Explaining Policyholders’ Chatbot Acceptance with an Unified Technology Acceptance and Use of Technology-Based Model

Jorge de Andrés-Sánchez 1,* and Jaume Gené-Albesa 2

1 Social and Business Research Group, Department of Business Administration, University Rovira i Virgili, Campus de Bellissens, 43204 Reus, Spain
2 Department of Business Administration, University Rovira i Virgili, Campus de Bellissens, 43204 Reus, Spain; jaume.gene@urv.cat
* Correspondence: jorge.deandres@urv.cat

Abstract: Conversational robots powered by artificial intelligence (AI) are intensively implemented in the insurance industry. This paper aims to determine the current level of acceptance among consumers regarding the use of conversational robots for interacting with insurers and seeks to identify the factors that influence individuals’ behavioral intention to engage with chatbots. To explain behavioral intention, we tested a structural equation model based on the Unified Theory of Acceptance and Use of Technology (UTAUT) model. It was supposed that behavioral intention is influenced by performance expectancy, effort expectancy, social influence, and trust, and by the moderating effect of insurance literacy on performance expectancy and effort expectancy. The study reveals a significant overall rejection of robotic technology among respondents. The technology acceptance model tested demonstrates a strong ability to fit the data, explaining nearly 70% of the variance in behavioral intention. Social influence emerges as the most influential variable in explaining the intention to use conversational robots. Furthermore, effort expectancy and trust significantly impact behavioral intention in a positive manner. For chatbots to gain acceptance as a technology, it is crucial to enhance their usability, establish trust, and increase social acceptance among users.

Keywords: insurtech; fintech; chatbots; Unified Theory of Acceptance and Use of Technology; structural equation modeling; partial least squares

1. Introduction

Fintech, which can be defined as the application of technological advances to improve financial products and services [1], has experienced explosive development in recent years because of the impact of digital technologies such as artificial intelligence on the financial industry [2]. Thus, because Insurtech is a branch of fintech that focuses on the insurance sector [3], it can be defined as the application of Industry 4.0 technologies to offer new solutions to the challenges that arise in the provision of products and services as well as in the development of processes in the insurance industry [4]. Although Insurtech has experienced sustained growth since the beginning of the 2010s [5,6], the SARS-CoV-2 crisis has strengthened this trend, as it has caused the introduction of many measures to provide promised services to customers who were hindered by mobility limitations [7].

The implementation of Insurtech must be understood comprehensively, given that Industry 4.0 encompasses all levels of the insurers’ value chain, including risk and data analysis, sales, fraud investigation, and asset and liability management [4]. In fact, companies that adopt an integrated approach in the implementation of the digital agenda tend to have better business performance [5]. As outlined in [8], it also embraces customer experience, since the use of instruments powered by artificial intelligence allows for an extremely personalized service.
Industry 4.0 techs have been applied intensively in the insurance industry [6]. This work is focused on the use of conversational robots, whose implementation in the insurance sector began in the mid-2010s [9]. They can be defined as software that mimics human conversations using artificial intelligence algorithms whose purpose is to allow the user to find answers to their queries [10]. It is expected that chatbots will play an important role in improving the services offered by the insurance industry, but their effectiveness will depend on customer satisfaction [11]. Therefore, it is relevant to understand the factors that affect the acceptance of chatbots for carrying out procedures related to current policies.

In theory, chatbots have become an effective tool for addressing user queries in an automated way [12], so many companies have employed them to provide assistance to customers [13]. Conversational bots are promising for enhancing consumer support since they can simplify processes, avoid waits, and offer unrestricted availability around the clock. This enables human agents to concentrate on handling nonroutine and intricate matters, thereby leveraging their expertise effectively [14]. In an insurance environment, chatbots can perform tasks such as searching for suitable products for potential customers based on some keywords, providing information to policyholders about current policies and processes, or speeding up and facilitating claims processes [15]. Several studies report that the implantation of a chatbot system in the organization is judged positively by the stockholders, since news in this regard is followed by an enhancement of the value or the organization [13].

Robotics science is experiencing rapid growth, and chatbots developed with artificial intelligence, which are the most advanced conversational agents, are programmed to interact like real human beings and continuously learn from conversations to provide more accurate answers [12]. In fact, there have been chatbot prototypes capable of passing the Turing test since 2014, when the Eugene Goodman prototype did so [16]. However, the technology behind chatbots is not yet fully mature and often presents flaws [17]. Chatbots are not capable of capturing the evolution of a conversation by interpreting the tones and inflections of the customer’s voice, lack emotional skills such as empathy, and cannot respond to complex requirements [18]. These limitations may explain the reluctance of many insurance consumers to interact with chatbots [9,19].

This article evaluates attitudes toward the use of chatbots in insured–insurer interactions in Spain through the use of a structured survey answered by more than 250 persons. The analysis is based on the Unified Theory of Acceptance and Use of Technology (UTAUT) by Venkatesh et al. [20], which has been used in a great number of papers to explain the acceptance of a great number of fintechs [21]. This paper also introduces trust into the analysis, which has been proven to be a key factor in explaining the acceptance of robot technology [22] and the adoption of artificial intelligence technology in the Insurtech setting [23]. We have also taken into account the possible moderating effect of insurance literacy on the effects of performance expectancy and effort expectancy on behavioral intention. This work addresses two research questions:

RQ1. What is the average intention to be assisted by conversational robots to make procedures such as declaring a claim?

RQ2. Does UTAUT, along with trust and insurance literacy, provide a statistically significant framework for the analysis of chatbot acceptance for policy management?

To answer both questions, we used a structured questionnaire that was answered by 252 respondents. Whereas RQ1 can be answered by simply observing the average punctuation of the questions linked to intention to use bots, the second research question was assessed by adjusting a structural equation model that was fitted using the partial least squares technique.

The rest of the article is structured as follows. In the second section, a review of the literature related to our analysis is carried out, which justifies the use of UTAUT to explain attitudes toward the use of chatbots with the addition of trust and insurance literacy. The third section describes the materials and analytical methodology adopted in the study.
Subsequently, the results of the data analysis are presented. Finally, the main findings are discussed, practical implications are outlined, and the limitations of the study are commented on, suggesting future research.

2. A Unified Technology Acceptance and Use of Technology-Based Model to Assess Behavioral Intention of Policyleholders to Use Chatbots

2.1. Initial Considerations

This study tests a model of policyholders’ chatbot tech acceptance based on the Unified Theory of Acceptance and Use of Technology (UTAUT) [20] that was refined in [24]. It is a synthesis of previous technology acceptance models, such as the Technology Acceptance Model (TAM) and its extensions [25,26], or the Theory of Planned Behavior [27]. UTAUT links performance expectancy (PE), effort expectancy (EE), and social influence (SI) with the behavioral intention (BI) to use a given technology. The fourth original explanatory factor, facilitating conditions, is supposed to directly impact the actual use of the technology and not on behavioral intention [20], which is the assessed output variable in this work. Likewise, the interaction with bots has negligible digital barriers since even a conventional phone is able to use them. In any case, chatbots have fewer impediments to their use than other digital technologies, such as apps [28]. Therefore, this factor is not considered in this analysis. On the other hand, to explain BI, we added trust. The reason for considering trust is that it has been proven to be a key factor in explaining customer acceptance of chatbots [22] and is a cornerstone factor of the insurance industry [29].

