Technological Forecasting & Social Change 198 (2024) 123028

Contents lists available at ScienceDirect

Technological Forecasting & Social Change

journal homepage: www.elsevier.com/locate/techfore

How to keep your information secure? Toward a better understanding of users security behavior

Álvaro Arenas [a], Gautam Ray [b], Antonio Hidalgo [c], Alberto Urueña [c,*]

[a] IE Business School, IE University, c/María de Molina, 12, 28006 Madrid, Spain
[b] Carlson School of Management, University of Minnesota, 321 19th Ave S, Minneapolis, MN 55455, USA
[c] Universidad Politécnica de Madrid, ETSI Industriales, c/José Gutiérrez Abascal, 2, 28006 Madrid, Spain

ARTICLE INFO

Keywords:
Information security
Routine activity theory
Malware infection
Cyberattacks
Personal computer

ABSTRACT

Use of computers and the Internet is an integral part of our lives, with business becoming more digital. As a result, individuals are using their home computers to perform diverse tasks and to store sensitive data. This paper investigates the relative efficacy of two strategies to protect home computers from security threats: security tools and security activities. For the analysis, we collected data from over 1900 individuals in Spain, following an approach combining self- reported data, via an online survey, with actual data collected directly from home users' computers. The main contribution of the paper is to provide a model, based on routine activity theory, explaining the role of security tools and security activities in protecting personal computers from malware infection, thus offering an in-depth understanding of users' security behavior. Using multivariate, logit and probit regressions, our study reveals that having security tools is positively related with higher risk activities and more infections, while pursuing security activities reduces malware infections. These results have important implications for policy makers and organizations, reinforcing the view that security tools are not sufficient to protect users from malware infection, and the need to develop security education and awareness programs for computer users.

1. Introduction

Computers and the Internet are integral parts of our lives. We use them for communication via email or instant messaging, to participate on social networking sites, to bank, and to buy products online, among other services. Such dependency on computers and the Internet has made users increasingly vulnerable and exposed to several threats, including information security threats. In the last decade, there has been an increase in the number of cyber-attacks on personal computers, and reports from consultants and the press state that recovering from computer malware and other security threats – such as credential reuse, denial-of-service attack, phishing, and SQL injection attacks (Humayun et al., 2020) – poses significant financial and social costs (CompTIA, 2022; Fleck, 2022). The exposure of confidential information or privacy-sensitive data during security breach incidents can result in serious losses to both users and organizations (Ou et al., 2022). This situation has been exacerbated by the increase in tele-working, tele-medicine, and tele-education, as part of the post-pandemic new normal (Lallie et al., 2021).

Home computer users are recognized as a point of weakness in achieving Internet security because dedicated technical staff do not maintain their computers, as is the case for employees within organizations (Anderson and Agarwal, 2010). A domestic user who purchases a computer may need to install additional software, select an appropriate password, and define particular settings for applications. Moreover, the user may need to perform regular updates to the software. Consequently, studies have concentrated on analyzing the factors that influence the security behavior of domestic users (Furnell et al., 2007; Liang and Xue, 2010; Mills and Sahi, 2019). These studies have made notable contributions to understand the security behavior of domestic users. However, most of these studies have focused on behavioral intentions and not on actual behaviors. A common methodological limitation is that many studies use a variable measurement strategy based exclusively on intentions. Measures of behavioral intentions based on self-reports may incorporate problems with recall or social desirability. Some studies have shown that behavioral intentions have systematic biases compared to real behaviors (Parry et al., 2021; Kormos and Gifford, 2014). In addition, when users are asked about potentially sensitive issues, there is

* Corresponding author at: Universidad Politécnica de Madrid, ETSI Industriales, c/José Gutiérrez Abascal, 2, 28006 Madrid, Spain.
E-mail addresses: alvaro.arenas@ie.edu (Á. Arenas), gautamr@umn.edu (G. Ray), antonio.hidalgo@upm.es (A. Hidalgo), alberto.uruena@upm.es (A. Urueña).

https://doi.org/10.1016/j.techfore.2023.123028
Received 1 June 2023; Received in revised form 12 November 2023; Accepted 13 November 2023
Available online 27 November 2023

a risk of systematically obtaining responses that are biased downward due to social desirability (Crossler et al., 2013). The present study follows an alternative approach where, in addition to collecting data through a survey, we also collect actual data by installing software in users' computers, with the approval of participants, to evaluate the actual risk and the security incidents suffered by the computer. Another important aspect is the study of the user's role in improving security aspects; this is still an emerging area of research interest that needs further attention. While many studies have aimed to provide technical advancements to improve software security, it is essential to understand the socio-technical aspects in order to identify the prevailing issues and improving the effectiveness of the software security process (Dissanayake et al., 2022). These are significant gaps that need addressing.

Guided by Routine Activity Theory (RAT) (Cohen and Felson, 1979), the goal of this paper is to understand the relative efficacy of security tools used versus the security activities performed by domestic users on the security incidents suffered by the user. We focus on one type of security incident, malware infection, since malware continues to be a major threat on the Internet and the underlying root cause of most Internet security problems (Symantec, 2022). Cohen and Felson's (1979) RAT is an environmental explanation of crime, where the behavioral patterns and intersections of people in time and space influence when and where crimes occur. The theory suggests that a suitable target and the presence or absence of capable guardians influences criminal activities. According to RAT, a capable guardian is a person or object that discourages a crime from taking place; guardians could be physical security measures such as security cameras and alarm systems, or they could be personal security activities such as locking doors (Tseloni et al., 2004). In the present study, we associate physical guardians with security tools such as antivirus software, firewall, antispam software, etc., and contrast the usage of security tools with users' security activities such as changing passwords, clearing browser history, etc. The analysis using data from over 1900 individuals in Spain collected via a survey and from the scan of individuals' computers suggests that security activities have a greater impact in mitigating security incidents than security tools do.

By using different types of regressions (multivariate, logit, probit), our study reveals that having security tools is positively related with higher risk activities and more infections, while pursuing security activities reduces malware infections. These results have important implications for policy makers and organizations, reinforcing the view that security tools are not sufficient to protect users from malware infection, and the need to develop security education and awareness programs for computer users.

The remainder of this paper is organized as follows. Section 2 describes the theoretical background related to the topic. Sections 3, 4, and 5 describe the research framework, hypothesis development, and the results, respectively. Section 6 discusses the work, Section 7 concludes the paper, and presents the strengths, limitations, and future research directions.

2. Theoretical background

This study is related to two major streams of research – (i) behavioral information security of individuals, and (ii) criminology theories – where we focus on routine activity theory and its application to information security.

2.1. Information security behavior of individuals

Previous work on security behavior for individuals has identified factors that distinguish home security from security behavior in the workplace. First, home users do not usually receive security training and do not have access to professional IT support for solving problems at home; instead, they tend to rely on their own experience to secure their PCs (Anderson and Agarwal, 2010). Second, although people may receive security training at work, there is no clear evidence that this may result in increased home security (Talib et al., 2010). Likewise, protection procedures are not mandatory at home and security is perceived as expensive (Furnell et al., 2007). Most homes do not administer compliance toward pre-defined rules and it is difficult to impose sanctions for poor behavior (Li and Siponen, 2011). As a result, the relevant studies on the topic of computer abuse (Willison and Warkentin, 2013) are less applicable in the home context.

Drawing on Protection Motivation Theory (PMT) and related health-belief models, previous research has shown that individuals' perceptions of risks have important implications for information security, as actions of individual users can compromise entire systems (Rogers, 1975). PMT posits that when an individual is confronted with a threat, he or she cognitively assesses the threat and a possible remedy. One of the seminal works in this area is that of Anderson and Agarwal (2010), who extended PMT to show that home computer user's intentions are formed by a combination of cognitive, social, and psychological components, and that these intentions can be enhanced through self-view and goal-frame message manipulations.

Liang and Xue (2010) used the Technology Threat Avoidance Theory (TTAT), which draws on PMT as its theoretical base, to explain how perceived threat severity and susceptibility contribute to the avoidance of spyware threats in personal computer usage. TTAT describes the process and the factors that influence IT users' threat avoidance behavior (Liang and Xue, 2009). According to TTAT, when users perceive an IT threat, they are motivated to actively avoid the threat by taking safeguarding measures, if they perceive that such methods can help avoid the threat, and they may also passively avoid the threat by performing emotion-focused coping.

Previous research on information security behavior has focused on protecting and mitigating threats to the information assets of individuals in organizational contexts, concentrating mainly on technical issues (Crossler et al., 2013). In addition, most studies are based on self-reported data gathered from surveys, with students as the sample in the majority of cases. Our study overcomes previous limitations by focusing on information security behavior of individuals in domestic contexts – a relevant topic given the increase of home-based activities in the post-pandemic world – and by using data collected by the Spanish government about computer use of their citizens, which include different segments of the population across the country. We also propose a complementary approach that combines self-reported data with actual data collected directly from home users' computers, which allows us to examine the relationship between risk perception, security behavior and actual security incidents.

2.2. Routine activity theory and information security

Criminology theories provide useful theoretical perspective to understand why certain individuals or groups are more likely to become involved in delinquency or crime. RAT postulates how changes in routine activities of society's members impact the levels of direct-contact predatory crimes; that is, crimes where one or more persons directly take or damage the person or property of another. Cohen and Felson (1979) introduced RAT as an ecological perspective on criminal behavior, where the behavioral patterns and intersections of people in time and space influence when and where crimes occur. Instead of focusing on the effect of punishment, as it is the case of Deterrence Theory, RAT argues that the features of environmental settings impact criminal activities.

