2008). E-learning and the digital divide: Perpetuating cultural and socioeconomic elitism in higher education. *Communications of the Association for Information Systems*, 22(1). Article 1 <http://aisel.aisnet.org/cais/vol22/iss1/23/>.
- Squire, L. (1973). Some aspects of optimal pricing for telecommunications. *Bell Journal of Economics and Management Science*, 4(2), 515–525. <https://doi.org/10.2307/3003051>. JSTOR.
- Taylor, L. D. (1994). *Telecommunications demand in theory and practice*. Springer Netherlands. <https://www.springer.com/us/book/9780792323891>.
- 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., López, R., & Pérez Amaral, T. (2020). Adoption of E-commerce by individuals and digital-divide. In J. Alleman, P. N. Rappoport, & M. Hamoudia (Eds.), *Applied Economics in the Digital Era: Essays in Honor of Gary Madden* (pp. 103–134). Springer International Publishing. [https://doi.org/10.1007/978-3-030-40601-1\\_4](https://doi.org/10.1007/978-3-030-40601-1_4).
- 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). In R. W. Crandall, & J. H. Alleman (Eds.), *Broadband: Should We Regulate high-speed internet access? Aei-brookings joint center for Regulatory studies* *The demand for Bandwidth: Evidence from the INDEX Project* (pp. 39–56). American Enterprise Institute.
- 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>
- 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 Economics Letters*, 15, 1015–1018.
- Volery, T., & Lord, D. (2000). Critical success factors in online education. *International Journal of Educational Management*, 14(5), 216–223. <https://doi.org/10.1108/09513540010344731>
- Walker, E. C., & Kelly, E. (2007). Online instruction: Student satisfaction, kudos, and pet peeves. *Quarterly Review of Distance Education*, 8(4), 309–319 (Scopus).
- Wei, K.-K., Teo, H.-H., Chan, H. C., & Tan, B. C. Y. (2010). *Conceptualizing and testing a social cognitive model of the digital divide*. Information Systems Research. <https://doi.org/10.1287/isre.1090.0273>
- Wooldridge, J. M. (2015). *Introductory econometrics: A modern approach* (6th ed.). Cengage Learning.