In all business sectors, at the time of writing this paper, the introduction of chatbots to provide customer service is not being implemented uniformly and consistently across all companies [13]. This statement also holds true in the insurance sector [5]. For instance, while some companies may make it mandatory for customers to interact with a chatbot initially, with the option to continue with the bot or request assistance from a human operator based on their satisfaction, other companies allow customers to choose their preferred mode of interaction right from the beginning. Therefore, we understand that the meaning of the variable usage may be heterogeneous in the sample, as it could initially be mandatory for some of the survey respondents. In contrast, “behavioral intention” is a homogeneous concept for all respondents and is particularly relevant for measuring the acceptance of technology by customers. In fact, in technology acceptance models such as the UTAUT model [20,24] or the TAM [25,26], perceived usefulness and effort expectancy influence behavioral intention. Furthermore, the majority of the literature reviewed, which is used to formulate our hypotheses, focuses in some cases primarily on explaining behavioral intention rather than current usage, which in many cases may be driven by mandatory requirements rather than behavioral intention.

It should also be noted that two demographic factors, age and gender, which are considered potential moderators in the work [20], are taken into consideration, but without making any concrete hypothesis about the sign of their moderating capability. The negligible presence of digital barriers in the use of bots makes it difficult to construct any hypothesis regarding the effect of age. Regarding the moderating capability of gender, it should stem from social aspects related to the roles assigned to a specific gender in the society within which the study is framed. In our case, it is difficult to identify gender-related issues to make insurance procedures in the population under study and at the time of the study (the 2020s). Thus, even though in an initial stage we will assess the mediating capacity of the two highlighted sociodemographic factors, only those factors that initially demonstrate a statistically significant effect will be definitively considered in the model.

Likewise, we have also considered possible moderating effects of insurance literacy on the effects of performance expectancy and effort expectancy on behavioral intention, which is supported by arguments provided in [30] in a fintech context and [31] in the evaluation of chatbot acceptance by customers of online travel agencies. It should be noted that, in [20,24], experience is considered a moderating variable, particularly in terms of effort expectancy. In these studies, it is assumed that experience pertains to the technology
being used rather than the task itself, which is the regular job. Therefore, it is implied that experience in the task exists within the target population under study. The variable insurance literacy should be considered in another way in which the construct of experience can be manifested. However, in this case, it pertains to the task being performed (managing active policies), as proficiency in this aspect varies among the target population due to it not being a regular task. On the other hand, using the technology does not require any special experience, as theoretically, it is sufficient to *speak or chat* [15,18].

In the reviewed literature on technological acceptance in finance and insurance, as well as the use of chatbots for customer service, the mainstream is built upon the foundation of the TAM or UTAUT, which are adapted to the specific technology and its intended use. While a TAM model suggests that behavioral intention is solely influenced by perceived usefulness and effort expectancy, with other relevant factors such as social influence (or subjective norm) impacting behavioral intention mediated by these basic constructs [26], UTAUT postulates that these constructs, including social influence, directly influence the intention to use [20]. Moreover, following UTAUT, this study does not consider a hypothetical indirect impact of effort expectancy on behavioral intention through performance expectancy [20], which TAM does take into account [25]. Consequently, the model examined in our paper, as illustrated in Figure 1, is grounded in the UTAUT framework rather than provided by TAM.

![Figure 1. Inner model tested in this paper. Source: Own elaboration with the ground of Venkatesh et al. [20].](image)

2.2. Direct Effects of Performance Expectancy, Effort Expectancy, Social Influence and Trust on Behavioral Intention

The explained outcome in this paper is behavioral intention. Following [20,25], behavioral intention is the main antecedent of usage and can be defined as “persons’ readiness to perform a given behavior” that, in our context, consists of interacting with chatbots to implement procedures about existing policies.

*Performance expectancy* (PE), which is equivalent to so-called perceived usefulness in TAM, can be defined as the degree of perception by potential users that the assessed system will help to improve the attained performance when developing a task [20]. The usefulness of chatbots to make insurance procedures can be justified with several arguments. First, it allows basic demands to be solved more quickly than if exclusively using human assistance [14]. Likewise, bots do not replace other communication channels for interacting with the insurance company but can be understood as an alternative tool to improve service to policyholders [32]. The diversification of communication channels is usually appreciated and fosters customer loyalty [33].
The capability of Insurtech to increase the value of services provided to policyholders and insureds and to enhance production and marketing efficiency [3] must allow the insurer to achieve advantages over its competitors, which could be offering better products and service and/or diminishing prices.

On the other hand, at the time of writing this work, chatbots have several disadvantages that can reduce their usefulness. Generally, chatbots cannot solve complex situations, and therefore, management must be resolved by a human operator [18]. In those cases, the first interaction with the chatbot is perceived as a waste of time and a decrease in its perceived usefulness [17].

Under certain conditions, the utilization of chatbots can have an adverse effect on the perception of the service. When policyholders experience a significant loss, such as the destruction of their home, they anticipate the interlocutor to provide attentive support [34]. Studies have highlighted a notable drawback in employing chatbots for customer service: their inherent absence of empathy [18].

To truly benefit from bot technology, companies must achieve good coordination between data input and processing in the back-end system so that the front-end system can provide reliable responses and appropriate service [9]. Otherwise, the poor quality of front-end system responses results in a poor perception of the usefulness of the interaction [26].

The papers examined reveal a considerable consensus that perceived ease of use (PE) plays a favorable role in shaping individuals’ perceptions of technological excellence. This outcome has been observed consistently across various domains, including blockchain applications in finance [35–38], digital banking [39–43], intention to use online applications in the insurance industry [22,44], and acceptance of conversational robot services [10,45–54]. Therefore, the following hypothesis is proposed in this study:

**Hypothesis 1 (H1).** The performance expectancy of chatbots positively influences behavioral intention toward their use in customer procedures with the insurance company.

Following Venkatesh et al. [20], effort expectancy (EE) can be defined as the intensity of the belief that to employ a given tech is not difficult, and is essentially the perceived ease of use of TAM [25]. In this study, EE should be understood as the absence of obstacles to communicating and carrying out transactions with the insurance company. Theoretically, chatbots provide some advantages over alternative channels. They are capable of providing assistance 24 h a day, 7 days a week, and are more flexible than humans in terms of when they are available [14]. Moreover, they have fewer barriers to use than other digital technologies, such as apps [28].

On the other hand, there is currently a consensus that chatbot technology is not sufficiently developed to allow for easy interaction in many circumstances. It is very common for chatbots to offer confusing responses to customers, which worsens their perception of usability and usefulness and therefore negatively impacts their acceptance [14]. Some issues that can be highlighted in this regard are technological robotic phobia, which has a significant impact on customers’ perception of usability and attitude toward bots [55], and the fact that chatbots do not capture voice tones or the direction of a particular conversation [18].

The drawbacks noted in the above paragraphs lead to bots frequently making errors in real interactions [56]. This problem hinders ease of use [17] and explains why more than half of the interactions with bots are not completed [19]. The literature has also shown the relevance of this variable to explain behavioral intention in contexts similar to those analyzed in this paper, such as blockchain technology [35–38], digital banking [39,42], insurtech [44], and attitudes toward chatbots by customers [22,46–48,52,53,57]. Therefore, the following hypothesis is proposed.

**Hypothesis 2 (H2).** The effort expectancy of chatbots positively influences behavioral intention toward their use in customer transactions with the insurer.