RAT assumes that three main components predict the likelihood of a crime: (i) a motivated offender, (ii) a suitable target, and (iii) the absence of capable guardianship. When a motivated offender encounters a suitable target in the absence of a capable guardian, crimes may occur. For instance, people may leave their house during the day for a patterned period of time because they need to go to work. A potential offender may perceive this daily recurrent activity as an opportunity to enter an unprotected empty house to steal valuables from the house.

Thus, the criminal risk for an individual is influenced by individuals' routine activities that bring them or their properties into direct contact with potential offenders.

RAT has been mainly applied to explain crime victimization in general, and cybercrime victimization in particular (Yar, 2005; Reyns, 2013; Leukfeldt, 2014; Pyrooz et al., 2015), usually in combination with lifestyle exposure theory (Hindelang et al., 1978).

In criminology, several studies have applied RAT to examine malware infection victimization, with no conclusive results. The earliest work was done by Choi (2008), who studied virus infection, establishing that performing risky online activities increases the odds of virus victimization, and that capable guardianship in the form of anti-virus software decreases virus victimization. By contrast, Bossler and Holt (2009) found that guardianship is not correlated with data loss from malware victimization. Leukfeldt and Yar (2016) found that online activities are not a factor in malware victimization, but that being on-line more frequently increases the chance of a malware infection. In the context of information systems, RAT has been applied to study system risks from a criminology perspective (Willison and Backhouse, 2006), to analyze cyber threats (Choo, 2011), and to study insider attacks to IS applications (Wang et al., 2015).

The present study departs from previous RAT studies by focusing on the domestic setting, and by including risk perception and the history of previous security and fraud incidents as instrumental variables to analyze malware infection. By contextualizing RAT in the domestic setting, integrating it with models of security and fraud management, and validating the resulting comprehensive security behavioral model with actual data from a country population, we hope to make a significant contribution to the understanding of the role of security tools and security management practices for domestic users.

3. Research framework and hypothesis development

In this section, we develop a research model to study how usage of security tools and security activities by Internet users influences their security incidents, as depicted in Fig. 1.

Specifically, drawing on the RAT and security behavior literature, we argue that usage of security tools and security activities both impact security incidents. However, the literature is not conclusive on which one has a higher impact, and we aim to clarify that point. We start by defining the specific type of security incident we examine in this paper: malware infection.

3.1. *Malware – malicious software*

The term *malware* refers to a variety of hostile, intrusive, and annoying software or program code designed to secretly access a computer system without the owner's informed consent (Souppaya and Scarfone, 2013). According to the Cybersecurity Predictions Report from DataProt (2022), "560,000 new pieces of malware are detected every day. There are now more than one billion malware programs out there. Every minute, four companies fall victim to ransomware attacks. Trojans account for 58% of all computer malware." Such constant growth in malware is becoming a significant problem for governments, businesses, and citizens around the world.

Malware is no longer simply used to damage, break, or intrude on computer systems; it now exists as a tool used by criminals to make a profit, making a shadow Internet economy that is worth over US$445

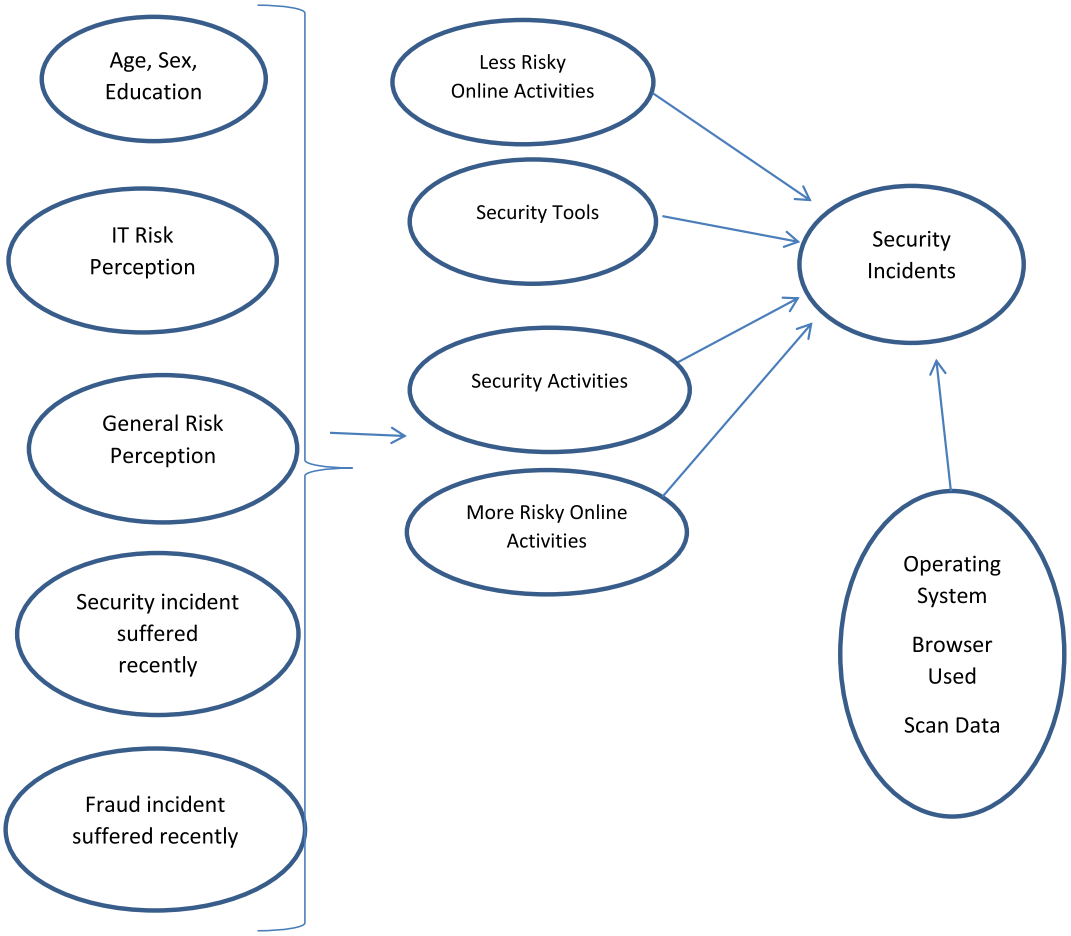

Fig. 1. Conceptual model.

billion, according to an estimate from the World Economic Forum's Global Risk Report (2016). Malware development and distribution is highly organized and controlled by criminal groups that have formalized and implemented business models to automate cybercrime.

The rapid development and popularity of the Internet has led to malware becoming increasingly complex, occurring in such forms as viruses, worms, Trojan horses, spyware, and ransomware. The earliest and best-known type of malware is a computer virus, a piece of code with malicious objectives that, when executed, replicates itself by modifying other computer programs and inserting its own code. A worm includes malicious code that propagate itself via LANs or the Internet, penetrating remote machines. A virus requires user intervention to spread, whereas a worm spreads automatically. A Trojan horse (or Trojan) is a type of malware that is often disguised as legitimate software; it appears to perform a certain action, but actually performs another action similar to a computer virus. A "rootkit" is a program (or combination of several programs) designed to take fundamental control of a computer system, without authorization by the system's owners or legitimate managers. Often, rootkits are also Trojans, which fools users into believing that they are safe to run on their systems. A backdoor is a method of bypassing normal authentication procedures. Once a system has been compromised (by any of the previous types of malwares), one or more backdoors may be installed to allow easier access in the future. Spyware is any software installed on the system without the owner's knowledge. Spyware collects information and sends it back to the attacker so the attacker can use the stolen information in a nefarious manner. Lockers, ransom malware, and ransomware are types of malware that prevent users from accessing their system or personal files and demand ransom payment in order to regain access (Siponen and Oinas-Kukkonen, 2007).

The relationships between these different types of malware can be complex. For example, a Trojan horse may be used to deliver and install ransomware on a victim's system; worms can help propagate viruses to other computers within a network; and spyware may be used in conjunction with a botnet to steal valuable information from infected machines (Aslan and Samet, 2020). In addition, the impact of malware on the security behavior of individuals can be varied. Malware attacks can lead to various negative consequences, which may influence how individuals perceive and approach security measures.

3.2. Online activities and malware infection

Cyberspace creates an environment in which computer criminals like hackers look for opportunities to find suitable targets to attack and to get valuable information. Therefore, while cyberspace can be considered a risky place, it may not be appropriate to argue that everybody in cyberspace has the same level of risk. Some individuals are more likely to be target of a cyberattack because of their proclivity to engage in risky online activities. RAT postulates that people's daily activities create the opportunity for potential criminals to commit crimes, with an individual having a greater likelihood of victimization if they are located in a certain place at a certain time (Cohen and Felson, 1979).

The risk of victimization depends on different lifestyle of individuals, which is defined as "routine daily activities", in which routine activities comprise both vocational activities (work, school, house-keeping activities) and leisure activities. One of the early works extending the concept of routine activities to study computer crime was carried out by Choi (2008), who found that some online activities, such as visiting unknown websites and downloading illegal videos and games, create a higher risk of victimization than other online activities, such as checking email or visiting online news channels. Other researchers have applied RAT to cybercrime, including online activities such as instant messaging and participation in chat rooms (Ngo et al., 2011) and social networking sites (Leukfeldt, 2014). Following these lines of work, we selected a set of classic online activities (see Appendix I) and classified them into more and less risky. The riskier activities include sharing files using P2P networks, direct download from servers or cyber-lockers, and online games. These activities contrast with email communication, participation in online social networks, and *E*-commerce, which are considered less risky. In light of the above discussion, we hypothesize that:

H1. Use of home computers for more risky online activities increases malware infection.

H2. Use of home computers for less risky online activities does not affect malware infection.

