Ca' Foscari University of Venice

Department of Economics

Working Paper

Cinzia Di Novi
Matija Kovacic
Cristina Elisa Orso

Online Health Information Seeking Behavior, Healthcare Access, and Health Status During Exceptional Times

ISSN: 1827-3580
No. 26/WP/2023

Working Papers
Department of Economics
Ca' Foscari University of Venice
No. 26/WP/2023
ISSN 1827-3580

Online Health Information Seeking Behavior, Healthcare Access, and Health Status During Exceptional Times

Cinzia Di Novi
European Commission, Joint Research Centre (JRC)

Matija Kovacic
European Commission, Joint Research Centre (JRC); Department of Economics Ca' Foscari University

Cristina Elisa Orso
Department of Law, Economics, and Cultures, University of Insubria

Abstract
Online health information seeking behavior (e-HISB) is becoming increasingly common and the trend has accelerated as a result of the COVID-19 pandemic when individuals strongly relied upon the Internet to stay informed by becoming exposed to a wider array of health information. Despite e-HISB having become a global trend, very few empirical investigations have analyzed its potential impact on healthcare access and individuals' health status. In this paper, we try to fill this gap. We use data from the second SHARE Corona Survey and estimate a recursive model of e-HISB, healthcare access, and individuals' health status that accounts for individuals' unobserved heterogeneity. The most interesting result concerns the e-HISB indirect effect on individuals' poor health through healthcare access, that is positive. Arguably, patients use information from the Internet to cope with their perceived vulnerability to illness, but they lack the ability to understand the medical information: an incorrect self-diagnosis may increase the likelihood of doctor visits for them, which, in turn, also increases the likelihood of perceiving a poor health status.

Keywords: health information seeking behavior, healthcare access, health status

JEL Codes: I10, I12

Address for correspondence:
Matija Kovacic
Department of Economics
Ca' Foscari University of Venice
Cannaregio 873, Fondamenta S.Giobbe
30121 Venezia - Italy
e-mail: matija.kovacic@unive.it

The Working Paper Series
is available only on line
(http://www.unive.it/pag/16882/)
For editorial correspondence, please contact:
wp.dse@unive.it

Department of Economics
Ca' Foscari University of Venice
Cannaregio 873, Fondamenta San Giobbe
30121 Venice Italy
Fax: ++39 041 2349210

Online Health Information Seeking Behavior, Healthcare Access, and Health Status During Exceptional Times

Cinzia Di Novi[+], Matija Kovacic[+], Cristina Elisa Orso[♣]

Abstract

Online health information seeking behavior (e-HISB) is becoming increasingly common and the trend has accelerated as a result of the COVID-19 pandemic when individuals strongly relied upon the Internet to stay informed by becoming exposed to a wider array of health information. Despite e-HISB having become a global trend, very few empirical investigations have analyzed its potential impact on healthcare access and individuals' health status. In this paper, we try to fill this gap. We use data from the second SHARE Corona Survey and estimate a recursive model of e-HISB, healthcare access, and individuals' health status that accounts for individuals' unobserved heterogeneity. The most interesting result concerns the e-HISB indirect effect on individuals poor health through healthcare access, that is positive. Arguably, patients use information from the Internet to cope with their perceived vulnerability to illness, but they lack the ability to understand the medical information: an incorrect self-diagnosis may increase the likelihood of doctor visits for them, which, in turn, also increases the likelihood of perceiving a poor health status.

JEL classification: I10, I12

Keywords: health information seeking behavior, healthcare access, health status

[+] European Commission, Joint Research Centre (JRC), Via Enrico Fermi 2749, TP 267, 21027 Ispra (VA), Italy.
[♣] Department of Law, Economics, and Cultures, University of Insubria, Como, Italy

Introduction

The share of individuals who are turning to the Internet to obtain information about health has risen in the vast majority of European countries in recent decades, due to the widespread adoption of smartphones, tablets, and laptops. According to the Eurostat, around 52% of individuals searched for health-related information and symptoms online in 2022 and this proportion is still growing (Eurostat, 2022). Despite the Internet being an important and rapidly evolving source of health-related information, with negligible monetary and opportunity (time) costs (Bundorf et al., 2006; Costa-Font, et al., 2009; Suenaga & Vicente, 2022), concerns about the quality of information available on the Internet and the ability of individuals to assess and understand its credibility and content are raising questions about the implications of expansions in its use (Suziedelyte, 2021; Chen & Liu, 2022).

Even hypothesizing that all the relevant medical information is available online, individuals do not have complete information about their own health conditions and may have limited abilities by which to utilize online information in such a way as to efficiently adopt health-improving medical decisions independently of their physicians (Arrow, 1963; Dwyer & Liu, 2013). This means that doctor-patient relationships cannot be replaced by patients self-diagnosing and medicating based on what they have found on the Internet, at least in principle. However, online information seeking behavior (henceforth e-HISB) might affect the likelihood of visiting a health professional as well as the frequency of visits and, ultimately, individuals' health status.

There are two potential and contrasting hypothesis regarding the effects of e-HISB on physician visits according to Lee (2008): (i) online information seeking, by responding to patients' needs for health information, may negatively affect the likelihood of visiting health professionals and the frequency of visits; (ii) conversely, e-HISB might increase their health concerns and consequently their demand for physician visits and other medical services by making individuals more acutely aware about their health conditions.

Despite e-HISB having become a global trend, only a few empirical investigations on how health information seeking from the Internet affects healthcare access and individuals' health currently exist. Suziedelyte (2012) has investigated whether the health information that people obtain from the Internet affects their demand for healthcare using data from the U.S. Health Information National Trends Survey (2003-07). In the estimation model, she considered the endogeneity of Internet health information seeking in the demand for healthcare access equation and used information on U.S. states' right-of-way regulations in order to construct an instrument for e-HISB. Her paper's findings suggest that Internet health information seeking has a positive effect on the demand for healthcare. According to her results, e-HISB makes patients more concerned about their health compared to non-seekers. Greater health awareness, in turn, drives e-health information seekers to increase their number of health professional visits.

Suenaga and Vicente (2021) examined the relationship between e-HISB and the demand for physician services, using data collected from the 2014 Eurobarometer survey on European citizens' digital health literacy. Their analysis distinguished individuals seeking health information exclusively from offline sources from those seeking both online and offline sources. They used an extended sample selection model that addresses both the sample selection issue created by the survey design (i.e., the Eurobarometer survey collected data on offline health information searches only for those individuals who never sought health information online) and the endogeneity of health information seeking variables in the healthcare demand equation. The empirical analysis revealed that the demand for physician services is positively associated with offline health information seeking only, and not with e-HISB, in contrast with previous findings by Suziedelyte (2012).