**Social influence** (SI) is “the degree to which an individual perceives that important others believe they must use a new tech” [20]. In this regard, the opinion of personal
financial advisors and insurance brokers is usually significant in decisions about insurance issues [21]. Although conversational robots have become commonplace in the provision of services to customers, a significant number of consumers still harbor doubt and exhibit hesitance in regard to interacting with them [58]. However, despite this skepticism, there continues to be strong commercial interest in chatbot technology, primarily driven by the recognized advantages it offers. Consequently, businesses actively promote the adoption of conversational bots among their customers [11]. In fact, that tech is commonly employed to deliver initial assistance to both customers and users of public services [18].

It is a widely established fact that the opinion of peers such as friends or family members impacts on the perception and attitude toward new techs because people often act to attain the approval of appreciated persons [20]. This also applies to areas related to financial communication channels [59], new financial instruments such as crowdfunding [60], and the acceptance of conversational bots [47].

As is frequently observed when disruptive technologies are integrated, numerous jobs associated with the insurance industry will become obsolete. Projections indicate that by 2030, the adoption of Industry 4.0 technologies such as chatbots will lead to a drastic reduction of 70% to 80% in the workforce required to cater to customers’ needs [61]. This consequence can be perceived as a negative social utility that is common to all Industry 4.0 technologies [62], thus negatively impacting the behavioral attitude toward the tech. Likewise, although many positive benefits can be obtained from techs powered by AI, it is also true that there is an extended social perception that AI implies several dangers, such as job losses. Some examples are that multinational companies may use AI to assemble enormous nonauditible influence and power; the fear that engines powered by AI could become more intelligent than humans, and this jeopardizes humanity; or issues with privacy [63]. Therefore, we propose to test the following hypothesis.

Hypothesis 3 (H3). Social influence has a positive influence on policyholders’ behavioral intention to use chatbots in their procedures with the insurer.

In empirical studies of technology acceptance, it is common to utilize theoretical frameworks such as the Technology Acceptance Model (TAM) and its extensions [25,26], the Theory of Planned Behavior [27], or UTAUT [20,24], while introducing new variables that are not covered by these theoretical frameworks but are relevant to the analyzed technology and its context. In our case, the relevance of trust (TRUST) in the relationships between the insurer and their company is significant, as it encompasses two aspects: the relational and cognitive dimensions [23]. This explains why, within the empirical literature related to our research, it is common to complement the technology acceptance theory used as the theoretical framework with the introduction of trust as a relevant construct. For instance, the acceptance of chatbots by consumers [47] has been studied from within the theory of reasoned action, and [10,46,57] introduce trust in a UTAUT formulation. In studies on the applications of blockchain technology in finance [35,38], TAM is used as the foundation for including trust as an explanatory factor. Within e-banking channels [39,41], trust is evaluated within the TAM framework. Likewise, [64] introduce trust in a UTAUT-based model to evaluate the acceptance of e-payment methods, and [44] also introduces that construct in their TAM model for online insurance evaluation.

Trust should have an impact on policyholders’ acceptance of Insurtech in two distinct ways. First, given the unique characteristics of financial and insurance industries, coupled with the fact that the interaction between the company and policyholders is mediated by technology, trust emerges as an exceptionally relevant factor in elucidating customers’ attitudes and behavioral intentions [23]. Second, trust grounds financial agreements [64], and this fact is especially true in the insurance industry. Here, both parties involved in the transaction, namely, the insurer and the policyholder/insured, must exhibit trust in each other within an environment characterized by a heightened level of moral hazard and a substantial likelihood of adverse selection. Consequently, trust is essential in insurance market development [29]. Within the realm of insurance, we can define trust on the part of
policyholders in insurance companies as their perception of the ability of these companies to offer secure compensation in the event of a loss and to ensure satisfactory interactions in this regard [29]. This facet of trust is commonly referred to as relational or organizational trust [65].

Similarly, the concept of trust plays a crucial role in understanding the adoption of emerging technologies such as chatbots, particularly in business-to-consumer (B2C) interactions [66]. Individuals often harbor doubts regarding these tools, primarily because of the absence of transparency [47]. In the realm of remote financial services, trust can be defined as customers’ confidence in a particular mode of interaction, wherein they believe that companies will be capable of delivering satisfactory service through this channel [39]. This aspect of trust is tied to what is known as cognitive trust [65].

Insurtech such as chatbots have among their objectives to improve the value of products offered to customers [7,15], so it is expected that greater trust in their benefits will imply a more positive attitude toward their adoption and greater perceived usefulness in their services. There are several empirical studies that show that the direct impact of trust on the attitude and intention to use novel technologies is significant in contexts such as blockchain technology [35,38], digital banking [39,41,64], and Insurtech [44]. In the field of chatbot acceptance, we can highlight the works [10,22,46,48,53,57,67]. Therefore, the following hypothesis is proposed:

**Hypothesis 4 (H4). Trust has a positive influence on policyholders’ behavioral intention to use chatbots in their procedures with the insurer.**

2.3. The Moderating Effects of Insurance Literacy, Gender, and Age

The significance of experience, which is linked to self-efficacy as a moderator of constructs influencing behavioral intention, has been postulated by widely recognized theories of technology acceptance such as TAM [26] and UTAUT [20,24]. In these works, experience exclusively pertains to the evaluated technology, as it is applied in tasks where it is assumed that the user possesses extensive experience, for instance, in their own work. Conversely, in our case, heterogeneity with respect to experience lies in an aspect where the consumer may lack knowledge (insurance management), but the utilization of technological advancements does not necessitate special technological skills [28]. In such cases, experience and literacy in the relevant domain can also exert influence, either interacting with other factors [68], through moderating effects [31,69], or through mediation [70] on behavioral intention.

Considering that the literature on technology acceptance is more prevalent in finance than in the insurance industry, which is logical due to the former being a broader and more general field, an initial definition of insurance literacy (IL) can be derived by adapting the definition of financial knowledge proposed by Stolper and Walter [71]. Consequently, insurance literacy can be understood as the extent of understanding regarding essential insurance concepts and the ability to apply them in practical situations. In greater detail, [72] identifies insurance literacy as encompassing dimensions such as knowledge and understanding of personal risk exposure, methods for managing risk, functioning of insurance mechanisms, selection of suitable insurance policies, understanding of the rights and obligations of policyholders, and familiarity with the insurance inquiry process. Furthermore, insurance literacy involves the utilization of this knowledge in decision-making processes related to insurance.

In any case, drawing an analogy with the effects of financial literacy on behavioral intention as presented in several papers, it is expected that insurance literacy will have a positive impact on the influence of performance expectancy on behavioral intention. In the field of m-banking, [73] reported the positive influence of financial literacy on intention to use through the mediation of perceived usefulness. Additionally, in the context of mobile banking, [30] argues that individuals with higher levels of financial literacy are more adept at assessing the financial implications of technology adoption, leading to greater performance expectancy, as they perceive clear advantages from its usage. Similarly, in the realm of fintech, [69] found a significant positive impact of literacy on behavioral intention.
by moderating perceived usefulness. This moderating effect is also reported in [74] when explaining attitudes toward taking financial risks.

In the specific context of insurance decisions, [75,76] also discovered a positive mediating effect of insurance literacy through factors associated with performance expectations on customers’ behavioral intention and attitude toward subscribing to insurance policies.

A similar perspective is presented in a study examining the acceptance of conversational chats in relationships with online travel agencies [31]. This study found that greater familiarity with online travel bookings can positively moderate the impact of expected usefulness on behavioral intention.