3.3. Security tools and security activities as guardians against malware infection

Guardians play a significant role in RAT. Bossler and Holt (2009) defined guardianship as "the capability of persons and/or objects that prevent the motivated offender from injuring or attacking the target". In criminology, a person or property that is less guarded is considered as a suitable target by offenders because such person or property offer little or no resistance to crime. Therefore, individuals or properties that have greater guardianship are less likely to become victims of crime (Cohen and Felson, 1979). Previous research applying RAT to cybercrime has distinguished two forms of guardianship: physical guardianship, denoting computer software developed to help protect computer systems from computer criminals (such as antivirus or firewalls), and personal guardianship, denoting users' skill level with technology to protect their computer systems.

Information security management includes provision of security tools, as well as the establishment of practices and procedures to protect information (Werlinger et al., 2009). Following RAT, we consider security tools as physical guardians, and security practices, procedures, and activities as a type of personal guardianship. In contrast to physical guardianship, personal guardianship is about awareness of the risks in the online environment and the soft skills and capabilities of users to follow procedures and practices to safeguard against different risks. Security behavior is continuously impacted by the introduction of new technologies, which makes it necessary for organizations to adapt to the fast pace of growth. Consequently, it is necessary to take soft skills into account to increase dynamic capabilities (Ogunrinde, 2022; de Miguel et al., 2022).

Previous research has shown that security tools and security activities have both strengths and weaknesses. On one hand, security tools, such as antivirus software, can provide real-time protection against known malware threats. They can detect and block malicious files or activities before they can cause harm. In addition, security tools can automate the scanning process, helping to detect and remove malware that might be present on the system, making the whole process more efficient and less time-consuming (Rowlingson, 2011). However, while security tools excel at dealing with known malware, they may struggle to detect and block new and previously unseen threats (Aslan and Samet, 2020). Cybercriminals constantly create new malware variants to evade detection. Furthermore, some security tools can be resource-intensive, leading to reduced system performance during scans or real-time protection. This can be particularly problematic for older or less powerful devices, which is often the case with computers used at home. Finally, the effectiveness of security tools relies heavily on regular updates, as failure to update the software promptly can leave the system vulnerable to the latest threats.

On the other hand, in the domestic setting, security activities usually involve practices such as setting passwords, checking security settings of the operating system and browsers, or backing up important files (Rowlingson, 2011). Security activities have certain strengths. Unlike security tools, which may struggle to detect new or previously unknown malware, security activities can be more adaptable to emerging threats (Peltier, 2016). Regular security assessments and updates to practices allow for a quicker response to new risks. Moreover, while security tools typically involve ongoing licensing fees, security activities can often be

implemented at a lower or no cost, making them an attractive option for domestic use. Nevertheless, implementing security activities in the domestic setting does have certain limitations. Even with the best security activities in place, human error can still lead to security breaches. A single mistake, such as clicking on a malicious link or falling for a phishing email, can bypass all preventive measures. Additionally, some security activities can be complex to implement to users without technical knowledge and require continuous monitoring and regular updates to stay effective. Prior research has also shown that security awareness and training programs are good practices to protect information (D'Arcy et al., 2009; Yoon et al., 2020). Thus, security activities can improve outcomes in the organizational setting where they can be monitored and sanctioned. However, it is unclear whether security activities can have a similar effect in the home setting in the absence of monitoring and sanctioning.

However, despite the conflicting arguments and findings from prior research in relation to security tools and security activities, we hypothesize that physical guardianship in the form of security tools and personal guardianship in the form of security activities are likely to reduce the likelihood of malware infection on home computers.

H3. Use of security tools in home computers is associated with a decrease in malware infection.

H4. Use of security activities in home computers is associated with a decrease in malware infection.

The main limitation of security tools is that they only protect against known malware, whereas security activities are difficult to monitor and sanction in the home setting. Thus, the question of whether security tools or security activities will be more effective at mitigating malware infection is ultimately an empirical one.

4. Research methodology

4.1. Methodology

This research is based on data from the *Cybersecurity and Confidence in Spanish Households* national study conducted by the National Observatory of Telecommunication and Information Society of Spain in June 2014. The study was designed to measure the cybersecurity behavior of Spanish households. Two types of data were collected in the study: self-reported data and scanned data. Self-reported data correspond to the responses that participants gave to an online questionnaire. The online questionnaire was piloted with 97 respondents. Scanned data refer to data obtained from scanning users' personal computers upon approval of participants. A specialized software called *iScan* was installed remotely in participants' equipment to search for resident malware and to collect data about the equipment's operating system, its up-to-date status, the security tools installed, and the security activities of the user. The scanning software was installed and the scanned data were collected after the participant had completed the online questionnaire, but scanning can be performed up to four weeks after the participant completes the questionnaire. This period is usually much shorter and although reminders are sent to expedite the survey and scanning process, these are done separately at the convenience of the user. Before the scanning takes place, a survey quality control process was carried out to avoid random responses from users. The *iScan* software analyzed the presence of malware using 46 antivirus engines and uses an online malware and virus-scanner aggregator as the base of its analysis. The process of installing the software and running the scan took approximately 1 h.

For data analysis, we used different multivariate linear regression and multivariate logistic binary regression (logit and probit). T4hese models provide researchers with more information on the directionality and size of an effect than the standard statistical techniques such as ANOVA. The models can deal with imbalanced data, which frees researchers from overly restrictive designs that affect the naturalness of the object of their study (Jaeger, 2008). All statistical calculations were performed with Stata version 16.

4.2. Participants

The participants in the study consisted of a sample of Internet users in Spain who were representative in terms of their geographic, demographic, and socio-economic distribution. A total of 3010 Internet users completed the online questionnaire; from this self-reported data, over 1900 individuals' computers were scanned and data from their computers were collected. The empirical analysis uses data from 1918 respondents who completed the online questionnaire and agreed to install the *iScan* software. The mean age of the respondents was approximately 43 years old, with a minimal age of 18 years and a maximal age of 75. Forty-seven percent of the respondents were females. In terms of educational background, 90 % had at least a high school degree and 30 % had a graduate degree. We aimed to have a demographically representative group of participants, since previous studies have shown that young users have a higher number of malware infections than older ones (Urueña et al., 2019). A market research company was employed to recruit participants; and as an incentive to encourage participation, individuals were given vouchers that could be used to buy items in several stores around the country.

The online activities, security tools installed, and the security activities of a user are likely to be endogenous. We used sex, age, education, security incidents suffered recently, fraud incidents suffered recently, general risk perception, and IT risk perception of the user as exogenous variables for the less and more risky online activities, security tools installed, and the security activities of the individual.

The basic model is shown in Fig. 1. The malware incidents on a computer are likely to be influenced by the different online activities, the security tools installed, and the security activities of the individual, controlling for the operating system, browser, and the risk faced by the computer.

4.3. Variables

4.3.1. Dependent variables

The software installed on the computer captured three measures of malware incidents suffered by the computer. The first measure is "Infected"; that is, whether the computer is infected with malware such as virus, worms, etc. The "Infected" variable indicated the presence of malware in respondents' computers (0 = no, 1 = yes). For a computer to be considered infected, it should have at least one malware that has been detected by at least seven anti-virus engines, corresponding to 15 % of the total of anti-virus engines in *iScan*. Malware must be spotted by a number of anti-virus engines to avoid the problem of false positives (false alarms), in which case the anti-virus thinks that a file is harmful when it is actually safe. By detecting a malware using several anti-virus engines, we decrease the likelihood of false positives. The second measure of malware is the "Risk Level" of the computer. The software installed measured the risk level on a an ordinal scale from 0 to 3, where a higher number meant a more risky/compromised computer. Anti-virus engines classify malware ranged from weak to very dangerous; for each engine, the classification of the malware was normalized from 1 (weak, low risk) to 3 (very dangerous, high risk). The risk level of a computer is the average of all the risk level of the malwares found by the 46 anti-virus engines. The third measure of malware is the total number of malware infections on the computer (total infections). Appendix I includes the list of all malware analyzed by the *iScan* tool and the risk level associated with each type of malware.

4.3.2. Independent variables

The key independent variables in the study are less and more risky online activities, security tools installed, and the security activities of the individual. The survey asks the individual to indicate (yes/no) what

online activities they perform on their computer (email, instant messaging, social network, e-commerce, online games, etc.); what security tools they have installed on their computer in the last three months (antivirus programs, firewall, spam blocker, pop-up blocker, etc.); and their security activities (block cookies, use digital certificates, create data partitions, updated the operating system, etc.). Online activities were divided into less risky and more risky online activities, following similar classification used in prior cybercrime studies (Yar, 2005; Choi, 2008). The measures for security activities were developed based on recommended security practices for home computer security (Wash, 2010), drawing from the CERT Home Computer Security and the US-CERT Cyber-Security guidelines. Please see Appendix II for the list of items for less risky and more risky online activities, security tools installed, and the security activities of the individual. Preliminary analyses showed that internal consistency for these scales was adequate (Cronbach's $\alpha \geq 0.75$). We assess less and more risky online activities, security tools installed, and security activities for each individual as the principal factor of their response to the items in the respective scales. Due to the binary nature of responses, a tetrachoric correlation matrix (rather than a Pearson correlation matrix) is used for identifying the principal factor (Knol and Berger, 1991).

4.3.3. Instrument variables

The less and more risky online activities, security tools installed, and the security activities of an individual are very likely to be endogenous and are likely to be influenced by the demographic characteristics like the sex, age, and education of the individual, instrumental variables that are usually used in cybercrime studies (Choi, 2008; Herrero et al., 2022). Some studies have also shown that young users have a higher number of infections than older ones (Urueña et al., 2019). IT and general risk perception of the individual; and the recent experiences like the security and fraud incidents recently suffered by the individual are also considered instrumental variables.

