Exploring how security and privacy affect IoT device purchase behaviour

By

N.D. Ho-Sam-Sooi

Student number: 4244629

in partial fulfilment of the requirements for the degree of

Master of Science
in Complex Systems Engineering and Management

at the Delft University of Technology,
to be defended publicly on Friday January 10, 2019 at 15:00 PM.

First Supervisor:	Dr. ir. W. Pieters, TU Delft
Second supervisor:	Dr. ir. M. Kroesen, TU Delft

Preface

Amsterdam, 27 December 2019

Dear reader,

In the context of the MSc Complex Systems Engineering and Management at Delft University of Technology, I have written this report to highlight the findings of the research that I have been conducting the last 6 months. The report is targeted at anyone interested in the field of cyber security and privacy. More specifically, my research has targeted behavioural aspects of security and privacy of IoT devices. As the field of IoT security and privacy field is mainly dominated by highly technical contributions, my aim was to take a first step towards including the consumer in the scientific discussion.

I am very grateful to PwC for granting me the opportunity to graduate at their firm and at the same time gaining an insight into how my knowledge can be put into practice. I would like to thank Sanne Amber Maas for her positive mindset and support during this period.

My thesis committee has also been a great support during my research. I would like to thank Wolter Pieters for his contributions in conceptual discussions, scoping and his critical view, which enabled me to vastly improve the quality of my research and reporting. I would like to thank Maarten Kroesen for his assistance in carrying out the stated choice experiment and analysing and interpreting the collected data. Being able to spread my survey via his BSc course enabled me to complete the data collection at a very early stage

Executive summary

The IoT concept is characterised by physical objects that are connected via internet connectivity in order to provide innovative functionalities to their end users. The market penetration and societal acceptance of IoT devices is ever-increasing, as more use cases are introduced and the affordability of the devices improves. IoT devices improve the quality of life for consumers by providing new and innovative functionalities. For example, smart thermostats enable consumers to remotely configure the heating in their home and in some cases remove the need for manual adjustments of the heating system entirely. Although such use cases are highly beneficial for consumers, the widescale adoption of IoT devices also introduces significant risks with regard to privacy and security. In many cases, the IoT devices lack basic security controls such as encryption or authentication schemes. If adversaries are able to gain access to the device or information that is stored on the device, they can harm the confidentiality, integrity or availability of highly sensitive information or even inflict physical harm. In addition, IoT devices often collect large volumes of highly sensitive data. Manufacturers often share this information with third parties or use it for the improvement of their services, thus harming the privacy of consumers.

The existing body of literature in the IoT security and privacy field is strongly focused on the analysis and design of technical measures that mitigate the privacy and security risks of IoT devices. However, for such measures to be successful, it is crucial that consumers buy IoT devices with sufficient protection mechanisms and opt for manufacturers that safeguard their privacy. Thus, it can be argued that nudging consumers towards buying more secure devices and taking privacy into account when purchasing devices is crucial to ensure the safety of consumers. Due to limited incentives for market parties, it seems sensible that governmental bodies should take an active role. However, this requires knowledge regarding the choice behaviour of consumers when purchasing IoT devices. More specifically, it should be clear how privacy and security influence the purchase behaviour of consumers. This study aims to generate this knowledge by answering the following research question:

""How do security and privacy influence the choice of consumers to buy an IoT device? And how sensitive is the effect of security and privacy to framing and personal factors?"

The study takes a mixed methods approach towards answering this research question. The quantitative part of the study focuses on the effect of security and assesses how security affects the choices of respondents in a stated choice experiment in comparison to other device attributes, such as the functionalities and price of the device. The stated choice experiment has been conducted by means of a survey that was spread by a group of BSc students at the faculty of Technology, Policy and Management in the context of a data analytics course. A total of 510 valid responses were provided to the survey. In the survey, the respondents were faced with a choice set containing two options for a smart thermostat, in the hypothetical scenario that their (smart) thermostat had broken and they were faced with the decision for a new smart thermostat. Per choice set, the respondents were asked to assess whether they would purchase either one of the smart thermostats and which of the two smart thermostats would have their preference if they had to make a choice between the two options. The smart thermostats varied on three attributes: Price, functionality and security.