Furthermore, experience and knowledge within the field where a new technology is implemented are expected to have a negative moderating role in effort expectancy, as they carry less significance in decision-making for individuals with greater experience [20,26,70]. The moderating influence of financial literacy on effort expectancy when adopting new bank channels has been observed in [69,77]. Based on these findings, the following hypotheses are proposed:

**Hypothesis 5a (H5a).** Insurance literacy positively moderates the effect of performance expectancy on behavioral intention.

**Hypothesis 5b (H5b).** Insurance literacy negatively moderates the effect of effort expectancy on behavioral intention.

Age and gender are considered potential moderators in [20]. This fact motivates us to take the first step in testing the hypothesis presented in Figure 1 by examining their moderating capabilities. In [20], it is justified that gender can serve as a moderating variable for PE, EE, and SI, as men are typically more task-oriented and less adverse to risk compared to women. Conversely, women tend to place greater value on the new system’s potential for enhancing ease to use and are more sensitive to subjective norms. Additionally, [20] justifies age as a moderating variable for PE, EE, and SI based on the observation that younger individuals tend to be more responsive to external rewards. Older users may have a reduced ability to interact with new systems while devoting attention to job-related responsibilities. Furthermore, they often possess a stronger sense of belonging to a social group.

However, we believe that the potential mediating role of sociodemographic variables depends on various factors, such as the specific technology being evaluated, the context in which it will be used, and the society in which it is tested. In the case of interacting with conversational bots, where there are no digital impediments compared to apps [28], constructing hypotheses regarding the moderating effect of age becomes challenging, as our context is not directly linked to job tasks.

Regarding the moderating effect of gender, [20,24] emphasize that it should arise from social aspects and the specific roles assigned to a particular gender within the societal framework of the study. In our particular case, it is difficult to identify gender-related issues within the population under study in the year 2023.

Only a limited number of studies have examined the moderating capability of age and gender on performance expectancy (PE), effort expectancy (EE), and social influence (SI) [36,42,52]. In [36], both variables are considered, but the significance of their moderating capacity is not reported. Conversely, in [42], the inclusion of the gender variable is motivated by the potential relevance of gender roles stipulated by religious factors in Kenya, and indeed, this moderation is found to be significant in that study. In [52], the potential moderating capacity of the gender variable is explored in relation to an experiment on the perception of gendered anthropomorphic design cues moderating behavioral intention toward chatbots. However, the results presented in this paper are inconclusive.

Regarding the mediating effect of sociodemographic variables on effort expectancy, there are also very few studies in the reviewed literature. While [36] does not report the results, [42] indicates the significance of gender. However, studies such as [52] and [57] either fail to reach conclusive results or reject the significance of the gender variable. Similar
patterns are observed with regard to social influence. The limited number of studies that have examined the mediating effect of age and gender [42,52] have reported contradictory findings. Thus, we propose the following hypothesis:

**Hypothesis 6 (H6).** There is no significant moderating effect of gender and age on the effect of performance expectancy, effort expectancy, and social influence on behavioral intention.

### 3. Materials and Methods

#### 3.1. Materials

The survey used a structured questionnaire that was written and provided in Spanish. The English version is shown in Table 1. It began with the following introductory text: “We require your opinion on the management of your current policies regarding contacting the insurer and using automated systems such as voice and text robots instead of a human operator. For example, consider a common procedure such as declaring a claim”. The questions were answered on an 11-point Likert scale that ranged from 0 (complete disagreement) to 10 (complete agreement). While it is recognized that the original UTAUT employed a seven-point Likert scale and that subsequent reviewed papers have utilized scales ranging from five to seven grades, current mainstream psychometric scales tend to follow the five–seven-point range [78]. However, a literature review on scales [78] states that scales with five or seven points do not adequately capture the full range of human capability to discern nuances. Typically, people tend to avoid using extreme values when responding, resulting in only one (or two) viable option(s) to express agreement or disagreement for many respondents when using a five (seven)-point scale. Additionally, scores on lower-point scales tend to be higher than those on scales with a greater number of grades. These considerations have led several authors to suggest that an 11-level scale is more suitable than scales with fewer points. Moreover, an 11-point scale allows for an “indifferent” response, and respondents can provide answers in a more intuitive manner, given the prevalence of decimal numerical systems in everyday situations [78]. It should be noted that the implementation of the UTAUT has also utilized 11-point scales in various studies. For instance, [79] applied the UTAUT theoretical framework to explain the acceptance of cyborg techs, [80] did so in the context of sport technology, and [68] evaluated the behavioral intention to invest in cryptocurrencies. Alesanco-Llorente et al. [81], adopting the cognitive–affective–normative (CAN) theory of technology, employed an 11-point Likert scale to assess behavioral intention toward mobile augmented reality techs, while [82] did the same in the field of wine innovations.

The survey was provided first to six practitioners of the Spanish insurance market who were members of the International Association for Insurance Law (Spanish section). After obtaining their feedback and taking it into account in the new version of the questionnaire, it was distributed to another 12 respondents who were not affiliated with the financial and insurance sector. After considering these latest comments, we established the final questionnaire. The comments from those 18 volunteers improved the readability of the questionnaire but did not imply substantial changes and, moreover, allowed us to make a first test of the measurement scales.

The next step consisted of sharing the survey through social networks (LinkedIn, Telegram, moderated mailing lists, etc.). The survey was completed online from 20 December 2022 to 12 March 2023. Since we were looking for opinions from consumers who were truly informed, we only accepted responses from people who were actually holders of at least two policies. The total number of responses obtained was 252. However, responses where there were unanswered items were not used. Therefore, the final number of responses analyzed in this paper was 226. It allowed a margin of error of ±5% to ±7.5%, which, given that the reference population is large (>5000), we judged to be correct.
Table 1. Constructs and items.

<table>
<thead>
<tr>
<th>Item</th>
<th>Foundation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Behavioral intention</strong></td>
<td></td>
</tr>
<tr>
<td>BI1. I agree in interacting with a bot to make procedures with my insurer.</td>
<td>Grounded in [20,25]</td>
</tr>
<tr>
<td>BI2. I believe that I will employ conversational bots to interact with the insurer regarding my policies.</td>
<td></td>
</tr>
<tr>
<td>BI3. I will opt to manage existing policies with conversational robots.</td>
<td></td>
</tr>
<tr>
<td><strong>Performance expectancy</strong></td>
<td></td>
</tr>
<tr>
<td>PE1. Chatbots make procedures with the insurance company easier.</td>
<td>Grounded in [20,24] and un used on [48] in the assessment of AI acceptance</td>
</tr>
<tr>
<td>PE2. Chatbots allow for a faster resolution of issues with my policies.</td>
<td></td>
</tr>
<tr>
<td>PE3. Chatbots allow for making procedures with the insurance company with less effort.</td>
<td></td>
</tr>
<tr>
<td><strong>Effort expectancy</strong></td>
<td></td>
</tr>
<tr>
<td>EE1. Using chatbots to communicate with the insurer is easy.</td>
<td>Based on [20,48,59]</td>
</tr>
<tr>
<td>EE2. Managing claims and other procedures with the insurer through chatbots is clear and understandable.</td>
<td></td>
</tr>
<tr>
<td>EE3. The help of chatbots in managing policies and claims is accessible and less prone to errors.</td>
<td></td>
</tr>
<tr>
<td>EE4. It is easy to use the communication channels of the insurer smoothly through chatbots.</td>
<td></td>
</tr>
<tr>
<td><strong>Social influence</strong></td>
<td></td>
</tr>
<tr>
<td>SI1. Persons who are important for me think that chatbots makes easier insurance procedures.</td>
<td>Grounded in [20] and used by [48,59]</td>
</tr>
<tr>
<td>SI2. The people who influence me feel that if there is possibility to choose a channel, better interact with bots.</td>
<td></td>
</tr>
<tr>
<td>SI3. Persons whose opinions I value feel that making insurance procedures with bots is a step forward.</td>
<td></td>
</tr>
<tr>
<td><strong>Trust</strong></td>
<td></td>
</tr>
<tr>
<td>TRUST1. I feel that conversational bots are reliable.</td>
<td>Based on [40] that was grounded in [83]. Used in [57] within a chatbot context.</td>
</tr>
<tr>
<td>TRUST2. The use of chatbots enable the insurance company to fulfil its commitments and obligations.</td>
<td></td>
</tr>
<tr>
<td>TRUST3. The use of chatbots to interact with the insurer considers the interests of policyholders.</td>
<td></td>
</tr>
<tr>
<td><strong>Insurance literacy</strong></td>
<td></td>
</tr>
<tr>
<td>IL1. I have a good level of knowledge about insurance matters.</td>
<td>Based on [84]</td>
</tr>
<tr>
<td>IL2. I have a high ability to apply my knowledge about insurance in practice.</td>
<td></td>
</tr>
</tbody>
</table>