The data for the instrument variable were collected using the survey. Education is measured on an ordinal scale from 1 to 7, where 1 stands for do not know how to read and write and 7 stands for advanced graduate degree. Security incidents suffered recently, fraud incidents suffered recently, general risk perception, IT risk perception were assessed based on responses to multiple item scales.

The measurement of general risk perception was done using the short version of the Domain-Specific Risk-Taking (DOSPERT) scale (Blais and Weber, 2006). The DOSPERT questionnaire measures risk perceptions on financial, health/safety, recreational, ethical and social dimensions. The risk perception scale measures the degree to which the respondent considers the activities in these domains as risky. A seven-point rating scale was used, with 1 being not risky and 7 being extremely risky. The measurement of IT risk perception was based on the work of Herrero et al. (2017), measuring risk perception about the use of IT in several activities such as accessing bank account from a public network or adding unknown people in online social networks (see Appendix III for the items used for each scale). The specific assessment for security incidents suffered recently, fraud incidents suffered recently, general risk perception, and IT risk perception was based on calculation of principal factor of each scale. Analyses showed that internal consistency for these scales was adequate (Cronbach's $\alpha \geq 0.74$)

4.3.4. Control variables

We were interested in the impact of self-reported less and more risky online activities, security tools installed, and the security activities on malware incidents. The malware incidents were also likely to be influenced by the security tools installed as well as the security activities performed on the computer. We assessed and controlled for the risk faced by a computer based on the software collected data from computer. The *iScan* software collected information about the presence of security tools such as antivirus, firewall, anti-spyware, and fraud detection software in the individual's computers (scan security tools). The *iScan* software also captured the use of security activities performed on the computer, such as the use of digital certificates, latest update to the OS, and cleaning of the browsing history (scan security activities). We also controlled for the software collected autorun status of the computer (scan autorun); see Appendix IV for the assessment of these variables. We assessed the scanned security tools, scanned security activities, and scanned autorun status for each individual as the principal factor of data from the scan of the computer. Due to the binary nature of this data, a tetrachoric correlation matrix (rather than a Pearson correlation matrix) was used in identifying the principal factor (Knol and Berger, 1991). We also controlled for the operating system and the browser used by the computer.

Table 1 presents the summary statistics and correlations between the key variables. The three key dependent variables – infected, risk level, and total infections – are highly correlated with each other. Thus, the dependent variables are internally consistent. As expected, the riskier online activities were positively correlated with the dependent variables whereas less risky online activities are not correlated with the dependent variables. Security activities are negatively correlated with the dependent variables. Hausman tests also indicate that online activities, security tools used, and the security activities of individuals are endogenous.

5. Results

In order to check hypotheses H1–H4, we performed 15 different regression models distributed in four tables, measuring variables related to the computer being "infected" (Table 2), the risk level of the computer (Table 3), the total number of infections on the computer (Table 4), and the drivers of less and more risky online activities (Table 5). For each table, we will explain the different regression models used and the relationship with the stated hypotheses.

We first examined whether self-reported less and more risky online activities, security tools, and security activities were related to the computer being infected. Table 2 presents the analysis. Model 1 presents a logit model and Model 2 presents a probit model. These models do not consider any potential endogeneity in less and more risky online activities, security tools, and security activities. These models suggest that, consistent with Hypothesis 2, less risky online activities are unrelated with the computer being infected; consistent with Hypothesis 1, more risky online activities are positively related with the computer being infected ($p < 0.01$ level); contrary to Hypothesis 3, security tools are positively related with the computer being infected ($p < 0.05$ level); and consistent with Hypothesis 4, security activities are negatively related with the computer being infected ($p < 0.01$ level). Model 1 suggests that a unit increase in more risky online activities increases the log odds of the computer being infected by 0.707. Similarly, a unit increase in security tools increases the log odds of being infected by 0.319, whereas a unit increase in security activities reduces the log odds of being infected by 0.604. The results in the probit model (Model 2) are also quite consistent with the results in Model 1. Models 3 and 4 in Table 2 use instrument variable probit models to consider the endogeneity of less and more risky online activities, security tools, and security activities in examining their impact on the computer being infected. Model 3 suggests that, contrary to Hypothesis 3, security tools used are positively related with the computer being infected ($p < 0.01$ level); whereas, consistent with Hypothesis 4, security activities are negatively related with the computer being infected ($p < 0.01$ level). Model 4 uses a two-step estimation and has quite consistent results. This analysis suggests that security tools used by the user are positively related with an increase in probability of the computer being infected ($p < 0.1$ level). However, the security activities of the user reduce the likelihood of the computer being infected ($p < 0.1$ level).

Table 3 conducts the analysis with the risk level of the computer as the dependent variable. This analysis uses an ordered probit model as the risk level is a four-point scale ordinal variable, with 0 signifying lowest level of risk and 3 signifying the highest level of risk. Model 5

Table 1
Descriptive statistics and correlation among key variables.

	Mean	Std Dev	1	2	3	4	5	6	7	8	9	10	11	12	13
1	0.599	0.490	1.00												
2	1.324	1.319	0.82	1.00											
3	2.304	3.949	0.48	0.56	1.00										
4	0.906	0.284	0.00	0.00	−0.01	1.00									
5	0.334	0.317	0.09	0.10	0.08	0.41	1.00								
6	0.430	0.362	0.03	0.05	0.01	0.31	0.33	1.00							
7	0.582	0.325	−0.05	−0.04	−0.05	0.31	0.26	0.54	1.00						
8	0.170	0.200	−0.01	0.00	−0.02	−0.01	0.01	0.09	0.01	1.00					
9	0.259	0.171	−0.06	−0.05	−0.01	0.04	0.05	0.11	0.15	0.04	1.00				
10	0.000	0.991	−0.05	−0.03	−0.01	0.05	0.03	0.05	0.02	−0.03	0.04	1.00			
11	0.468	0.499	−0.03	−0.05	−0.03	−0.06	−0.14	−0.17	−0.16	−0.05	−0.08	−0.04	1.00		
12	42.565	11.189	−0.01	−0.01	0.02	−0.09	−0.18	0.06	0.10	0.02	0.07	0.01	−0.14	1.00	
13	5.724	1.044	−0.08	−0.08	−0.06	0.13	0.05	0.04	0.08	0.02	−0.02	0.02	0.05	−0.14	1.00

1 = Infected; 2 = Risk Level; 3 = Total Infections; 4 = Less Risky Online Activities; 5 = More Risky Online Activities; 6 = Security Tools; 7 = Security Activities; 8 = scan Security Tools; 9 = Scan Security Activities; 10 = Autorun; 11 = Sex; 12 = Age; 13 = Education.

Table 2
The computer is "infected".

	Logit	Probit	Instrument variable probit	Instrument variable probit
	Robust standard error	Robust standard error		Two-step estimation
	Model 1	Model 2	Model 3	Model 4
Less risky online activities	−0.212 (0.180)	−0.130 (0.112)	−1.024 (0.995)	−2.087* (1.195)
More risky online activities	0.707*** (0.159)	0.431*** (0.0969)	0.210 (0.763)	0.981 (0.909)
Security tools	0.319** (0.157)	0.195** (0.0962)	2.759*** (0.414)	3.248* (1.707)
Security activities	−0.604*** (0.172)	−0.370*** (0.106)	−2.547*** (0.571)	−2.940* (1.695)
Scan security activities	−0.430 (0.275)	−0.260 (0.169)	0.0271 (0.198)	−0.102 (0.286)
Scan security tools	−0.340 (0.235)	−0.213 (0.145)	−0.485*** (0.130)	−0.683** (0.293)
Autorun	−0.0177 (0.0504)	−0.0117 (0.0309)	−0.0500* (0.0285)	−0.0635 (0.0513)
Constant	0.706*** (0.227)	0.436*** (0.141)	1.285** (0.530)	2.130*** (0.744)
Operating system	Yes	Yes	Yes	Yes
Browser	Yes	Yes	Yes	Yes

*** p < 0.01, ** p < 0.05, * p < 0.1.

Table 3
The risk level of the computer.

Variables	Robust standard errors	Robust standard errors	Bootstrap standard errors
	Model 5	Model 6	Model 7
Less risky online activities	−0.172* (0.0991)	−0.151 (0.104)	−0.151 (0.111)
More risky online activities	0.426*** (0.0867)	0.429*** (0.0943)	0.429*** (0.0962)
Security tools	0.222*** (0.0858)	1.885 (1.154)	1.885 (1.239)
Security activities	−0.347*** (0.0950)	−1.226 (1.139)	−1.226 (1.229)
Scan security activities	−0.238 (0.154)	−0.249 (0.161)	−0.249 (0.160)
Scan security tools	−0.0842 (0.133)	−0.103 (0.138)	−0.103 (0.145)
Autorun	−0.00454 (0.0285)	−0.0181 (0.0306)	−0.0181 (0.0327)
Security tools residual		−1.646 (1.155)	−1.646 (1.240)
Security activities residual		0.852 (1.142)	0.852 (1.232)
Less risky online activities residual		−2.412*** (0.802)	−2.412*** (0.900)
More risky online activities residual		0.904* (0.531)	0.904* (0.544)
Constant cut1	−5.671*** (0.247)	−7.225*** (0.466)	−7.225*** (0.631)
Constant cut2	−5.063*** (0.259)	−6.615*** (0.482)	−6.615*** (0.628)
Constant cut3	−5.043*** (0.253)	−6.593*** (0.478)	−6.593*** (0.628)
Operating system	Yes	Yes	Yes
Browser	Yes	Yes	Yes

*** p < 0.01, ** p < 0.05, * p < 0.1.

presents analysis that does not consider the endogeneity of the less and more risky online activities, security tools used, and the security activities of the individual user. In this model, consistent with Hypothesis 1, more risky online activities are positively related with the risk level of the computer (p < 0.01 level); and consistent with Hypothesis 4, security activities are negatively related with the risk level (p < 0.01 level). However, contrary to Hypothesis 3, the security tools used are positively related with the risk level (p < 0.01 level).