Hone et al. (2016), tried to understand whether e-HISB affects the likelihood of bad self-rated health with a logit regression model by using data collected from the 2014 Eurobarometer survey on European citizens' digital health literacy. They distinguished between online seeking for general health and online seeking for disease-specific information. Their results show that searching for general information is less likely to be associated with self-reported bad health, whereas searching for disease-specific information increases the likelihood of self-reported bad health. However, the authors did not control for the e-HISB's endogeneity in the health equation and the related problems of reverse causality.

The previous literature failed to take the fact that e-HISB, individuals' health and healthcare access might be determined simultaneously into account. For instance, individuals striving to deal with health challenges, such as an illness diagnosis or chronic disease management, tend to be much more motivated in engaging in e-HISBs (see for instance Ayers et al, 2007; Weaver III et al, 2010): their health status may determine the demand for information to learn about a health or about illness-related concerns; e-HISB, as stated above, may affect (negatively or positively) the demand for healthcare services that, in turn, may influence individuals health status. Moreover, e-HISB is likely to be correlated to other variables that can also affect individuals' demand for health and healthcare. Individuals who are more efficient producers of health, such as those who are highly educated and who have a higher level of health literacy, for instance, also have a greater ability to find and to act upon online health information, but are also simultaneously more likely to have a greater demand for health and healthcare services (Bundorf et al., 2006; Costa-Font, et al., 2009).

The above discussion suggests that a step toward a complete understanding of the effects described requires a complex model that considers the simultaneous relationships between e-HISB, healthcare access, and an individual's health status. As such, we used a simultaneous equation model for binary variables; specifically, we constructed a joint model of e-HISB, healthcare access, and an individual's health status that considers individual's unobserved characteristics that are likely to be correlated with health information seeking, an individual's health status, and healthcare utilization. We examined the direct association between health information seeking via the internet and healthcare access and then we examined the direct association between e-HISB and an individual's health status and its indirect one through healthcare access by using a recursive multivariate probit design.

In this study, we specifically focused on adults aged 50 and over. Although older adults show lower rates of Internet adoption, when compared to younger adults, online health information seeking is becoming increasingly common among them and this trend has accelerated as a result of the COVID-19 pandemic (Lee & Jang, 2022; Symeonaki et al., 2022). Health deteriorates with age, so older adults may be more motivated to seek health-related information in order to cope with uncertainty, to stay informed about preventing diseases, and to look for others with similar health concerns; at the same time, however, older adults might be more reluctant to depart from traditional paternalistic models of healthcare due to their limited proficiency with both computer usage and the Internet (Bundorf et al., 2006; Mesch et al., 2012).

We used data collected in the second SHARE Corona Survey, and supplemented them with data from the previous 8th wave of SHARE, in order to assess both the potential merits and shortcomings of seeking health information online and how doing so may affect older adults healthcare access and health status. The second wave of the SHARE Corona Survey contains questions related to Internet access and the types of digital services used since the COVID outbreak (such as such as online banking, paying bills, or paying taxes, buying or selling goods, etc.) including questions concerning searches for information on health-related issues.

Individuals strongly relied on the Internet to stay informed during the COVID-19 pandemic outbreak and digital engagement grew in importance, especially among older adults because of lockdown mandates and social isolation (Suh et al., 2022); hence, COVID-19 serves as an exogenous source of variation. At the same time, the COVID-19 outbreak was characterized by the so-called *infodemic* phenomenon or an overabundance of health information available from a variety of (not always official or objective in nature) digital platforms that served to overwhelm the average person (WHO, 2020). We exploited the advantage of the SHARE Corona Survey in order to contribute to the understanding of e-HISB during exceptional times, such as the COVID-19 outbreak, and aim to highlight the importance of paying attention to the information needs of vulnerable groups such as the elderly.

Consistent with the previous literature, our findings show a positive effect of e-HISB on healthcare access that indicates that patients consider e-HISB and health professional visits as complements rather than substitutes. The effect of e-HISB on health, on the other hand, appears to be more complex. Indeed, our results show that while the direct effect of e-HISB on the likelihood of reporting poor health is negative, its indirect effect, through healthcare access, is positive. Arguably this is due to the fact that patients use information from the Internet to cope with their perceived vulnerability to illness but they lack the ability to understand the medical information: an incorrect self-diagnosis may increase the likelihood of doctor visits for them, which, in turn, also increases the likelihood of perceiving a poor health status.

The rest of the paper is organized as follows. Section 2 describes the data and variables used in this study and the empirical strategy deployed, including the estimation method. The results are discussed in Section 3. Finally, Section 4 summarizes and concludes the paper.

2. Data and Methods

2.1 Data

This study makes use of individual-level data drawn from the second SHARE Corona Survey. The first SHARE Corona Survey was implemented as a quick response within the SHARE study in order to understand the COVID-19 pandemic's effects. The interviews took place between June and September 2020 via a Computer Assisted Telephone Interview (CATI), partly to collect a set of basic information as in the regular SHARE questionnaire, and partly to elicit information on life circumstances amidst COVID-19. Respondents who participated in the first SHARE Corona Survey were interviewed again and participated in the second SHARE Corona Survey from June to August 2021 which contained questions on their use of the Internet, including its use for information about matters pertaining to health. In addition to the second SHARE Corona Survey dataset, we use data from the regular 8th wave of SHARE which collected information on the health, demographic, and socio-economic status of respondents aged 50 years and over. The interviews took place between October 2019 and March 2020.

The final sample consisted of 13.829 observations across 18 European countries after conditioning for having no missing values on any dependent variable and/or covariate, namely: Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, France, Italy, Latvia, Lithuania, Netherlands, Romania, Slovakia, Slovenia, Spain, and Sweden. The sample was restricted to exclude respondents living in European countries for which sub-national geographies were not available since our identification strategy is based on information collected from the Eurostat survey on the use of Information and Communication Technologies (ICT) in households and by individuals and the Eurostat data on the density of physicians at the Nomenclature of territorial units for statistics (NUTS) 2 level (see Subsection 2.5).[1]

2.2. Outcome Variables

We identified three classes of dependent variables for the empirical model: e-HISB, healthcare access, and individuals' general health status.

e-HISB was defined as a binary indicator of whether respondents had looked for information on health-related issues on the Internet since the COVID-19 outbreak. According to their answer, they were classified as either e-HI seekers or as non-e-HI seekers.