Table 2 shows the profile of the people who responded to the survey. It can be observed that a very high proportion of them are over 40 years old. It can also be seen that the distribution by gender is balanced. Additionally, a percentage close to 50% claim to have more than four policies. Most respondents (approximately 90%) reported being at least college graduates. This high proportion could be attributed to the fact that a significant portion of responses come from LinkedIn, which is composed of a very high proportion of professionals with a university degree. Regarding the most common income profile, it can be considered high (≥EUR 3000, 35%) or medium (between EUR 1750 and EUR 3000), which is consistent with the fact that we were seeking responses from decision-makers with at least two insurance policies.

Table 2. Sociodemographic profile of the respondents.

<table>
<thead>
<tr>
<th>Gender</th>
<th>Age</th>
<th>Academic degree</th>
<th>Income</th>
<th>Number of policies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male: 52.03%</td>
<td>≤40 years: 14.375</td>
<td>At least graduate: 88.62%</td>
<td>≥EUR 3000: 34.55%</td>
<td>&gt;4 contracts: 52.03%</td>
</tr>
<tr>
<td>Women: 45.53%</td>
<td>&gt;40 years and &lt;55 years: 53.89%</td>
<td>Undergraduate: 11.38%</td>
<td>≥EUR 1750 and &lt;EUR 3000: 35.77%</td>
<td>≥2 and &lt;4: 47.97%</td>
</tr>
<tr>
<td>Other/NA: 2.44%</td>
<td>≥55 years: 29.94%</td>
<td></td>
<td>&lt;EUR 1750: 29.67%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>NA: 1.80%</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
3.2. Data Analysis

RQ1. what is average intention to be assisted by conversational robots to make procedures such as declaring a claim?, was answered by comparing the central tendency measures of the responses, mean and median, of the items of behavioral intention with 5, which can be considered the neutral value on an eleven-point Likert scale. This analysis was implemented using Student’s t and Wilcoxon tests.

RQ2: Does UTAUT, along with trust and insurance literacy, provide a statistically significant framework for the analysis of chatbot acceptance for policy management? was answered by fitting a structural equation model that was adjusted with partial least squares (PLS-SEM) to the model proposed in Figure 1 and developed in Section 2. One positive aspect of using PLS-SEM is that it can fit complex models without the need for explicit assumptions about the data distribution. Additionally, it does not require a high number of observations [85].

In the initial stage, the model presented in Figure 1 was adjusted, also taking into account the hypothetical mediating effect of age and gender as postulated in the original formulation of UTAUT [20]. This allows the testing of H6. Subsequently, we re-adjusted the model developed in Section 2, considering only the age and gender moderation effects that were found to be statistically significant in the first step. In this regard, age was defined as a dichotomous variable, taking a value of 1 if the individual was not older than 50 years, and 0 otherwise. Gender was also defined as a dichotomous variable, taking a value of 1 if the respondent identified as female and 0 otherwise.

Note that the construct that absorbs the highest number of direct interactions (represented by the arrows) is behavioral intention. It absorbs four direct interactions (from perceived usefulness, effort expectancy, social influence, and trust), along with the additional two interactions corresponding to the mediation of insurance literacy, resulting in a total of six interactions and six more mediation effects that come from age and gender. Therefore, applying the 10 times rule, we can determine that a minimum of 120 observations is required to implement the proposed model using PLS-SEM if all the possible interactions are considered that are reduced to 60 whether effects from age and gender are not under analysis [86].

Additionally, if we consider the inverse R squared method and set a minimum R² value of at least 10% for the constructs being predicted (in this case, only behavioral intention), then following [86], for a significance level of 5% and a power of 80%, our sample size also fulfills the criterion, because N > 203.

To fit the model in Figure 1, we used SmartPLS4.0 software following a standard procedure. First, we evaluated the measurement model by verifying the internal consistency of the scales and their discriminant validity. To check internal consistencies, we used the usual measures: Cronbach’s alpha, composite reliability (CR), Dijkstra and Henseler’s ρ_A, [87] which should all be above 0.7, and the average extracted variance (AVE), which should be greater than 0.5. Furthermore, the factor extraction of the items should be above 0.7. The discriminant validity of the constructs was assessed with Fornell–Larker’s criterion [88] and the value of heterotrait–monotrait ratios (HTMT), which follows [89], must be <0.9 [89].

The estimation of path coefficients β is performed using the bootstrapping technique with 5000 subsamples and the studentized method. The p values allowed us to test the hypotheses we proposed in the second section. The overall quality of the fit is quantified with the coefficient of determination R², which is complemented by the standardized root mean square residual (SRMR) and normed fit index (NFI). The predictive capability of the model was stated with the Stone–Geisser Q² coefficient and, following [90], by using the cross-validated predictive ability test (CVPAT).

4. Results

From the results displayed in Table 3, we can deduce that the measurement scales of BI, PE, EE, SI, TRUST, and IL are reliable because Cronbach’s alpha, CR, and ρ_A > 0.7, VME > 0.5. Moreover, the factor loadings for all items were >0.7. From Table 4, we can conclude that
Fornell and Larker’s criteria [88] indicate that the constructs have discriminant capability because the correlations of constructs are always below the square root of their AVE. The HTMT ratio criterion also suggests the existence of general discriminant power by the constructs proposed. The exception is the pair PE and TRUST, with a HTMT, despite being <1, of >0.90.

Table 3. Descriptive statistics and analysis of internal consistency of scales.