In Models 6 and 7, we control for the endogeneity of the less and more risky online activities, security tools used, and the security activities of the user using the control function approach. This involves running separate linear regressions for less and more risky online activities, security tools used, and security activities of the user on the instrument variables described above in the first-stage regressions and then including the residuals in the second-stage ordered probit models. We label the residual of the first stage linear regressions of less risky online activities, more risky online activities, security tools used, and security activities as less risky online activities residual, more risky online activities residual, security tools residual, and security activities residual, respectively. In Model 6, consistent with Hypothesis 1, more risky online activities are positively related with higher risk ($p < 0.01$); and consistent with Hypothesis 2, less risky online activities are unrelated with the risk level. In Model 6, security tools are positively related with higher risk and security activities are negatively related with higher risk. However, the coefficients are not statistically significant. In this model, the coefficients of less risky online activities' residual and more risky online activities residual are significant, suggesting that endogeneity is a concern. Thus, in Model 7 we present the models with bootstrapped standard errors. Bootstrap methods can be applied in many statistical analyses, often providing answers of equal or higher quality than alternative methods, even though the price to be paid is computer time (see Wehrens et al., 2000 for discussion). The results in

Table 4
The total number of infections on the computer.

Variables	Poisson model	NBREG model	Poisson model	NBREG model	Poisson model	GMM estimation
	Robust standard error	Robust standard errors	Robust standard errors	Robust standard errors	Bootstrap standard errors	Robust standard errors
	Model 8	Model 9	Model 10	Model 11	Model 12	Model 13
Less risky online activities	−0.221	−0.211	−0.237	−0.221	−0.237	−1.531
	(0.147)	(0.143)	(0.153)	(0.149)	(0.170)	(1.799)
More risky online activities	0.511***	0.546***	0.548***	0.602***	0.548***	0.0668
	(0.119)	(0.128)	(0.131)	(0.140)	(0.129)	(1.001)
Security tools	0.136	0.153	3.741**	4.967***	3.741**	2.951**
	(0.130)	(0.130)	(1.717)	(1.670)	(1.665)	(1.244)
Security activities	−0.413***	−0.403***	−3.417**	−4.450***	−3.417**	−2.673***
	(0.138)	(0.139)	(1.684)	(1.673)	(1.637)	(0.838)
Scan security tools	−0.289	−0.260	−0.250	−0.238	−0.250	−0.737
	(0.205)	(0.195)	(0.214)	(0.204)	(0.219)	(0.463)
Scan security activities	0.123	0.0556	0.107	0.0665	0.107	0.341
	(0.231)	(0.222)	(0.254)	(0.240)	(0.249)	(0.382)
Autorun	0.0223	0.0255	0.0322	0.0450	0.0322	−0.00522
	(0.0456)	(0.0432)	(0.0506)	(0.0487)	(0.0541)	(0.0776)
Less risky online activities residual			−1.540	−1.931*	−1.540	
			(1.061)	(1.111)	(1.018)	
More risky online activities residual			−0.193	−0.385	−0.193	
			(0.774)	(0.751)	(0.741)	
Security tools residual			−3.594**	−4.798***	−3.594**	
			(1.720)	(1.677)	(1.679)	
Security activities residual			2.947*	3.967**	2.947*	
			(1.706)	(1.685)	(1.662)	
Constant	1.023***	1.028***	2.654***	3.153***	2.654***	2.275*
	(0.207)	(0.192)	(0.659)	(0.724)	(0.639)	(1.214)
Operating system			Yes	Yes	Yes	Yes
Browser			Yes	Yes	Yes	Yes

*** p < 0.01, ** p < 0.05, * p < 0.1.

Table 5
The drivers of less and more risky online activities.

Variables	Less risky online activities	More risky online activities
	Robust standard error	Robust standard error
	Model 14	Model 15
Security tools	−0.0345***	0.0776***
	(0.00799)	(0.0191)
Security activities	−0.00382	0.0193
	(0.00855)	(0.0204)
Online activities	0.895***	0.718***
	(0.00782)	(0.0179)
Education	0.00689***	−0.0112**
	(0.00216)	(0.00513)
Sex	0.0147***	−0.0388***
	(0.00483)	(0.0115)
Age	0.00113***	−0.00254***
	(0.000219)	(0.000512)
IT risk propensity	−0.0119***	0.0259***
	(0.00300)	(0.00712)
General risk propensity	−0.00196	0.00232
	(0.00275)	(0.00652)
Security incidents suffered recently	−0.0171	0.0332
	(0.0177)	(0.0420)
Fraud suffered recently	0.0187**	−0.0476***
	(0.00757)	(0.0181)
Scan security activities	−0.00213	0.00344
	(0.0134)	(0.0316)
Scan security tools	−0.00414	0.0161
	(0.0121)	(0.0285)
Autorun	0.000916	−0.000335
	(0.00264)	(0.00644)
Constant	0.0552**	−0.0882
	(0.0247)	(0.0595)
R-squared	0.880	0.527
Operating system	Yes	Yes
Browser	Yes	Yes

*** p < 0.01, ** p < 0.05, * p < 0.1.

Model 7 are very consistent with the results in Model 6. Overall, the analysis in Table 3 indicates that more risky online activities drive the risk level of the computer. Also, though security tools used are associated with increase in risk level and the security activities are associated with a decrease in the risky level, in the models that consider the endogeneity of less and more risky online activities, security tools and security activities, the coefficients are not statistically significant. This suggests that it is the riskier online activities of the user that drive the risk level of the computer.

Table 4 performs the analysis with the total number of malware infections as the dependent variable. Given that the dependent variable is a count variable, we use count data models: Poisson and Negative Binomial models. Models 8 and 9 do not consider the endogeneity of the less and more risky online activities, security tools and security activities. Consistent with Hypothesis 1, Models 8 and 9 suggest that more risky online activities are positively associated with the number of malware infections ($p < 0.01$ level); and consistent with Hypothesis 2, less risky online activities are unrelated to the number of malware infections. However, inconsistent with Hypothesis 3, Models 8 and 9 suggest that security tools are unrelated with the number of malware infections, whereas, consistent with Hypothesis 4, security activities are negatively related with the number of malware infections ($p < 0.01$ level). In Models 10 and 11 the regressions consider the endogeneity of less and more risky online activities, security tools, and security activities by including the residuals from the first stage models as in Table 3. Model 10 presents the Poisson model and Model 11 presents the negative binomial regression. Consistent with Hypothesis 1, the results in Models 10 and 11 indicate that more risky online activities are positively related with the total number of infections ($p < 0.01$ level). Consistent with hypothesis 2, less risky online activities are unrelated to the number of malware infections. However, contrary to Hypothesis 3, security tools are positively related with the number of malware infections. Consistent with Hypothesis 4, security activities are negatively ($p < 0.01$ level) related with the number of malware infections. Since the residual terms security tools residual, and security activities residual are

significant in Models 10 and 11, Model 12 presents the analysis with bootstrapped standard errors. Model 12 also produces very consistent results. Finally, Model 13 presents a general method of moments analysis to control for endogeneity, which also produces very consistent results. The analysis indicates that security tools are positively related with the number of malware infections ($p < 0.05$ level). However, the security activities are negatively related with the number of malware infections ($p < 0.01$ level).

In the majority of the above analysis, security tools are positively related and security activities are negatively related with the computer being infected, risk level, and the number of malware infections. These models control for the actual security tools, security activities, autorun state, and the browser and operating system of the computer using scanned data from these computers. This raises the question of why self-reported security tools are positively related with the computer being infected, the risk level of the computer, and the number of malware infections, and why self-reported security activities of the user negatively are related to the computer being infected, the risk level of the computer, and the number of malware infections.

One plausible explanation is that users who believe that they have security tools to protect their computer from malware engage in more risky online activities. Thus, we found a positive relationship between security tools and the computer being infected, the risk level of the computer, and the number of malware infections. However, other users who perform security activities to protect their computer from malware engage in less risky online activities; thus, we find a negative relationship between security activities and the computer being infected, the risk level of the computer, and the number of malware infections.

To further explore these propositions, we examine the relationship between the security tools and security activities of the user and the riskiness of the online activities of the user while controlling for all the online activities of the user, and the sex, education, age, IT and general risk perception, the security incident and fraud incidents recently suffered by the user, scan of security tools, scan of security activities, autorun state of the computer, and the browser and the operating system of the computer. Table 5 presents the analysis. The analysis suggests that security activities are unrelated to the less or more risky online activities. However, while security tools are negatively related with less risky online activities ($p < 0.01$ level), security tools are positively related with more risky online activities ($p < 0.01$ level).

6. Discussion

This study extended and empirically validated RAT to understand malware infection. Consistent with RAT, the results indicate that online lifestyle is related with malware infection. The study also adds a new understanding about the role of guardianship, contrasting physical and personal guardianship in the case of malware infection.

The study reveals that more risky online activities are positively related to malware infection, and less risky online activities are unrelated to malware infection. The relation between risky online activities and malware infection is present in other RAT studies. For instance, Choi (2008) showed that risky online leisure activities increase the odd of virus infections, and Bossler and Holt (2009) showed that watching online pornography increases the risk of malware infection. However, results of previous studies about less risky online activities and malware infections are not conclusive. Some studies have shown that usual activities, such as on-line shopping and visiting forums and social network sites, increase the odds of on-line threat victimization (Van Wilsem, 2011), while other studies have shown that most routine activities on the computer are not correlated with data loss from malware infection (Bossler and Holt, 2009). In looking for an explanation for the lack of relation between less risk online activities and malware infection, we noted that engaging in less risky online activities, such as visiting trusted websites, avoiding suspicious links or downloads, and being cautious with email attachments, reduces the likelihood of encountering malware-laden content (Souppaya and Scarfone, 2013). This cautious behavior makes individuals less attractive as potential targets for cybercriminals, as their efforts to exploit vulnerabilities are less likely to succeed. In addition, individuals who engage in less risky online activities are more likely to implement and maintain security measures that act as capable guardianship (Anderson and Agarwal, 2010), making it more difficult for cybercriminals to infiltrate their systems.