We created a binary variable indicating whether respondents went to a doctor's office or a medical facility in the last twelve months prior to the interview as a measure of health professional visits access.

We used the self-assessed health (SAH) as a measure of an individual's health status. The SAH is supported by literature that shows a strong predictive relationship between people's self-rating of their health and mortality or morbidity (Idler & Benyamini, 1997). Moreover, the self-assessed health measurement correlates strongly with more complex health indices, such as functional ability or indicators derived from health service use (Unden & Elofosson, 2006). The following standard self-assessed health status question was asked: 'Would you say that in general your health is: 1. Excellent, 2. Very good, 3. Good, 4. Fair, 5. Poor?" We dichotomized the multiple-category responses and constructed a binary indicator with a value of 1 if individuals reported that their health was fair or

[1] The EU survey on the use of Information and Communication Technologies (ICT) in households and by individuals has been conducted every year since 2002 and collects harmonized and comparable data on households' access to, and individuals' use of, the Internet.

poor, and 0 otherwise (i.e., excellent, very good, or good) because the answers could not simply be scored (for example as 1, 2, 3, 4, 5) because the true scale will not be equidistant between categories (O'Donnell et al., 2008) according to previous literature (see, for instance, Di Novi, 2010; Di Novi, 2013).[2]

All of the outcome variables were constructed according to the information included in the second SHARE Corona Survey.

2.3 Explanatory Variables

In our model, we controlled for a rich set of individuals' demographic and socio-economic characteristics, general health literacy, computer skills, and health conditions collected from the 8th wave of SHARE. For demographics, we included the respondent's sex, age, family size, geographic location (rural vs urban area), and a dummy variable for the region of residence. For socioeconomic characteristics, we included individuals level of education, marital status, occupation, and income.

The International standard classification of education (Isced) was used to classify the education variable. Three levels of education were considered: (1) low education (no educational certificates or primary school certificate or lower secondary education) as a reference category; (2) medium education (upper secondary education or high school graduation); and (3) high education (university degree or postgraduate). Marital status was categorized as 'living with a spouse or a partner in the same household' vs 'living as single' (reference category). Occupations were categorized into three groups: employed, retired, and other occupational status (namely unemployed, sick or disabled, homemakers, or other) as a reference category. Income information is based on total annual household income and was obtained by adding up its different components assessed in the questionnaire after deductions for income tax and social or national insurance contributions. It mainly comprises labor income, public pensions, and income from assets. Income was split into quartiles, with the lowest one as a reference category.

Suffering from health conditions is one of the most common reasons for accessing healthcare services, but also for gaining knowledge regarding health on the Internet (Rice, 2006). In our model,

[2] We carried out a sensitivity analysis re-running the model with a different cut-off point for SAH: we constructed a binary indicator that took value one for individuals who reported that their health was poor and zero otherwise (excellent, very good, good, and fair). This construction did not significantly affect the results: the multivariate probit coefficients with this new dependent variable were fairly similar to those presented in the paper. For the sake of brevity, the results of the sensitivity analysis are not included, but they are available on request.

we controlled for suffering from a chronic disease (high blood pressure; high blood cholesterol; stroke; diabetes; chronic lung disease; asthma; arthritis, osteoporosis; cancer; peptic ulcer; Parkinson's disease; cataracts; hip fracture; or other conditions). Specifically, we created a dummy variable that takes value 1 if respondents reported that they suffer from at least one chronic condition. We also included an indicator in the model of pre-existing general health conditions that were identified by using the SAH dummy indicator (fair and poor vs excellent, very good, and good) from the 8th wave of SHARE.

Health literacy and computer skills has been consistently reported as strong predictors for online health information seeking (Arnold et al., 2009; Kim, 2015). Moreover, health literacy can influence not only individuals' e-HISB, but also its associated demand for health and healthcare services according to the previous literature (see among others Sørensen et al., 2012; Ilic et al., 2022),. Indeed, health literacy affects the individuals' ability to "*access, understand, appraise, and apply health information*" to what concerns health behaviors, health care access, and ultimately health outcomes (Sørensen et al., 2012 page 3). Hence, the model also included indicators of respondents' computer skills and general health literacy. Concerning computer skills, respondents were asked: "How would you rate your computer skill? Would you say they are ...". A five-point scale was used for the response, ranging from poor to excellent. An additional category was "I never used a computer". We then derived a binary indicator that takes value 1 when respondents have at least good computer skills (i.e., when they reported 1. Excellent, 2. Very good, 3. Good) and zero otherwise (4. Fair, 5. Poor, 6. I never used a computer") (Cavapozzi & Dal Bianco, 2022). General health literacy was measured by using the Single-Item Literacy Screener (SILS) which was designed to identify adults in need of help with written or printed health material. Respondents were asked: "How often do you need to have someone help you when you read instructions, pamphlets or other written material from your doctor or pharmacy?" with answering options: 1. Never, 2. Rarely, 3. Sometimes, 4. Often and 5. Always. One again, we constructed a dummy variable with a value of one if respondents reported "Never" and zero otherwise ("Rarely", "Sometimes", "Often" and "Always"). Responses "Never" were selected to represent a good level of health literacy.

We observed individuals' e-HISB, health, and healthcare access during an exceptional time, namely the COVID-19 pandemic. The COVID-19 context was characterized by uncertainty and a strong need for information about the pandemic's evolution, the risks associated with coronavirus exposure, the community-level policies, and restrictions. The local virus spread might also have been a key factor in determining e-HISB, healthcare access, and individuals' health (especially in terms of psychological distress). Therefore, the model considered a variable related to the COVID-19 experience and to the spread of COVID-19 among respondents' contacts. This dummy indicator has a value of one if a respondent or anyone close to a respondent had suffered from the Coronavirus or had been hospitalized due to the infection or anyone close to a respondent died after having become infected by the Coronavirus, and 0 otherwise.

Finally, following Di Novi et al. (2023), we also included the COVID-19 Government Response Stringency Index (SI) from the Oxford Coronavirus Government Response Tracker (OxCGRT) in the model (Hale et al., 2021).[3] This index captures the day-to-day variation in the containment and closure policies adopted by national governments worldwide to tackle the pandemic. The index scores between 0 and 100, with a higher score indicating a more stringent response. The SI relies on the following measures: closures of schools and universities, closures of workplaces, cancelling public events, limits on gatherings, closures of public transport, orders to "shelter-in-place" and otherwise confined at home, restrictions on internal movement between cities/regions, restrictions on international travel, and the presence of public information campaigns.