<table>
<thead>
<tr>
<th>Items</th>
<th>Mean</th>
<th>Median</th>
<th>Std. Deviation</th>
<th>Factor Loading</th>
<th>Cronbach’s α</th>
<th>CR</th>
<th>ρA</th>
<th>AVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>BI</td>
<td>1.27</td>
<td>0</td>
<td>1.87</td>
<td>0.921</td>
<td>0.891</td>
<td>0.895</td>
<td>0.932</td>
<td>0.822</td>
</tr>
<tr>
<td>BI1</td>
<td>2.24</td>
<td>1</td>
<td>2.70</td>
<td>0.861</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BI2</td>
<td>1.38</td>
<td>0</td>
<td>2.06</td>
<td>0.935</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.918</td>
<td>0.921</td>
<td>0.948</td>
<td>0.859</td>
</tr>
<tr>
<td>PE1</td>
<td>2.71</td>
<td>2</td>
<td>2.66</td>
<td>0.929</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PE2</td>
<td>2.57</td>
<td>2</td>
<td>2.58</td>
<td>0.925</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PE3</td>
<td>2.46</td>
<td>2</td>
<td>2.61</td>
<td>0.926</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.928</td>
<td>0.932</td>
<td>0.949</td>
<td>0.823</td>
</tr>
<tr>
<td>EE1</td>
<td>2.88</td>
<td>2.5</td>
<td>2.82</td>
<td>0.869</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EE2</td>
<td>2.69</td>
<td>2</td>
<td>2.61</td>
<td>0.942</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EE3</td>
<td>2.16</td>
<td>2</td>
<td>2.27</td>
<td>0.906</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EE4</td>
<td>2.64</td>
<td>2</td>
<td>2.64</td>
<td>0.910</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.927</td>
<td>0.929</td>
<td>0.953</td>
<td>0.872</td>
</tr>
<tr>
<td>SI1</td>
<td>1.75</td>
<td>1</td>
<td>1.94</td>
<td>0.922</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SI2</td>
<td>1.61</td>
<td>1</td>
<td>2.05</td>
<td>0.953</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SI3</td>
<td>2.03</td>
<td>1</td>
<td>2.15</td>
<td>0.927</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TRUST</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.83</td>
<td>0.865</td>
<td>0.897</td>
<td>0.745</td>
</tr>
<tr>
<td>TRUST1</td>
<td>2.07</td>
<td>1</td>
<td>2.50</td>
<td>0.912</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TRUST2</td>
<td>3.46</td>
<td>3</td>
<td>3.04</td>
<td>0.836</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TRUST3</td>
<td>2.08</td>
<td>2</td>
<td>2.18</td>
<td>0.839</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.969</td>
<td>0.988</td>
<td>0.985</td>
<td>0.971</td>
</tr>
<tr>
<td>IL1</td>
<td>5.3</td>
<td>5</td>
<td>2.78</td>
<td>0.988</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IL2</td>
<td>5.05</td>
<td>5</td>
<td>2.88</td>
<td>0.987</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: In the three questions regarding behavioral intention, the hypothesis that the mean and median are 5 (p < 0.01) has been rejected.

Table 4. Results of discriminant validity with Fornell and Larker (1981) and HTMT ratios by Henseler et al. (2015) criteria.

<table>
<thead>
<tr>
<th>BI</th>
<th>PE</th>
<th>EE</th>
<th>SI</th>
<th>TRUST</th>
<th>IL</th>
</tr>
</thead>
<tbody>
<tr>
<td>BI</td>
<td>0.906</td>
<td>0.769</td>
<td>0.791</td>
<td>0.782</td>
<td>0.836</td>
</tr>
<tr>
<td>PE</td>
<td>0.724</td>
<td>0.909</td>
<td>0.838</td>
<td>0.685</td>
<td>0.948</td>
</tr>
<tr>
<td>EE</td>
<td>0.723</td>
<td>0.799</td>
<td>0.907</td>
<td>0.75</td>
<td>0.888</td>
</tr>
<tr>
<td>SI</td>
<td>0.713</td>
<td>0.701</td>
<td>0.641</td>
<td>0.934</td>
<td>0.786</td>
</tr>
<tr>
<td>TRUST</td>
<td>0.734</td>
<td>0.841</td>
<td>0.794</td>
<td>0.69</td>
<td>0.863</td>
</tr>
<tr>
<td>IL</td>
<td>0.158</td>
<td>0.183</td>
<td>0.215</td>
<td>0.045</td>
<td>0.246</td>
</tr>
</tbody>
</table>

Note: Principal diagonal is the squared root of AVE. Below the principal diagonal are correlations between variables, and above the principal diagonal are HTMT ratios.

Figure 2 shows that the fit of BI reaches an R² = 65.6%, which, following [85], can be considered substantial. Note that for these values of the determination coefficients, and under the minimum R² criterion, a sample of less than 33 responses would yield significant results [86]. The values of SRMR = 0.054, and thus <0.08, and NFI = 0.85, (i.e., >0.8) also suggest that the model has good adherence to the sample [91].
Table 5. Testing the moderating effects of gender and age over performance expectancy, easiness expectancy, and social influence in the model proposed in Section 2.

<table>
<thead>
<tr>
<th>Moderate effects</th>
<th>$\beta$</th>
<th>Student’s t</th>
</tr>
</thead>
<tbody>
<tr>
<td>$PE \times IL \rightarrow BI$</td>
<td>0.042</td>
<td>0.457 (0.647)</td>
</tr>
<tr>
<td>$EE \times IL \rightarrow BI$</td>
<td>-0.001</td>
<td>0.007 (0.995)</td>
</tr>
<tr>
<td>$PE \times gender \rightarrow BI$</td>
<td>-0.065</td>
<td>0.391 (0.696)</td>
</tr>
<tr>
<td>$EE \times gender \rightarrow BI$</td>
<td>-0.022</td>
<td>0.126 (0.901)</td>
</tr>
<tr>
<td>$SI \times gender \rightarrow BI$</td>
<td>0.102</td>
<td>0.629 (0.529)</td>
</tr>
<tr>
<td>$PE \times age \rightarrow BI$</td>
<td>0.159</td>
<td>0.979 (0.328)</td>
</tr>
<tr>
<td>$EE \times age \rightarrow BI$</td>
<td>-0.187</td>
<td>1.121 (0.262)</td>
</tr>
<tr>
<td>$SI \times age \rightarrow BI$</td>
<td>0.074</td>
<td>0.498 (0.618)</td>
</tr>
</tbody>
</table>

Note: The $p$-value of the coefficient is in parenthesis.

Figure 2 and Table 6 show the values of $\beta$ and their level of statistical significance in the final model adopted in this paper, which leads to accepting hypotheses H2, H3, ($p < 0.01$); and H4 ($p < 0.05$). However, unexpectedly, although the sign of the effect of performance expectancy on behavioral intention is as expected, it was not statistically significant, so hypothesis H1 was rejected. We have also found that the sign of the mediation of insurance literacy on PE and EE was as we expected (positive in the first case and negative in the second) but that mediation was not significant, i.e., H5a and H5b were rejected.
Table 6. Path coefficients and results of testing the hypothesis of Section 2.

<table>
<thead>
<tr>
<th>Effect</th>
<th>( \beta )</th>
<th>Student's t</th>
<th>Decision on the Hypothesis</th>
</tr>
</thead>
<tbody>
<tr>
<td>PE( \rightarrow )BI</td>
<td>0.034</td>
<td>0.407</td>
<td>H1: Reject</td>
</tr>
<tr>
<td>EE( \rightarrow )BI</td>
<td>0.288</td>
<td>3.586 **</td>
<td>H2: Accept</td>
</tr>
<tr>
<td>SI( \rightarrow )BI</td>
<td>0.339</td>
<td>4.764 ***</td>
<td>H3: Accept</td>
</tr>
<tr>
<td>TRUST( \rightarrow )BI</td>
<td>0.23</td>
<td>2.441 **</td>
<td>H4: Accept</td>
</tr>
<tr>
<td>PE \times IL( \rightarrow )BI</td>
<td>0.057</td>
<td>0.635</td>
<td>H5a: Reject</td>
</tr>
<tr>
<td>EE \times IL( \rightarrow )BI</td>
<td>−0.018</td>
<td>0.193</td>
<td>H5b: Reject</td>
</tr>
</tbody>
</table>

Note: *** and ** Stands for statistical significance at 1% and 5% probability level, respectively.