As discussed in the literature review, the findings of studies examining RAT in the context of online activities and malware infection are not conclusive, and some of the reasons for such uncertainty may be related to the type of sample, since prior studies have primarily used students as their sample. We use a representative sample of the Spanish population as our main sample. Our results are consistent with those of Leukfeldt and Yar (2016), which is the only other study applying RAT to malware infection that employs a similar sample (a representative sample of Dutch population); however, their study is based on self-reported data in the form of survey and secondary data.

Our study reveals that security tools (physical guardianship) are positively related with the computer being infected, higher risk level, and having more infections. Existing literature emphasizes that investing in technology is not always the solution; there should be an emphasis on individuals to improve their security behaviors (Balapour et al., 2020). Jampen et al. (2020) found that, despite the various advanced capabilities that tools currently available in the anti-phishing domain offer, such tools only support a limited subset of the potential factors identified as necessary to yield the desired training effects and greater awareness of phishing techniques and means of addressing them increases overall security. A plausible explanation for this behavior may come from moral hazard (Rowell and Connelly, 2012), since it is commonly noted that "protection from harm induces reckless behavior" (Romanosky, 2013). It is likely that security tools give users a false sense of security – "illusion of control" (Kellner et al., 2019), encouraging users to engage in more risky online activities that lead to security incidents. This finding makes an important contribution to our understanding of the role of security tools in information security. The literature has documented that, in general, security tools are not enough to protect information systems (Peltier, 2016). However, to the best of our knowledge, the existing literature is short on evidence regarding the negative consequences of the use of security tools.

Our study reveals that security activities (personal guardianship) reduce malware infection. Previous work using RAT has found that personal guardianship plays a role in cybercrime prevention, considering it as a primary form of defense (Grabosky, 2001; Bossler and Holt, 2009). More generally, the findings of the present study suggest that criminology theories can play a role in developing frameworks to understand and prevent information security incidents; in our case, malware infection. Programs to decrease malware infection may require a combination of several approaches including information technology, behavioral security, and criminology. In addition, by contrasting security tools and security activities, our study reveals the significance of security activities over security tools in preventing malware infections.

This study also contributes to the discussion about self-reported versus actual data in information security research. The rapid progress and widespread use of Internet technology have promoted the generation and accumulation of large-scale log data (Miao and Li, 2022). Previous studies have suggested that, in an information security context, it is preferable to measure actual behaviors rather than intentions (Anderson and Agarwal, 2010; Mahmood et al., 2010; Warkentin et al., 2011, 2012) because intentions do not always lead to behaviors (Crossler et al., 2013; Herrero et al., 2023). In the present study, we collected data in both ways, via an online survey and via data captured via the *iScan* monitoring software. The study uses the so-called black-hat data, such as malware infection, proper updates to security tools, as well as operating system and browsers.

The main theoretical contribution of this work is to present and validate a model rooted in RAT to study the relation between online

activities, security tools, security activities, and malware infection. We used demographic information as instrumental variables, as well as risk perception, and previous security and fraud incidents. There are several theoretical implications in the use of such a model. First, in relation to the counterintuitive result that more security tools are related to malware infection, we used concepts of moral hazard as a theoretical explanation of such a result. It would be interesting to further explore that via experiments, enriching our model with other measures of risk such as risk propensity. Second, the independent variables security tools and security activities were analyzed. The literature has shown that some security tools may be more effective than others, and that some security activities could be more effective (Rowlingson, 2011). It would be interesting to discriminate the different types of security tools and activities to analyze their individual relation with online activities and malware infection.

Our study has implications for public policy makers and organizations in general. First, the results illustrate the consequences of online activities and the importance of personal guardianship in the form of security activities to protect domestic users from malware infection. Policy makers should create policies and campaigns to make domestic users aware of the possible risks and consequences of their online behavior, as well as of the preventive measures that can be taken to decrease those risks. Second, our results reinforce the view that security tools are not sufficient to protect users from malware infection since security tools can only protect against known malware. Organizations should include security education and awareness programs for their main stakeholders – employees and customers – promoting security activities as a complement to the use of security tools.

7. Conclusion

The battle to secure cyberspace has transcended technological artifacts, with importance given to the behavioral aspect of cybersecurity. With growing use of home computers for teleworking, vocational and leisure activities, the security of home computers is becoming increasingly important. This study investigates the impact of security tools and security activities on malware infection for the case of domestic users. Using RAT as our theoretical lens, we found that risky online activities are associated with malware infection, whereas less risky online activities are not associated with malware. These findings confirm the intuition that it is the activities of the user that influence the security outcomes of home computer users. Interestingly, we found that security tools are positively related with malware infections, whereas security activities have a negative impact on malware infections. Collectively, our results underline the relevance of security activities for protecting home computers from malware infection and show that over-reliance on security tools could have negative consequences in relation to malware infection.

Two main strengths distinguish our work from previous work applying RAT to information security. First, the participants belong to a nationally representative sample of Internet users, which is not common in this field, where convenience samples are the rule rather than the exception. The large dataset used in the analyses should also add consistency to the study's findings and help the potential generalization of the results. Second, we complemented self-reported data, in the form of an online survey, with actual data, in the form of scanned data, which adds generalizability to the study findings.

Future research can take this study further by addressing several limitations of our study. First, we have focused on personal computers and there is an increase in the use of mobile devices around the world. Future work can extend this research to cover both PCs and mobile devices. In the present study, we only measured whether a computer is infected by malware; it would be interesting to also examine the drivers and consequences of the malware infection, such as for data loss. Second, the data used for the study was collected in 2014. The nature of online security risks and tools available to combat them evolve rapidly. Because data from 2014 may not accurately reflect the current situation, we call for new studies. Third, our sample was from one country, Spain. While the sample is representative of the Spanish society in terms of demographic, socio-economic and geographical distribution, future research can take this investigation further by drawing research subjects from different countries. Fourth, online activities, security tools installed, and security activities were reduced to binary variables. However, these variables may be more complex, with degrees of usage and effectiveness that may not be captured in a binary format. Finally, we focused on two types of guardianship – physical and personal – but other works on RAT include the so-called social guardianship, which refers to the availability of others who may prevent personal crimes by their mere presence or by offering assistance to ward off an attack. It would be interesting to analyze how interaction with others, either face to face interaction or digital interaction via social networks, may act as a form of guardianship to protect users from malware infection.

CRediT authorship contribution statement

Álvaro Arenas: Investigation, Conceptualization, Writing – original draft, Writing – review & editing. **Gautam Ray:** Methodology, Validation, Data curation, Writing – original draft. **Antonio Hidalgo:** Methodology, Validation, Writing – review & editing. **Alberto Urueña:** Supervision, Conceptualization, Data curation, Methodology, Validation, Writing – review & editing.

Declaration of competing interest

The authors declare no conflict of interest.

Data availability

Data will be made available on request.

Acknowledgements

Alberto Urueña thanks Foundation for the Scientific and Technical Research (FECYT, SV-PA-21-AYUD/2021/51411) for its contribution in supporting this research.

Appendix I. Dependent variables

List of malware detected by the *iScan* software and threat level.

High threat - risk level 3
Trojans
Keyloggers, keystrokes
Dialers
Virus

(*continued on next page*)

(*continued*)

Worm or worm
Rootkits
Exploits
Lockers, ransomware
Medium threat - risk level 2
Adware
Spyware
Scripts
Suspicious files detected heuristically
Low Threat - Risk Level 1
Intrusion tools
Jokes

Appendix II. Independent variables

Online Activities: From the following list of services offered by Internet, say if you've used them in the last three months.

Less Risky Online Activities	
Email (ex: Hotmail, Gmail, Yahoo…)	Choi (2008), Bossler and Holt (2009)
Social networks (Facebook, Twitter, LinkedIn…)	Ngo et al. (2011), Leukfeldt (2014)
E-commerce (buying and selling over the Internet) or auctions (eg: eBay)	Choi (2008), Bossler and Holt (2009)
Online banking (balance inquiry, transfers…)	Bossler and Holt (2009)
Watch Videos or listen to music online through Web pages (e.g.,: YouTube, Last.fm) or programs (e.g.: Spotify, Filmin)	
Online or video telephony (i.e.: Skype, Google Talk, ooVoo)	Van Wilsem (2011)
Forums, blogs and collaborative 2.0 platforms	Leukfeldt (2014)
Courses and online training	
E-Administration: information, procedures and formalities with Government over the Internet	
More Risky Online Activities	
Download files (Programs, videos, music) through P2P networks (eg: e, BitTorrent)	Choi (2008)
Direct download from servers or cyberlockers (e.g., Rapidshare, Uploaded, Fileserve)	Choi (2008)
Games online (multiplayer or games that need connection)	Choi (2008), Bossler and Holt (2009)
Online casinos betting services (e.g., Bwin, Unibet, Sportingbet)	

Security Tools: Thinking about the computer with which you usually access Internet at home, indicate if it has used the following in the past three months.

(a) Antivirus programs
(b) Firewall
(c) Programs or settings to block anti-spam/delete unwanted emails
(d) Anti-spy or anti-fraud programs
(e) Programs or browser settings to block pop-ups (pop-ups) and/or advertising
(f) Operating system updates
(g) Plug-ins, plugins, or Internet browser security extensions (e.g.: to block Java Script)
(h) Virus updates
(i) Programs, configurations or add-ins/extensions to block online advertising (eg: AdBlock, AdvertBan)

Security Activities: Thinking about the computer with which you usually access Internet at home, indicate if you've used or implemented the following in the last three months.