It was possible to know each participant's interview month from the SHARE Corona Survey questionnaire. The average value of the SI was computed over the month of the interview in the respondent's country of residence. This value was then compared with the value of the SI in the same country by March, 12 2020 (the day after WHO declared COVID-19 as a pandemic) to compute the relative change in the SI which takes the potential mitigation/tightening in the COVID-19 restrictions over time into account from the beginning of the pandemic; this might have influenced an individual's healthcare access, e-HISB, and their health (especially in terms of psychological distress). We then constructed a binary variable that takes value 1 if the stringency index has been declining from the beginning of the COVID-19 pandemic to the period of observation at country level and 0 otherwise.

Table 1 sets out a full description of the variables used in the model.

[Table 1 about here]

2.4 Empirical Strategy

Identifying a causal relation between e-HISB, health professional visits, and health may be complicated by the presence of endogeneity, as stated previously. Indeed, e-HISB may be correlated

[3] Free publicly-accessible data collected by the OxCGRT was used; it is available here: https://www.bsg.ox.ac.uk/research/research-projects/ covid-19-government-response-tracker.

to either unobserved health characteristics or to unobserved preferences that are likely to influence the demand for health and healthcare services. As such, we estimated the model using a recursive multivariate probit design.[4] The multivariate probit model's recursive structure builds on two structural-form equations that determine the probability of bad health status and healthcare access, and one reduced-form equation for the potentially endogenous dummy variable measuring e-HISB. In the health professionals' equation, the e-HISB indicator is included as an explanatory variable. The inclusion of this indicator allowed us to test whether patients treat e-HISB as a substitute or in a complementary fashion for physicians access. e-HISB and access to physicians are included as regressors in the structural equation for health.

We constructed and estimated a system of three equations with one reduced-form equation and two structural equations. Thus:

$$\begin{gathered} Health\ Status_i = \delta_1 y Healthcare\ Access_i + \delta_2 \text{e-HISB}_\text{i} + \alpha_1' Z_{1i} + \varepsilon_{1i} \\ Healthcare\ Access_i = \gamma_2\, e\text{-}HISB_I + \alpha_2' Z_{2i} + \varepsilon_{2i} \\ e\text{-}HISB_i = \alpha_3' Z_{3i} + \varepsilon_{3i} \end{gathered} \tag{1}$$

where $\boldsymbol{Z}_{hi}$ (with h = 1, 2, 3) are vectors of exogenous variables, α_h are parameter vectors, and $\boldsymbol{\delta_o}$ (with o = 1, 2) and γ_2 are scalar parameters. The error terms distributed as multivariate normal are ε_{hi}, each with a mean zero and variance covariance matrix Σ. Σ has values of 1 on the leading diagonal and correlations $\rho_{jk} = \rho_{kji}$ on the off-diagonal elements (where ρ_{jk} is the covariance between the error terms of equation *j* and *k*).

The exogeneity condition is stated in terms of the correlation coefficients in the setting mentioned previously, which can be interpreted as the correlation between the different equations' unobservable explanatory variables. All equations in system (1) can only be estimated separately as single probit models in the case of independent error terms (i.e., the coefficient ρ_{jk} is not significantly different from zero).

[4] A recursive model is a special case of a system of equations in which the endogenous variables are determined in a sequence. Thus, the right-hand side of the reduced-form equations for the endogenous variables include exogenous variables only. The right-hand side of the structural equation includes the exogenous variables and the endogenous variables estimated by the reduced-form equations. The model's development may be traced back to the pioneering work of Heckman (1978), and it is a common approach to deal with the endogeneity of binary-dependent variables. See Di Novi et al., (2020) and Di Novi et al. (2023) for applications that use the multivariate probit model to estimate a recursive system similar to the one used here.

Conventionally, the identification of a recursive multivariate probit model has been based on exclusion restrictions in order to obtain a more robust identification of the parameters. According to Maddala (1983), at least one of exogenous variables of the e-HISB and physicians access equations (i.e., in the vectors $z_{2i}$and z_{3i}) are not included in the health equation as explanatory variables. However, more recent work by Wilde (2000) shows that identification is achieved even if the same regressors appear in all equations, providing that there is sufficient variation in the data (i.e., providing that each equation contains at least one varying exogenous regressor). However, this result is valid in the context of multivariate normal distribution and, in the absence of additional instruments, identification strongly relies upon functional form—i.e., the normality of the stochastic disturbances, commonly referred to as identification by functional form (Li et al., 2019). It is, therefore, common practice to impose exclusion restrictions in order to improve the identification of the causal parameters δ_1 and δ_2. These exclusion restrictions (instruments) should be causally linked to e-HISB and physicians access and should affect an individual's general health through their effects on e-HISB and access to a physician exclusively.[5] The instruments are discussed in detail in Subsection 2.5.

2.5 Exclusion Restrictions

This subsection describes the exclusion restrictions that we adopted for both e-HISB and healthcare access equations.

a. e-HISB equation

We exploited the heterogeneity in regional NUTS-2 on broadband coverage in order to deal with the potential endogeneity of e-HISB. Specifically, we used data from the Eurostat database on Information and Communication Technology (ICT) usage in households and by individuals and measure broadband internet diffusion with the variable that refers to the percentage of households with broadband internet access (*isoc_r_broad_h*).

In recent years, broadband infrastructures and network speed across the European countries have improved substantially. Most of the European countries have at least 80 percent of their households with broadband access enjoying high-speed connections (> 30 Mbps) and very high-speed

[5] The econometric literature, however, does not provide any formal test to show the contribution of the excluded instruments to the identification of the parameters.

connections (> 100 Mbps). Bulgaria and Italy show lower levels between 82 and 85 percent respectively.

We assumed that the high-speed connection increases the frequency of Internet use and the engagement with Internet activities, thereby facilitating information searches; moreover, the high-speed connection might also mean that individuals can access more content in a given amount of time. Hence, we expected that we might observe an association between the high-speed connection and e-HISB because of the enhanced internet access enabled by faster broadband speeds (McDool et al., 2020).

b. Health professional visits equation

In order to address the potential endogeneity of health professional visits' binary indicator in the health equation we included, in the vector Z_{2i}, an indicator of healthcare supply at the regional level (NUTS-2); namely, the number of medical doctors, including generalists and specialist medical practitioners per 1,000,000 inhabitants provided by Eurostat.

We expected that the number of doctors and their geographic distribution might influence the likelihood of accessing a health professional in normal circumstances and even more so during exceptional times, such as during the COVID-19 pandemic.