Moreover, the respondents were randomly divided into two groups. For the first group, the security attribute was framed in terms of gains, while the description of the security attribute focused on losses in for the second group. Additionally, the survey contained a set of indicators that aimed to measure the values of personal factors that might influence the effect of security on choice behaviour. The following factors were identified: Privacy/Security Consciousness,

Technology Acceptance and Conservativeness. The Privacy/Security Consciousness factor measures to what extent a respondent is aware of privacy/security risks. Additionally, the factor takes privacy/security concerns and actions to mitigate the risks into account. Secondly, the Technology Acceptance factor evaluates whether a respondent is willing to make use of the newest innovative technologies. Finally, the conservativeness factor provides an indication of a respondent's perception of the value of innovation for society. Respondents who do not value innovation strongly score high on this factor.

From the data that has been collect from the stated choice experiment, Multinomial Logit (MNL) models were developed in order to quantify the effect of three device attributes: Price, Functionality and Security and assess whether personal factors or framing moderate these effects. The final model is constructed from the responses to the question that asked the respondents whether they would purchase a specific smart thermostat. The resulting causal model, including parameter estimates, is displayed below.

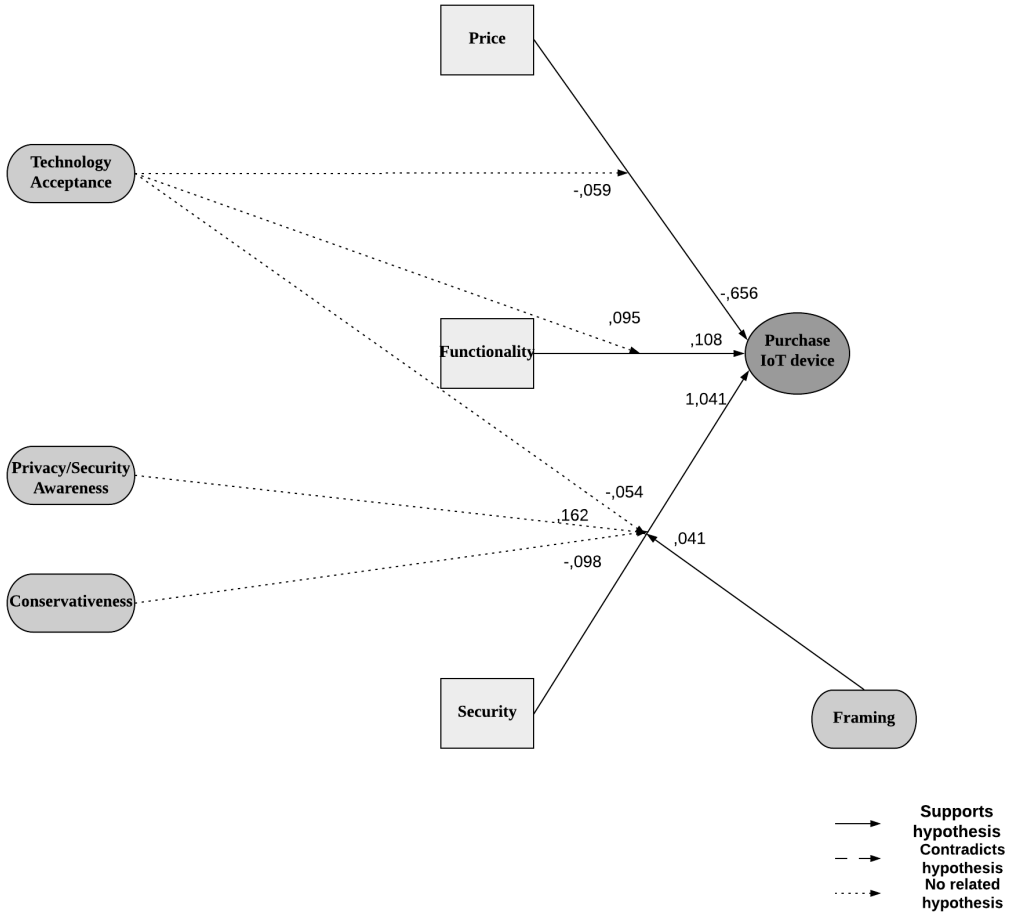


Figure 1.1: Causal model

In line with the hypotheses, the functionality and security have a positive influence on the utility of an alternative, while the price attribute negatively affects the utility of an alternative. The effect of security is exceptionally strong in comparison to other device attributes, which implies that security has an exceptionally strong effect on the purchase decision.