(a) I have password (access to equipment and documents).
(b) I make backup copies of important files (backup).
(c) I remove or block cookies or temporary files (manually or automatically).
(d) I use my computer with a user profile permission reduced (rather than an administrator profile).
(e) I use the electronic ID card to sign or electronically signed operations online on the Internet.
(f) I use digital certificates for identification or electronic signature (FNMT/CERES, etc.) with the exception of electronic ID.
(g) I use documents or data encryption tools.
(h) I have a partition, apart from the operating system, only for data in the hard disk.

Appendix III. Instrument variables

Security Incidents Suffered Recently: In the last three months had any of the following problems of security occurred with the computer with which you access Internet at home.

(a) Computer viruses or other codes malicious/malware (Trojans, worms…)
(b) I have lost or deleted data or files.

(c) I have received e-mails requested/unwanted (spam)
(d) Anyone has come on my email or on my social networking profile and has been impersonating me.
(e) Someone has accessed without my consent to my computer and/or device (intrusion)
(f) I have not been able to access service online because of cyber-attacks (fallen pages, etc.)

Fraud Incidents Suffered Recently: In the last three months had any of the following problems of security occurred on the computer with which you access Internet at home?

(a) They asked me for my personal information or user keys.
(b) They have offered me a service or a product that I have not requested.
(c) They've asked me to click on a suspicious link.
(d) I received a suspected of being fraudulent job offer.
(e) They have offered me products from pages of e-commerce that made me suspect that they are false.
(f) I have accessed web pages trying to impersonate banks, commerce or public administrations.

General Risk and IT Risk Perception: 30 questions from the Domain-Specific Risk Taking Scale (DOSPERT) (Blais and Weber, 2006), and 13 questions for IT Risk Perception (Herrero et al., 2017), ranging from 1 = not risky to 7 = extremely risky.

Appendix IV. Control variables

Security Tools, Security Activities, and Autorun State from iScan: Software capture settings on the following dimensions.

Variable	Items
Scan security tools	State and use of Firewall
	State and use of Anti-virus
	State and use of Parental control
	State and use of Anti-spy software
	State and use of Anti-fraud software
Scan security activities	State of system updates
	State and use of partitions
	Use of cleaning history
	Use of personal certificate – electronic ID
Scan autorun state	External units
	Removable units
	Fixed units
	Network units
	RAM units

References

Anderson, C.L., Agarwal, R., 2010. Practicing safe computing: a multimedia empirical examination of home computer user security behavioral intentions. MIS Q. 34 (3), 613–643. https://doi.org/10.2307/25750694.

Aslan, Ö.A., Samet, R., 2020. A comprehensive review on malware detection approaches. IEEE Access 8, 6249–6271. https://doi.org/10.1109/ACCESS.2019.2963724.

Balapour, A., Nikkhah, H.R., Sabherwal, R., 2020. Mobile application security: role of perceived privacy as the predictor of security perceptions. Int. J. Inf. Manag. 52, 102063 https://doi.org/10.1016/j.ijinfomgt.2019.102063.

Blais, A.-R., Weber, E.U., 2006. A domain-specific risk-taking (DOSPERT) scale for adult populations. Judgm. Decis. Mak. 1, 33–47 (Available at SSRN). https://ssrn.com/abstract=1301089.

Bossler, A.M., Holt, T.J., 2009. On-line activities, guardianship, and malware infection: an examination of routine activities theory. Int. J. Cyber Criminol. 3 (1), 400 (Available at). https://digitalcommons.georgiasouthern.edu/crimjust-criminology-facpubs/259.

Choi, K.S., 2008. Computer crime victimization and integrated theory: an empirical assessment. Int. J. Cyber Criminol. 2 (1), 308 (Available at).

Choo, K.K.R., 2011. The cyber threat landscape: challenges and future research directions. Comput. Secur. 30 (8), 719–731. https://doi.org/10.1016/j.cose.2011.08.004.

Cohen, L.E., Felson, M., 1979. Social change and crime rate trends: a routine activity approach. Am. Sociol. Rev. 588-608 https://doi.org/10.2307/2094589.

CompTIA, 2022. State of cybersecurity 2022 (Available at). https://www.comptia.org/content/research/cybersecurity-trends-research.

Crossler, R.E., Johnston, A.C., Lowry, P.B., Hu, Q., Warkentin, M., Baskerville, R., 2013. Future directions for behavioral information security research. Comput. Secur. 32, 90–101. https://doi.org/10.1016/j.cose.2012.09.010.

D'Arcy, J., Hovav, A., Galletta, D., 2009. User awareness of security countermeasures and its impact on information systems misuse: a deterrence approach. Inf. Syst. Res. 20 (1), 79–98. https://doi.org/10.1287/isre.1070.0160.

DataProt, 2022. Malware statistics 2022. Available at. https://dataprot.net/statistics/malware-statistics/.

de Miguel, P.M., Martínez, A.G., Montes-Botella, J.L., 2022. Review of the measurement of dynamic capabilities: a proposal of indicators for the automotive industry. ESIC Market 53 (1), e283. https://doi.org/10.7200/esicm.53.283.

Dissanayake, N., Jayatilaka, A., Zahedi, M., Babar, M.A., 2022. Software security patch management-a systematic literature review of challenges, approaches, tools and practices. Inf. Softw. Technol. 144, 106771 https://doi.org/10.1016/j.infsof.2021.106771.

Fleck, A., 2022. Cybercrime expected to skyrocket in coming years. Available at, Statista. https://www.statista.com/chart/28878/expected-cost-of-cybercrime-until-2027/.

Furnell, S.M., Bryant, P., Phippen, A.D., 2007. Assessing the security perceptions of personal internet users. Comput. Secur. 26 (5), 410–417. https://doi.org/10.1016/j.cose.2007. 03.001.

Grabosky, P.N., 2001. Virtual criminality: old wine in new bottles? Soc. Leg. Stud. 10 (2), 243–249. https://doi.org/10.1177/a017405.

Herrero, J., Urueña, A., Torres, A., Hidalgo, A., 2017. My computer is infected: the role of users' sensation seeking and domain-specific risk perceptions and risk attitudes on computer harm. J. Risk Res. 20 (11), 1466–1479.

Herrero, J., Torres, A., Vivas, P., Urueña, A., 2022. Smartphone addiction, social support, and cybercrime victimization: a discrete survival and growth mixture model. Psychosoc. Interv. 31 (1), 59. https://doi.org/10.5093/pi2022a3.

Herrero, J., Rodriguez, F.J., Urueña, A., 2023. Use of smartphone apps for mobile communication and social digital pressure: a longitudinal panel study. Technol. Forecast. Soc. Chang. 188, 122292 https://doi.org/10.1016/j.techfore.2022.122292 (doi:10.1080/13669877.2016.1153504).

Hindelang, M.J., Gottfredson, M.R., Garofalo, J., 1978. Victims of personal crime: An empirical foundation for a theory of personal victimization. Ballinger, Cambridge, MA.

Humayun, M., Niazi, M., Jhanjhi, N.Z., Alshayeb, M., Mahmood, S., 2020. Cyber security threats and vulnerabilities: a systematic mapping study. Arab. J. Sci. Eng. 45, 3171–3189. https://doi.org/10.1007/s13369-019-04319-2.