3. Results

Table 2 shows a simple descriptive analysis that presents sample means and standard deviations for the variables used in the model. About 36% of the study sample (60% female; mean age: 71 years) used the Internet during the COVID-19 outbreak to search for health information. We might note that the prevalence of bad health, based on SAH, increased from around 40% at the time of wave 8 to 42% by the time of the second COVID Survey. About 50% of respondents went to a doctor's office or a medical facility in the previous twelve months prior to the interview.

About 31 % of e-HI seekers reported suffering from fair or poor health against about 49% of non-seekers. The proportion of those who have accessed health professionals and medical facility is higher among e-HI seekers: about 56% against 47% among non-seekers. e-HI seekers are younger (mean age of e-HI seekers 67 years against 73 years of non-seekers) and have a higher level of health literacy and computer skills compared with non-seekers: about 85% of e-HI seekers reported having good health literacy against 66% of non-seekers; 55% of e-HI seekers also reported having higher computer skills against 18% of non-seekers respectively.

[Table 2 about here]

Table 3 shows the estimated marginal effects for the structural equations for bad health status and health professional visits and medical facility access and the reduced-form equation for an individual's e-HISB.

[Table 3 about here]

With specific reference to the reduced-form equation, our findings show that the indicator for broadband internet diffusion has a positive and significant effect on an individual's e-HISB, as expected: it increases the probability of accessing the Internet for searching health information by about 1.7%. Indeed, as stated previously, high-speed connections increase the frequency of Internet use by facilitating health information searches by reducing the opportunity time cost of accessing information on the Internet.

While the dummy indicator of pre-existing general health conditions, based on SAH, does not influence e-HISB, having been diagnosed with a chronic health condition increases the probability of being an e-HI seeker of about 3.4%. This result is consistent with the findings of previous studies that reported that one of the main reasons to go online to search for health information was having been diagnosed with a specific medical condition (see Bundorf et al., 2006; Rice, 2006; McMullan, 2006).

According to our results, individuals responded to the local spread of the coronavirus by searching for health information online (with a marginal effect of about 5.2%). The uncertainty of the nature of the disease and the method of transmission and treatment might have led individuals to seek out health-related information on COVID-19, coronavirus symptoms, and its treatment.[6]

e-HISB is negatively affected by age and living in a rural area, while it is positively affected by being female, married, highly educated, with a higher level of income, with a good level of health literacy, and computer skills as expected (which increased the probability of being an e-HI seeker from about 5.8% to 19.2% respectively).

With reference to the structural equation for the likelihood of visiting a health professional or accessing a medical facility (Column 2 in Table 3), our results show that an increasing number of medical doctors had a positive and significant effect on healthcare professionals access (+0.4%).

[6] One of the limitations of the second Corona survey database is that it does not allow for distinctions to be made between the type of health information that respondents looked for on the Internet. To that end, we were unable to identify those who looked for general health information, health information related to COVID-19 pandemic, and information strictly related to a specific disease.

Respondents in poor health have a greater demand for health professionals than those in better health and were more likely to go to either a doctor's office or a medical facility as expected (i.e., self-reported health status increases the probability of accessing a healthcare professional by about 4.1% while suffering from at least one chronic condition raised this likelihood by about 13.2%). According to our results, the likelihood of visiting a health professional or of accessing a medical facility is also positively affected by being female and highly educated and by the mitigation in the COVID-19 restrictions from the beginning of the pandemic (+3.3%).[7]

Finally, the results of the empirical analysis show that searching for health information on the Internet, all other factors being equal, has a positive and statistically significant effect on an individual's demand for healthcare professionals and medical facility access with a marginal effect of about 6%: it is apparent that e-HI seekers demand more health care than non-seekers. These findings seem to corroborate the hypothesis that e-HISB, by making individuals more aware about their health conditions, increases their health concerns and that this, in turn, may drive e-HI seekers to visit either a health professional or a medical facility (Lee, 2008; Suziedelyte, 2012). Hence, patients do not see the Internet as a replacement for the health professional, but as a complementary component according to our results.

With reference to the structural equation for individuals' SAH (Column 3 in Table 3), our results show that while e-HISB negatively affects the probability of perceiving bad health with a marginal effect of about -1.7%, visiting a health professional or accessing medical facilities increases the likelihood of perceiving bad health by about 5.7%. This result seems counterintuitive at first glance. Some studies actually suggest that the indicator of self-assessed health is strongly associated with psychosocial factors, such as positive mood, negative mood, and perceived vulnerability to illness. These factors appear to be significant contributors to self-assessed health independently of an individual's physical dimensions, such as physical symptoms and diseases. According to these studies, psychosocial factors play a relatively major role in how "healthy" we feel compared to physical discomfort (see, for instance, Andersen & Lobel, 1995; Benyamini et al., 2000). It may be that, in terms of direct effect of e-HISB on an individual's self-perceived health, information from the Internet led individuals to manage their health conditions more effectively, thereby decreasing their perceived

[7] Healthcare systems experienced major disruptions during the initial months of the COVID-19 pandemic; this included restrictions, implied temporary closures of medical practices, and cancellation or postponement of non-emergency and elective procedures. The spread of information about the threat of the coronavirus exacerbated the fear of infection, especially among the elderly and chronically ill individuals, who, in several cases, forewent medical treatments (Di Novi & Santos, 2023).

vulnerability to illness and consequently the likelihood of reporting a poor health status. Conversely, in terms of indirect effect of e-HISB through healthcare access, it may also be that patients who lack the ability to understand medical information that they find on the Internet, in principle, decided to make health-related decisions independently. However, incorrect online health information and a wrong self-diagnosis increases the likelihood for them to visits doctor and that this, in turn, also increases the likelihood of perceiving poor health status.

We tried to empirically test this hypothesis by disentangling the indirect effect of e-HISB through healthcare access on individuals' health status. More precisely, we decomposed the total effect of healthcare access on the probability of reporting poor general health conditions into its direct effect, and an its effect conditional on e-HISB (i.e., the indirect effect of e-HISB on individuals health status through healthcare access). These results are included in Table 4.

[Table 4 about here]

As expected, we found that healthcare access itself has a negative effect on the likelihood of reporting poor health of about 1% . However, conditional on e-HISB, its effect turns out to be positive (+11%) and absorbs the direct beneficial effect of medical visits on the health status, resulting in a positive and statistically significant total effect (+10%). This evidence is in line with our hypothesis: patients use information from the Internet to cope with their perceived vulnerability to illness, but they lack the ability to understand the medical information. As a consequence, an incorrect self-diagnosis increase the probability of visiting a health professional or accessing medical facilities for them, which, in turn, also increases the likelihood of perceiving a poor health status.