Figure 2 and Table 7 show that the structural equation model estimated with partial least squares has a predictive capacity that is at least acceptable. We found that \( Q^2 > 0 \); therefore, the model provides significant predictions [85]. The use of CVPAT shows that PLS-SEM provides good out-of-sample predictions, as they are more accurate than the benchmark of taking the average of behavioral intention (average loss difference = −2.516, \( p < 0.01 \)). However, we must also acknowledge that the predictions of the presented model do not surpass those of the alternative linear model, so from the perspective of CVPAT, the predictive capacity of our model is not excellent.

Table 7. Predictive capability of the PLS-SEM proposed in Section 2.

<table>
<thead>
<tr>
<th>Predictive Measures</th>
<th>Model vs. Average</th>
<th>Model vs. Linear</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( Q^2 )</td>
<td>RMSE</td>
</tr>
<tr>
<td>BI</td>
<td>0.618</td>
<td>0.627</td>
</tr>
<tr>
<td>Overall model</td>
<td>−2.477</td>
<td>&lt;0.01</td>
</tr>
</tbody>
</table>

RMSE = root mean square error, MAE = mean absolute error and ALD = average loss difference.

5. Discussion and Implications

5.1. Discussion

This article explains the behavioral intention (BI) of policyholders in managing their existing policies with the help of the UTAUT model [20], along with the consideration of the variable trust, which theoretically should have a significant impact on BI, both due to the nature of the insurance business and the relevance that this factor has in the acceptance of chatbots by consumers. The responses obtained denote a widespread resistance to the use of chatbots, and the proposed model allows for a good explanation of the intention to use since it can explain almost 70% of the variance of the sample. Although the utility of the UTAUT model in the adoption of Fintech has been widely demonstrated, empirical analyses in the Insurtech environment are not as frequent.

Regarding RQ1, *what is average intention to be assisted by conversational robots to make procedures such as declaring a claim?*, a clear attitude of rejection toward interaction with chatbots was observed. This finding is consistent with the literature review conducted in [58] and the results reported in the German insurance market [9] and the Canadian market [19].

Regarding RQ2, *does UTAUT, along with trust and insurance literacy, provide a statistically significant framework for the analysis of chatbot acceptance for policy management?*, we can note that the model proposed in the second section satisfactorily explains the intention to use, as shown by the usual indicators such as the coefficient of determination, standardized root mean square residual, and normed fit index. It also presents relevant predictive capacity, since the value \( Q^2 > 0 \) and the cross-validated predictive ability test reveal that the proposed model has a predictive capacity superior to the benchmark of taking the average of the BI items. However, it does not make significantly better predictions than the linear benchmark.
Performance expectancy (PE), as indicated by the literature review conducted, is usually the most relevant variable for explaining the adoption of new technologies. Our results show that, although the impact of PE on BI is positive, it does not have statistical significance. Although uncommon, it should be noted that this result has been previously obtained in fields such as m-banking [92], crowdfunding for funding sustainable technology [93], and the adoption of car telematic devices in an insurance setting [94]. Likewise, in the context of chatbot acceptance, [57,67] also do not report a significant direct impact of PE on behavioral intention.

The obtained results show that effort expectancy (EE) has a significant and positive impact on behavioral intention, which is consistent with the revised literature. Along these lines, we found studies related to blockchain applications in finance [35–38], digital banking [39,42], digital insurance intermediation [44], and chatbot adoption by consumers [22,46–48,52,53,57].

The variable with the greatest influence on behavioral intention is undoubtedly social influence (SI). This result is consistent with the studies reviewed in the context of blockchain [35,36], consumers’ acceptance of conversational robots [10,44,46–48,51,53,55], m-banking [40,64,92], and innovations related to insurance [21,44].

We have verified that trust also has a significant and positive impact on behavioral intention. This was the expected result given the confluence of the intrinsic nature of the insurance business, which is based on trust [21], and the relevance of this construct in the acceptance of robotic technologies [22], which make this factor highly relevant in Insurtechs powered by AI [23]. Our results coincide with [10,22,46,48,57] in the acceptance of conversational bots, with [35,38] in the field of blockchain technology, with [35,38] in the scope of digital banking, and with [44] in the acceptance of new techs in insurance settings.

It is noteworthy that, on the one hand, as argued in Section 2, the moderating effects of gender and age are not significant, contrary to the postulates of the original UTAUT framework [20]. On the other hand, although the sign of the moderation effect of insurance literacy on the link of performance expectancy and effort expectancy with behavioral intention is as we expected, that finding has no statistical significance. This analytical result contradicts the arguments of H5a and H5b.

The responses, on average, report low ratings both in aspects that potentially influence the formation of behavioral intention (usefulness perception) and in those that have shown statistical significance in our sample (effort expectancy, social influence, and trust). We have observed that the most relevant variable for explaining behavioral intention is social influence. This finding could indicate that in 2023, AI-based technologies are viewed with caution due to the potential problems they can bring, which are exacerbated by the lack of regulation [63], and that disruptive technologies often face significant social opposition [62].

The results also indicate that the development of chatbots for managing existing policies, particularly in areas such as claims, is not sufficiently mature. This explains why surveyed policyholders display resistance to be assisted by conversational bots, since they display low ratings for effort expectancy and trust. This result aligns with the findings of [56], who point out that conversational robots can often create more problems than they solve.

The perception of chatbots observed in our sample aligns with [18] in their study on the use of conversational bots in the relationship between Norwegian citizens and public administrations. Chatbots can be useful in providing assistance for simple customer requests, but require human supervision for slightly more complex issues. Therefore, they can be used by the company as a complementary interaction channel that is able to obtain support from insurance customers in very concrete circumstances.

5.2. Theoretical and Practical Implications

From a theoretical standpoint, we have demonstrated that the combination of the UTAUT analytical framework with trust can capture a significant proportion of the behavioral intention to use chatbots. Note that, whereas we have attained an $R^2$ of approximately
65%, studies in which TAM and UTAUT are tested in the original formulation [20,25] attain a coefficient of determination that is rarely above 60%. Moreover, the fact that effort expectancy, social influence, and trust are key variables and that insurance culture is not a relevant factor in the behavioral intention to interact with chatbots provides valuable information for firms about the main factors to take into account when implementing conversational chatbots in policyholder services.

We have found that the sociodemographic variables age and gender are not significant in the acceptance of chatbots for assisting policyholders. This finding supports the mainstream literature in the field of technological acceptance in financial and insurance contexts, as well as the use of conversational robots for consumer assistance, where these variables are often not considered in the analysis.