Jaeger, T.F., 2008. Categorical data analysis: away from ANOVAs (transformation or not) and towards logit mixed models. J. Mem. Lang. 59 (4), 434–446. https://doi.org/10.1016/j.jml.2007.11.007.
Jampen, D., Gür, G., Sutter, T., Tellenbach, B., 2020. Don't click: towards an effective anti-phishing training. A comparative literature review. HCIS 10 (1), 1–41. https://doi.org/10.1186/s13673-020-00237-7.
Kellner, A., Horlboge, M., Rieck, K., Wressnegger, C., 2019. False sense of security: a study on the effectivity of jailbreak detection in banking apps. In: 2019 IEEE European Symposium on Security and Privacy (EuroS&P). IEEE, pp. 1–14.
Knol, D.L., Berger, M.P., 1991. Empirical comparison between factor analysis and multidimensional item response models. Multivar. Behav. Res. 26 (3), 457–477. https://doi.org/10.1207/s15327906mbr2603_5.
Kormos, C., Gifford, R., 2014. The validity of self-report measures of proenvironmental behavior: a meta-analytic review. J. Environ. Psychol. 40, 359–371. https://doi.org/10.1016/j.jenvp.2014.09.003.
Lallie, H.S., Shepherd, L.A., Nurse, J.R., Erola, A., Epiphaniou, G., Maple, C., Bellekens, X., 2021. Cyber security in the age of COVID-19: a timeline and analysis of cyber-crime and cyber-attacks during the pandemic. Comput. Secur. 105, 102248 https://doi.org/10.1016/j. cose.2021.102248.
Leukfeldt, E.R., 2014. Phishing for suitable targets in the Netherlands: routine activity theory and phishing victimization. Cyberpsychol. Behav. Soc. Netw. 17 (8), 551–555. https://doi.org/10.1089/cyber.2014.0008.
Leukfeldt, E.R., Yar, M., 2016. Applying routine activity theory to cybercrime: a theoretical and empirical analysis. Deviant Behav. 37 (3), 263–280. https://doi.org/10.1080/01639625.2015.101 2409.
Li, Y., Siponen, M.T., 2011. A call for research on home users' information security behaviour. In: PACIS, p. 112.
Liang, H., Xue, Y., 2009. Avoidance of information technology threats: a theoretical perspective. MIS Q. 71–90. https://doi.org/10.2307/20650279.
Liang, H., Xue, Y., 2010. Understanding security behaviors in personal computer usage: a threat avoidance perspective. J. Assoc. Inf. Syst. 11 (7), 394. https://doi.org/10.17705/1jais.00232.
Mahmood, M.A., Siponen, M., Straub, D., Rao, H.R., Raghu, T.S., 2010. Moving toward black hat research in information systems security: an editorial introduction to the special issue. MIS Q. 34 (3), 431–433. https://doi.org/10.2307/25750685.
Miao, R., Li, B., 2022. A user-portraits-based recommendation algorithm for traditional short video industry and security management of user privacy in social networks. Technol. Forecast. Soc. Chang. 185, 122103 https://doi.org/10.1016/j.techfore.2022.122103.
Mills, A., Sahi, N., 2019. An empirical study of home user intentions towards computer security. In: Proceedings of the 52nd Hawaii International Conference on System Sciences. Available at. https://scholarspace.manoa.hawaii.edu/bitstream/10125/59921/1/0481.pdf.
Ngo, F.T., Paternoster, R., Cullen, F.T., Mackenzie, D.L., 2011. Life domains and crime: A test of Agnew's general theory of crime and delinquency. J. Crim. Justice 39 (4), 302–311. https://doi.org/10.1016/j.jcrimjus.2011.03.006.
Ogunrinde, A., 2022. The effectiveness of soft skills in generating dynamic capabilities in ICT companies. ESIC Market 53 (3), e286. https://doi.org/10.7200/esicm.53.286.
Ou, C.X., Zhang, X., Angelopoulos, S., Davison, R.M., Janse, N., 2022. Security breaches and organization response strategy: exploring consumers' threat and coping appraisals. Int. J. Inf. Manag. 65, 102498 https://doi.org/10.1016/j.ijinfomgt.2022.102498.
Parry, D.A., Davidson, B.I., Sewall, C., Fisher, J.T., Mieczkowski, H., Quintana, D., 2021. A systematic review and meta-analysis of discrepancies between logged and self-reported digital media use. Nat. Hum. Behav. 5, 1535–1547. https://doi.org/10.1038/s41562-021-01117-5.
Peltier, T.R., 2016. Information Security Policies, Procedures, and Standards: Guidelines for Effective Information Security Management. Auerbach Publications.
Pyrooz, D.C., Decker, S.H., Moule Jr., R.K., 2015. Criminal and routine activities in online settings: gangs, offenders, and the internet. Justice Q. 32 (3), 471–499. https://doi.org/10.1080/07418825.2013.778326.
Reyns, B.W., 2013. Online routines and identity theft victimization: further expanding routine activity theory beyond direct-contact offenses. J. Res. Crime Delinq. 50 (2), 216–238. https://doi.org/10.1177/0022427811425539.
Rogers, R.W., 1975. A protection motivation theory of fear appeals and attitude change. J. Psychol. 91 (1), 93–114. https://doi.org/10.1080/00223980.1975.9915803.
Romanosky, S., 2013. Comments to the department of commerce on incentives to adopt improved cybersecurity practices. Docket number 130206115-3115-01 (Available at), Information Law Institute, University of New York. https://www.ntia.doc.gov/files/ntia/romanosky_comments.pdf.
Rowell, D., Connelly, L.B., 2012. A history of the term "moral hazard". J. Risk Insur. 79 (4), 1051–1075. https://doi.org/10.1111/j.1539-6975.2011.01448.x.
Rowlingson, R.R., 2011. The Essential Guide to Home Computer Security. BCS, The Chartered Institute.
Siponen, M.T., Oinas-Kukkonen, H., 2007. A review of information security issues and respective research contributions. ACM SIGMIS Database DATABASE Adv. Inf. Syst. 38 (1), 60–80. https://doi.org/10.1145/1216218.1216224.
Souppaya, M., Scarfone, K., 2013. Guide to malware incident prevention and handling for desktops and laptops. Int. J. Comput. Res. 20 (4), 417. https://doi.org/10.6028/NIST.SP.800-83r1.
Symatec, 2022. Internet security threat report. Volume 23. Available at. https://www.symantec.com/content/dam/symantec/docs/reports/istr-23-executive-summary-en.pdf.
Talib, S., Clarke, N.L., Furnell, S.M., 2010. An analysis of information security awareness within home and work environments. In: Availability, reliability, and security, 2010. ARES'10 international conference on. IEEE, pp. 196–203. https://doi.org/10.1109/ARES.2010.27.
Tseloni, A., Wittebrood, K., Farrell, G., Pease, K., 2004. Burglary victimization in England and Wales, the United States, and the Netherlands a cross-national comparative test of routine activities and lifestyle theories. Br. J. Criminol. 44 (1), 66–91. https://doi.org/10.1093/bjc/44.1.66.
Urueña, A., Mateo, F., Navío-Marco, J., Martínez-Martínez, J.M., Gómez-Sanchís, J., Vila-Francés, J., Serrano-López, A.J., 2019. Analysis of computer user behavior, security incidents and fraud using Self-Organizing Maps. Comput. Secur. 83, 38–51. https://doi.org/10.1016/j.cose.2019.01.009.
Van Wilsem, J., 2011. Worlds tied together? Online and non-domestic routine activities and their impact on digital and traditional threat victimization. Eur. J. Criminol. 8 (2), 115–127.
Wang, J., Gupta, M., Rao, H.R., 2015. Insider threats in a financial institution: analysis of attack-proneness of information systems applications. MIS Q. 39 (1), 91–112. https://doi.org/10.25300/MISQ/2015/39.1.05.
Warkentin, M., Straub, D., Malimage, K., 2011. Measuring the dependent variable for research into secure behaviors. In: Decision Sciences Institute Annual National Conference. Boston, MA.
Warkentin, M., Straub, D., Malimage, K., 2012. Measuring secure behavior: a research commentary. In: Proceedings of the Annual Symposium on Information Assurance, pp. 1–8.
Wash, R., 2010. Folk models of home computer security. In: Proceedings of the Sixth Symposium on Usable Privacy and Security. ACM, p. 11. https://doi.org/10.1145/1837110.1837125.
Wehrens, R., Putter, H., Buydens, L.M., 2000. The bootstrap: a tutorial. Chemom. Intel. Lab. Syst. 54 (1), 35–52. https://doi.org/10.1016/S0169-7439(00)00102-7.
Werlinger, R., Hawkey, K., Beznosov, K., 2009. An integrated view of human, organizational, and technological challenges of IT security management. Inf. Manag. Comput. Secur. 17 (1), 4–19. https://doi.org/10.1108/09685220910944722.
Willison, R., Backhouse, J., 2006. Opportunities for computer crime: considering systems risk from a criminological perspective. Eur. J. Inf. Syst. 15 (4), 403–414. https://doi.org/10.1057/palgrave.ejis.3000592.
Willison, R., Warkentin, M., 2013. Beyond deterrence: an expanded view of employee computer abuse. MIS Q. 37 (1), 1–20. https://doi.org/10.25300/MISQ/2013/37.1.01.
World Economic Forum, 2016. Global risk report. Available at. https://www.weforum.org/reports/the-global-risks-report-2016.
Yar, M., 2005. The novelty of 'cybercrime' an assessment in light of routine activity theory. Eur. J. Criminol. 2 (4), 407–427. https://doi.org/10.1177/147737080556056.
Yoon, J., Vonortas, N.S., Han, S., 2020. Do-It-Yourself laboratories and attitude toward use: the effects of self-efficacy and the perception of security and privacy. Technol. Forecast. Soc. Chang. 159, 120192 https://doi.org/10.1016/j.techfore.2020.120192.

Alvaro Arenas is Professor and Chair of the Information Systems and Technology Department at IE Business School, IE University, Madrid, Spain. He holds a Ph.D. in Computation from Oxford University, UK. His main research interests include digital innovation, information security, and distributed information systems design. His research has been published in top journals such as the European Journal of Information Systems, Information & Management, Journal of Business Research, International Journal of Information Management, IEEE Computer, Internet Computing Journal, and Journal of the Association for Information Science and Technology, among others. Alvaro is very active in the European Research Area, having got leading positions in several EU-funded projects (LEAD, GridTrust, CoreGRID, XtreemOS). Currently, he is the PI for IE of the EU Digymatex project, investigating the use of digital technologies by children and teenagers.

Gautam Ray is a Professor at the Carlson School of Management at the University of Minnesota. His research interests are in the area of impact of IT on firm scope and structure, and how does IT create value. His research has appeared in Communications of the ACM, Information Systems Research, Management Science, Marketing Science, MIS Quarterly, Journal of Management Information Systems and the Strategic Management Journal. He received his Ph.D. from the Ohio State University in year 2000.

Antonio Hidalgo is full professor of Technology Strategy and Innovation, IPR and Technology Policy research group (INNOPRO), and Director of the Master and Doctorate in Economics and Innovation Management at Universidad Politécnica de Madrid (UPM). He received an MBA and a PhD in Industrial Engineering from the UPM. He has worked in several EU funded projects and participated as international consultant related to technological transfer in several projects financed by the World Bank and the Interamerican Development Bank. He is author of several books and papers about technology management published in different international journals.

Alberto Urueña is associate professor at ETSI of Industrial Engineering at Universidad Politécnica de Madrid (UPM, Spain) and holds a PhD in Economics and Innovation Management (UPM), and an Executive MBA at IE Business School. Alberto has published in leading scientific journals including, among others, International Journal of Information Management, Social Science Computer Review, Journal of Business Research, Journal of Risk Research and Telecommunications Policy. His research is oriented to the socio-technical impact of technological transformation with a sound interest in the drivers of Industry 4.0, digital innovation, and information security.