Concerning the other variables included in the structural equation, our findings show that a good level of health literacy, higher computer skills, and a higher socioeconomic status in general were associated with a lower probability of reporting poor health that increases with higher age and previously existing poor health conditions.

As discussed previously, we constructed a simultaneous equation model for three binary variables. The multivariate probit estimation allowed us to test for unobserved heterogeneity that may characterize the relationship between e-HISB and an individual's healthcare access and health status. The unobserved heterogeneity is captured by the correlation between the error terms from the single equation models. Table 5 shows the full recursive model's correlation coefficients.

[Table 5 about here]

The null hypothesis of exogeneity is rejected in only one case. According to our results, there exists a positive and statistically significant correlation between the disturbance of the e-HISB equation and the structural equation for individuals' health status—i.e., unobservable variables that increase the likelihood of bad health and also increase the probability of searching for health information online.

4. Conclusions

This study investigated whether access to health information on the Internet is likely to affect an individual's health and healthcare-related decisions in an exceptional time such as the COVID-19 pandemic which has had an enormous impact on people worldwide, subjecting the global population to health risks, fear, anxiety, and to an incredible amount of health information. This was done using data from the 8th wave of the Survey of Health, Ageing, and Retirement in Europe (SHARE) and from the second wave of SHARE Corona Survey.

This paper's main contribution consisted in analyzing the simultaneous relationship between e-HISB, an individual's health, and healthcare access. A multivariate probit approach was used to estimate recursive systems of equations for self-assessed health, e-HISB, and healthcare access.

Consistently with the previous literature, the results of this analysis show that the effect of Internet health information seeking on health care utilization is both positive and statistically significant. Thus, patients do not see the Internet as a replacement for a healthcare professional, but as a complement thereto.

The effect of e-HISB on health appeared to be much more complex. Indeed, while the estimated direct effect of e-HISB on an individual's poor self-perceived health is negative, the indirect effect of e-HISB on individuals' poor health status through healthcare access is positive. These results are consistent with the interpretation that, in principle, patients use information from the Internet to cope with their perceived vulnerability to illness but they lack the ability to understand the medical information that they find on the Internet. Incorrect online health information and an incorrect self-diagnosis increases the likelihood to doctor visits for them; this, in turn, also increases the likelihood of perceiving a poor health status.

Our findings highlight the importance of paying attention to the informational needs of vulnerable groups such as the elderly, especially those with a lower level of education and health literacy for whom the misinterpretation of health-related information still remains an open issue.

References

1. Andersen, M., & Lobel, M. (1995). Predictors of health self-appraisal: what's involved in feeling healthy?. *Basic and Applied Social Psychology*, 16(1-2), 121-136.
2. Arnold, M., Hsu, L., Pipkin, S., McFarland, W., & Rutherford, G. W. (2009). Race, place and AIDS: the role of socioeconomic context on racial disparities in treatment and survival in San Francisco. *Social Science & Medicine*, 69(1), 121-128.
3. Arrow, K. J. (1963). Uncertainty and the welfare economics of medical care. *The American Economic Review*, 53(5), 941–973.
4. Au, N., & Johnston, D. W. (2014). Self-assessed health: what does it mean and what does it hide?. *Social science & medicine*, 121, 21-28.
5. Ayers, S. L., & Kronenfeld, J. J. (2007). Chronic illness and health-seeking information on the Internet. *Health*:, 11(3), 327-347.
6. Bahia, K., & Castells, P. (2022). The impact of spectrum assignment policies on consumer welfare. *Telecommunications Policy*, 46(1), 102228.
7. Benyamini, Y., Idler, E. L., Leventhal, H., & Leventhal, E. A. (2000). Positive affect and function as influences on self-assessments of health: Expanding our view beyond illness and disability. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 55(2), P107-P116.
8. Bris, M., Pawlak, J., & Polak, J. W. (2017). How is ICT use linked to household transport expenditure? A cross-national macro analysis of the influence of home broadband access. *Journal of Transport Geography*, 60, 231-242.
9. Eurostat – Statistics Explained (2022) https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Digital_economy_and_society_statistics_-_households_and_individuals – accessed March 2023.
10. Bundorf, M. K., Wagner, T. H., Singer, S. J., & Baker, L. C. (2006). Who searches the internet for health information?. *Health Services Research*, 41(3p1), 819-836.
11. Cavapozzi, D., & Dal Bianco, C. (2022). Does retirement reduce familiarity with Information and Communication Technology? *Review of Economics of the Household*, 20(2), 553-577.
12. Chen, L., & Liu, W. (2022). The effect of Internet access on body weight: Evidence from China. *Journal of Health Economics,* 85, 102670.

13. Costa-Font, J., Mossialos, E., & Rudisill, C. (2009). When is the Internet a valued communication device for health information in Europe?. *Economics of Innovation and New Technology*, 18(5), 429-445.
14. Di Novi, C., & Santos, R. (2023). Health and aging: The sustainability and equity trade-off. *Journal of Economic Behavior & Organization,* 205: A8-A10;
15. Di Novi C., Paruolo P., Verzillo S. (2023) "Does labour protection influence mental-health responses to employment shocks? Evidence on older workers in Europe", *Economics Modelling,* 126: 106406.
16. Di Novi, C., Leporatti, L., & Montefiori, M. (2020). Older patients and geographic barriers to pharmacy access: When nonadherence translates to an increased use of other components of health care. *Health Economics*, 29, 97-109.
17. Di Novi C. (2013). The Indirect Effect of Fine Particulate Matter on Health through Individuals' Life-style, *Journal of Socioeconomics*, 44: 27-36.
18. Di Novi C. (2010). The Influence of Traffic-related Pollution on Individuals' Life-Style: Results from the BRFSS, *Health Economics*, 19:1318–1344.
19. Dwyer, D. S., & Liu, H. (2013). The impact of consumer health information on the demand for health services. *The Quarterly Review of Economics and Finance*, 53(1), 1-11.
20. Hale, T., Angrist, N., Goldszmidt, R. et al. (2021). A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker). *Nature Human Behavior*, 5: 529–538.
21. Heckman, J.J., (1978). Dummy endogenous variables in a simultaneous equation system. *Econometrica* 46 (6), 931.
22. Idler E. L., Benyamini Y. (1997). Self-rated health and mortality: a review of twenty-seven community studies, *Journal of Health and Social Behavior*, 38: 21–37.
23. Kim, Y. M. (2015). Is seeking health information online different from seeking general information online? *Journal of Information Science*, 41(2), 228-241.
24. Hone, T., Palladino, R., & Filippidis, F. T. (2016). Association of searching for health-related information online with self-rated health in the European Union. *The European Journal of Public Health*, 26(5), 748-75.
25. Ilic, A., Roser, K., Sommer, G., Baenziger, J., Mitter, V. R., Mader, L., ... & Michel, G. (2022). COVID-19 Information-Seeking, Health Literacy, and Worry and Anxiety During the Early

Stage of the Pandemic in Switzerland: A Cross-Sectional Study. *International Journal of Public Health,* 67, 1604717.