The evolution of technology continues to reshape the insurance industry, with one key development being the emergence of chatbots [3,6]. While chatbots promise efficiency and lower operational costs [15], enhancing the value of insurers in financial markets [13], there exists an underlying reluctance to adopt them [19], as demonstrated in this paper. Our research illustrates that customers still prefer human interaction, considering it more reliable and personalized. Their hesitance to engage with chatbots reflects a broader distrust in automated customer service, partly due to the current limitations of chatbot technology. These bots may lack the capacity to fully understand complex inquiries or complaints, resulting in suboptimal service. This highlights the importance of effort expectancy in understanding the intention to use chatbots in interactions with the insurer. It is important to note that effort expectancy should not be interpreted as a complication in implementing the interaction channel but rather as an assessment of the communication drawbacks.

However, it is essential not to disregard the practical benefits of chatbots. As technology advances, these AI-based entities will become more sophisticated, potentially surpassing human capabilities in terms of speed, accuracy, and availability. Integrating chatbots can also lead to significant cost savings for companies, which could be redirected to enhance customer experiences. This ambivalence necessitates a delicate balance. Companies should prioritize maintaining high customer satisfaction levels in the short term, which may require human-led customer services. A challenge for the insurance industry is to unequivocally demonstrate that the advantages of using chatbots in policyholders’ service benefit not only firms but also policyholders and insurers. Thus, the relevance of organizational trust, i.e., the belief that firms’ actions are intended to benefit the customer, aligns with this perspective.

Furthermore, it is important to acknowledge that, as consumers encounter chatbots across various services, not just in insurance, their comfort level with this technology is likely to increase. Familiarity can help overcome resistance and improve user acceptance, facilitating a smoother transition toward automated customer service in the insurance sector. Currently, chatbots powered by artificial intelligence are a rising disruptive force that presents several ethical and social concerns in the short term, such as job losses [62], privacy issues, and challenges related to transparency and interpretability [95]. The general social resistance toward emerging disruptive technologies [62], including chatbots, is evident from the high statistical significance of social influence in our empirical application.

This phase of reluctance should be considered an opportunity for the insurance sector to refine their chatbot technology. To achieve this, a closer examination of the customer experience is necessary. By understanding specific customer needs, frustrations, and expectations, companies can tailor their chatbot algorithms to enhance user engagement, increase reliability, and deliver personalized solutions. This approach will make the technology more user-friendly and build trust in the capabilities of chatbots.

In conclusion, the integration of chatbots into the insurance sector is not a matter of “if” but rather “when” and “how”. To navigate this transition smoothly, an iterative approach combining technological advancement with the preservation of human elements of customer service may prove most effective. Studies such as the present one can assist
the insurance sector in leveraging the advantages of chatbots without compromising customer satisfaction.

6. Conclusions, Limitations, and Future Research Directions

This study analyzed the perception of policyholders regarding their interaction with conversational bots when performing procedures related to their in-force policies using voice robots. We observed a widespread reluctance toward interacting with voice robots in this context. To establish the drivers of behavioral intention to use chatbots for procedures related to in-force policies, we used the theoretical framework provided by the well-known technology acceptance model UTAUT [20,24]. Specifically, we evaluated the influence of performance expectancy, effort expectancy, and social influence on behavioral intention. Additionally, we included trust and insurance literacy as variables to evaluate, which are not part of the original four variables of the UTAUT. The results show that our model is useful for explaining the behavioral intention to use bots.

We are aware of the limitations of the empirical work presented. This study was conducted in a specific country, Spain, and most responses were obtained from social networks such as LinkedIn. The users of these networks are often individuals with university degrees and a professional background that typically ranges from middle management to executive positions. Thus, the level of education and economic position may bias our conclusions about the behavioral intention to use chatbots. Therefore, it is advisable to exercise caution when extrapolating our findings to policyholders/insureds from cultures that are not close to the Spanish culture or to individuals with professional and educational profiles different from those of the sample used. To obtain broader conclusions, it would be necessary to expand the number of countries represented and the socioeconomic profiles of the surveyed individuals.

Another limitation of this paper is its exclusive reliance on correlational quantitative methods. However, in order to achieve a deeper understanding of the behavioral intention to use chatbot services in financial settings, it would be beneficial to enhance the study with a configurational approach that can identify profiles associated with acceptance and resistance towards chatbot assistance [96]. Additionally, conducting a qualitative analysis of the arguments provided by the surveyed individuals would also be valuable [97]. Undoubtedly, in this regard, UTAUT offers a robust theoretical framework to support such analyses [24].

The conclusions extracted in this study cannot be generalized to the medium and long term because it is grounded on a cross-sectional survey. Industry 4.0 technology and artificial intelligence are fields in which advances are extremely rapid, and attitudes and intention to use a given technology are highly volatile. A more comprehensive perspective will require similar analyses at different stages of chatbot technology development.

Since late 2022 and early 2023, chat systems based on AI technologies such as natural language processing, exemplified by ChatGPT, have experienced exponential development. Prominent figures in the field of information technology, such as Bill Gates, consider this to be one of the most disruptive technological advancements since the 1980s [96].

Undoubtedly, these technologies offer advantages such as their ability to simulate human conversation, user-friendly interfaces, and the capacity to comprehend the contextual nuances of a conversation with a user based on previous interactions [98,99]. However, they also present disadvantages, including the lack of continuous knowledge updating, as ChatGPT is not connected to databases, and the potential for biased or inaccurate responses to the extent that the training databases contain such biases [99].

The characteristics of this technology suggest that its implementation could enhance the consumer experience [98], which, in the insurance industry, could have implications for both the underwriting phase (e.g., assisting in selecting the most suitable coverage based on consumer profiles) and the management of existing policies, which is the focus of this study. Nevertheless, it is essential to consider the drawbacks resulting from the aforementioned issues. Therefore, a natural extension of this research could involve analyzing the potential
acceptance of ChatGPT-like technologies for decision-making support and management in the insurance domain from the perspectives of both insurers and policyholders.


**Funding:** This research has benefited from the Research Project of the Spanish Ministry of Science and Technology “Sostenibilidad, digitalización e innovación: nuevos retos en el derecho del seguro” (PID2020-117169GB-I00).

**Institutional Review Board Statement:** (1) All participants received detailed written information about the study and procedure; (2) no data directly or indirectly related to the health of the subjects were collected, and therefore the Declaration of Helsinki was not mentioned when informing the subjects; (3) the anonymity of the collected data was ensured at all times; (4) the research received a favorable evaluation from the Ethics Committee of the researchers’ institution (CEIPSA-2022-PR-0005).

**Informed Consent Statement:** All respondents gave permission for the processing of their responses.

**Data Availability Statement:** Data available by requiring it to the authors.

**Conflicts of Interest:** The authors declare no conflict of interest.

**References**

17. de Sá Siqueira, M.A.; Müller, B.C.N.; Bosse, T. When Do We Accept Mistakes from Chatbots? The Impact of Human-Like Communication on User Experience in Chatbots That Make Mistakes. Int. J. Hum.–Comput. Interact. 2023. [CrossRef]


25. Davis, F.D. Perceived usefulness, perceived ease of use, and user acceptance of technology. MIS Q. 1989, 13, 319–340. [CrossRef]


56. Xing, X.; Song, M.; Duan, Y.; Mou, J. Effects of different service failure types and recovery strategies on the customer response mechanism of chatbots. *Technol. Soc. 2022, 70, 102049.* [CrossRef]


64. Tomić, N.; Kalinić, Ž.; Todorović, V. Using the UTAUT model to analyze user intention to accept electronic payment systems in Serbia. *Port. Econ. J. 2023, 22, 251–270.* [CrossRef]