26. Lee, J., & Jang, S. N. (2022). Have changes in Internet use during the COVID-19 pandemic affected older adults' self-rated health? A cross-sectional study of young-old and old-old populations in Korea. *Geriatric Nursing*, 48, 141-145.
27. Li, C., Poskitt, D. S., & Zhao, X. (2019). The bivariate probit model, maximum likelihood estimation, pseudo true parameters and partial identification. *Journal of Econometrics*, 209(1), 94-113.
28. Maddala G. S. (1983). Methods of estimation for models of markets with bounded price variation, *International Economic Review*, 361-378.
29. McDool, E., Powell, P., Roberts, J., & Taylor, K. (2020). The internet and children's psychological wellbeing. *Journal of Health Economics*, 69, 102274.
30. McMullan, M. (2006). Patients using the Internet to obtain health information: how this affects the patient–health professional relationship. *Patient education and counseling*, 63(1-2), 24-28.
31. Mesch, G., Mano, R., & Tsamir, J. (2012). Minority status and health information search: A test of the social diversification hypothesis. *Social Science & Medicine*, 75(5), 854-858.
32. O'Donnell O. et al. (2008). Analyzing Health Equity Using Household Survey Data. The World Bank: Washington, DC.
33. Rice, R. E. (2006). Influences, usage, and outcomes of Internet health information searching: multivariate results from the Pew surveys. *International Journal of Medical Informatics*, 75(1), 8-28.
34. Sørensen, K, Van den Broucke, S, Fullam, J, Doyle, G, Pelikan, J, Slonska, Z, et al. Health Literacy and Public Health: A Systematic Review and Integration of Definitions and Models. *BMC Public Health* (2012) 12(1):80.
35. Suenaga, H., & Vicente, M. R. (2022). Online and offline health information seeking and the demand for physician services. *The European Journal of Health Economics*, 1-20.
36. Suh, J., Horvitz, E., White, R. W., & Althoff, T. (2022). Disparate impacts on online information access during the Covid-19 pandemic. *Nature Communications*, 13(1), 7094.
37. Suziedelyte, A. (2012). How does searching for health information on the Internet affect individuals' demand for health care services? *Social Science & Medicine*, 75(10), 1828-1835.
38. Undon A.L., Elofsson S. (2006). Do different factors explain self-rated health in men and women?, *Gender Medicine*, 3(4).

39. Symeonaki, M., Filandrianos, G., & Stamou, G. (2022). Visualizing key information and communication technologies (ICT) indicators for children and young individuals in Europe. *Humanities and Social Sciences Communications*, 9(1), 1-12.
40. Weaver III, J. B., Mays, D., Weaver, S. S., Hopkins, G. L., Eroğlu, D., & Bernhardt, J. M. (2010). Health information–seeking behaviors, health indicators, and health risks. *American Journal of Public Health*, 100(8), 1520-1525.
41. Wilde, J. (2000). Identification of Multiple Equation Probit Models with Endogenous Dummy Regressors. *Economics Letters* 69:309–12.
42. World Health Organization (2020). Managing the COVID-19 infodemic: Promoting healthy behaviours and mitigating the harm from misinformation and disinformation. https://tinyurl.com/k4zrca9e.

TABLES

Table 1: Variables Description and Data Source

e-HISB:	1 if respondent uses internet to look for health information (data source: second SHARE Corona Survey)
tech01:	percentage of individuals, at NUTS2, that have broad band access(data source: Eurostat survey on the use of ITC)
health professional visits	1 if respondent declares to have seen doctor/medical facility other than hospital since last interview (data source: second SHARE Corona Survey)
n. of physicians	n. of physicians/1.000.000 inhabitants, at NUTS2 level (data source: Eurostat data on the density of physicians)
SAH	1 if respondent suffered from fair-poor health (data source: second SHARE Corona Survey)
SAH_{t-1}	1 if respondent suffered from fair-poor health (data source: 8^{th} wave of SHARE)
chronic_condition	1 if respondent suffered from at least one chronic disease in 2019/2020 (data source: 8^{th} wave of SHARE)
rural	1 if respondent lives in rural area (data source: 8^{th} wave of SHARE)
health_literacy	1 if respondent declares no need to help with reading health information (data source: 8^{th} wave of SHARE)
female	1 if respondent is female (data source: 8^{th} wave of SHARE)
hhsize	n. of individuals within the household (data source: 8^{th} wave of SHARE)
age2021	age as continuous variable (data source: 8^{th} wave of SHARE)
local_spread_covi d	1 if respondent or anyone close to a respondent had suffered from the Coronavirus or was hospitalized due to the infection or anyone close to a respondent died after being affected by the Coronavirus (data source: second SHARE Corona Survey)
decline_stringency _index	1 if the stringency index in the country of residence has been declining from the beginning of the COVID-19 pandemic to the month of the interview (data source: Oxford Coronavirus Government Response Tracker - OxCGRT)
d_high_pcskill	1 if respondent has at least good computer skills, and zero otherwise (data source: 8^{th} wave of SHARE)
marital_status	1 if the individual is lives with , 0 otherwise (data source: 8^{th} wave of SHARE)
retired	1 if respondent is retired (data source: 8^{th} wave of SHARE)
employed	1 if respondent is employed (data source: 8^{th} wave of SHARE)
other_occupations	1 if respondent l is unemployed, sick or disabled, home maker or is doing other (data source: 8^{th} wave of SHARE)
low_ed	1 if respondent reported low education level (data source: 8^{th} wave of SHARE)
med_ed	1 if respondent reported medium level of education (data source: 8^{th} wave of SHARE)

high_ed	1 if respondent reported high education level (data source: 8th wave of SHARE)
quartiles	dummies for income quartiles (data source: 8th wave of SHARE)

Table 2: Descriptive Statistics.

Panel A: Full sample

Variable	Mean	Std. Dev.	N
e-HISB	0.362	0.481	13,829
tech01	89.783	3.578	13,829
health professional visits	0.501	0.500	13,829
n. of physicians	363.876	68.574	13,829
SAH	0.424	0.494	13,829
SAH_{t-1}	0.408	0.491	13,829
chronic_conditions	0.744	0.436	13,829
rural	0.361	0.480	13,829
health_literacy	0.731	0.443	13,829
d_high_pcskill	0.323	0.468	13,829
female	0.602	0.489	13,829
hhsize	2.062	0.980	13,829
age2021	71.206	9.108	13,829
local_spread_covid	0.402	0.490	13,829
decline_stringency_index	0.302	0.459	13,829
marital_status	0.675	0.469	13,829
retired	0.691	0.462	13,829
employed	0.199	0.399	13,829
other_occupations	0.098	0.298	13,829
low_ed	0.322	0.467	13,829
med_ed	0.457	0.498	13,829
high_ed	0.221	0.415	13,829
1°quartile	0.277	0.447	13,829
2°quartile	0.279	0.448	13,829
3°quartile	0.241	0.428	13,829
4°quartile	0.202	0.402	13,829

Panel B: Non -HI seekers.

Variables	Mean	Std. Dev.	N
tech01	89.593	3.657	8,812
health professional visits	0.466	0.499	8,812
n. of physicians	361.820	69.055	8,812
SAH	0.487	0.500	8,812
SAH_{t-1}	0.471	0.499	8,812
chronic_conditions	0.777	0.416	8,812
rural	0.402	0.490	8,812
health_literacy	0.664	0.472	8,812
d_high_pcskill	0.179	0.383	8,812
female	0.603	0.489	8,812
hhsize	2.035	1.021	8,812
age2021	73.422	9.155	8,812
local_spread_covid	0.357	0.479	8,812
decline_stringency_index	0.328	0.470	8,812
marital_status	0.634	0.482	8,812
retired	0.758	0.428	8,812
employed	0.121	0.326	8,812
other_occupations	0.108	0.311	8,812
low_ed	0.417	0.493	8,812
med_ed	0.444	0.497	8,812
high_ed	0.138	0.345	8,812
1°quartile	0.347	0.476	8,812
2°quartile	0.304	0.460	8,812
3°quartile	0.209	0.406	8,812
4°quartile	0.140	0.347	8,812

Panel C: e-HI seekers

Variables	Mean	Std. Dev.	Obs
tech01	90.118	3.410	5,017

health professional visits	0.563	0.496	5,017
n. of physicians	367.499	67.574	5,017
SAH	0.315	0.464	5,017
SAH_{t-1}	0.296	0.457	5,017
chronic_conditions	0.687	0.464	5,017
rural	0.287	0.452	5,017
health_literacy	0.849	0.358	5,017
d_high_pcskill	0.576	0.494	5,017
female	0.602	0.490	5,017
hhsize	2.111	0.900	5,017
age2021	67.302	7.587	5,017
local_spread_covid	0.481	0.500	5,017
decline_stringency_index	0.257	0.437	5,017
marital_status	0.746	0.435	5,017
retired	0.572	0.495	5,017
employed	0.338	0.473	5,017
other_occupations	0.082	0.274	5,017
low_ed	0.154	0.361	5,017
med_ed	0.478	0.500	5,017
high_ed	0.368	0.482	5,017
1°quartile	0.155	0.362	5,017
2°quartile	0.234	0.424	5,017
3°quartile	0.299	0.458	5,017
4°quartile	0.311	0.463	5,017

Table 3: Multivariate Probit Model—Marginal effects _ SAH 2021 as dependent variable

	e-HISB	Health Professional Visits	SAH
SAH_{t-1}	-0.026	0.041***	0.350***
	(0.037)	(0.031)	(0.040)
chronic_conditions	0.034***	0.132***	0.110***
	(0.034)	(0.029)	(0.037)
rural	-0.055***	-0.001	0.018*
	(0.049)	(0.041)	(0.028)
health_literacy	0.058***	0.021	-0.080***
	(0.069)	(0.040)	(0.038)
d_high_pcskill	0.192***	-0.001	-0.036***
	(0.019)	(0.016)	(0.015)
female	0.023***	0.016**	-0.003
	(0.031)	(0.023)	(0.028)
hhsize	-0.016*	-0.009*	0.002
	(0.022)	(0.015)	(0.018)
age2021	-0.013***	-0.001	0.003***
	(0.002)	(0.002)	(0.002)
local_spread_covid	0.052***	0.039***	0.020***
	(0.031)	(0.030)	(0.028)
decline_stringency_index	0.012*	0.033***	-0.009
	(0.033)	(0.034)	(0.031)
marital_status	0.038***	0.008	-0.003
	(0.038)	(0.035)	(0.046)
retired	0.047***	0.040***	-0.013
	(0.049)	(0.042)	(0.043)
employed	0.038***	0.008	-0.043**
	(0.046)	(0.046)	(0.058)
med_ed	0.097***	0.012	-0.025**
	(0.036)	(0.032)	(0.033)
high_ed	0.205***	0.045***	-0.049***
	(0.041)	(0.041)	(0.039)
2°quartile	0.030**	0.024**	-0.021
	(0.054)	(0.034)	(0.049)
3°quartile	0.072***	0.027*	-0.030**
	(0.055)	(0.042)	(0.048)
4°quartile	0.077***	0.012	-0.033
	(0.055)	(0.057)	(0.064)

tech01	0.017***		
	(0.008)		
n. of physicians		0.004***	
		(0.001)	
e-HISB		0.061***	-0.017***
		(0.052)	(0.062)
health professional visits			0.057***
			(0.072)
N	13829	13829	13829

Notes: Significance levels: * $p<0.1$, ** $p<0.05$, *** $p<0.01$. All the reported coefficients are average marginal effects. Standard errors are clustered at NUTS2 level.

Table 4: Total, Direct and Indirect Marginal Effects - Health Care Utilization and Health Information Seeking

Total effect	Marginal Effects	SE
Health Care Utilization	0.104***	0.013
Direct effect		
Health Care Utilization	-0.008*	0.006
Indirect Effect		
Health Care Utilization	0.112***	0.013

*Notes: Significance levels: * $p<0.1$, ** $p<0.05$, *** $p<0.01$. All the reported coefficients are average marginal effects.* Standard errors are clustered at NUTS2 level.

Table 5: Multivariate Probit Model - Correlation between the error terms

	e-HISB	Health professional visits	SAH
e-HISB	1	0.008(0.028)	0.070*(0.029)
Health professional visits		1	-0.026(0.036)
SAH			1

Standard errors in parentheses
*Legend: * = 10% significance level*