In addition, the results show that the description of security that focuses on gains is more effective in nudging the respondents towards buying more secure devices. This finding is in

line with Prospect Theory, which postulates that people are more risk averse when faced with possible gains. With regard to the personal factors, five interactions were found with the device attributes. The technology acceptance factor negatively moderates the effect of the price and security attribute and positively moderates the effect of the functionality attribute. This implies that respondents with a high score on this factor are willing to make concessions on price and security in order to make use of devices that provide them with innovative functionalities. The Privacy/Security Consciousness factor positively moderates the effect of security, which indicates that security contributes more strongly to the value of an alternative for respondents who are more aware of privacy and security risks of smart thermostats and act upon their knowledge. Finally, the role of security in the decision-making process is relatively small for respondents with a high score on the conservativeness factor. A possible explanation for this result is that respondents who do not value innovation strongly possess less technical knowledge and therefore are less likely to take security into account when purchasing devices.

The qualitative study aimed at revealing the underlying rationales that determine the effect of security and privacy of the choice behaviour of consumers. In the context of this study, 27 responses were provided to a web-based survey. In this survey, the respondents were asked questions regarding the role of security and privacy in their decision to buy or not to buy a smart thermostat. Surprisingly, security and privacy were only once mentioned as a reason to buy or not to buy a smart thermostat. The respondents were also asked for their awareness of security or privacy related risks relate smart thermostats. Although their responses strongly lacked detail, they were able to mention some high-level risks using general terms such as “hacking” or “information leaks”. Lastly, the respondents were asked to rate the severity of several risks that were presented to them via a set of hypothetical scenarios. From the results, five factors have been identified that influence the risk assessment process of the respondents: perception of security/privacy level, probability of occurrence, third party benefits and impact.

The results of both studies are to some extent contradictory. The quantitative study concluded that consumers take security strongly into account when purchasing devices. However, security and privacy were mentioned only once as a reason to purchase or not to purchase a smart thermostat in the qualitative study. A possible explanation for this discrepancy is that information regarding security in the stated choice experiment is available and presented in such a way that it is easy to compare the alternatives in the choice sets with regard to security. Moreover, the respondents in the quantitative study were actively triggered to think about security by including it as an attribute in the stated choice experiment. In the qualitative study, the respondents were asked an open question that did not specifically mention privacy and security. Even after actively being triggered to think about the importance of security and privacy in their purchase decision, many respondents indicated that privacy and security did not play a significant role.

The results of this study suggests that security and privacy can have a strong effect in the case that security and privacy related information is available and communicated in a simple manner that allows for comparison of devices. This provides support for the argument that governmental bodies could nudge consumers towards buying more secure devices and taking privacy into account by defining standards or legislation that define what information should be communicated and how this information should be communicated. As IoT security and privacy is a highly complex topic, it is advised to include market parties, such as manufacturers and retailers, into the development process of such legislations or standards. In addition, the quantitative study illustrated that improving the risk awareness of consumers can assist in nudging them towards buying more secure devices and taking privacy into account when purchasing devices. Finally, the qualitative study found four possible factors that influence the

risk perception of consumers: Perceived security and privacy level, probability of occurrence, third party benefits and impact. These factors could form the basis of risk awareness efforts.

Further research could build upon this study by evaluating the effects of other device attributes, such as privacy, ease of use, cost reduction or compatibility with other devices on the purchase behavior of consumers. This creates a more comprehensive overview of how the effect of security compares to the effect of other device attributes and shows whether similar conclusions hold for the effect of privacy on purchase behavior. Secondly, future efforts could opt for a different operationalisation of the security attribute in the stated choice experiment to assess which operationalisation is most effective in nudging consumers towards buying more secure devices. Thirdly, real-world choice data can be used as input for choice models to evaluate whether the choice behavior in the case of real-world choices resembles the choice behavior in stated choice experiments.

Table of contents

List of figures

List of tables

Chapter 1 Introduction

Security is among the most significant challenges surrounding the development of innovative digital technologies. To illustrate this, a recent report by PwC has shown that the number of cyber-attacks on a global scale has risen enormously in recent years (PwC, 2018). These cyber-attacks have resulted in a significant amount of costs for organisations in various sectors. The costs of cyber-attacks for the Dutch government and the largest businesses located in the Netherlands are estimated at €10 billion per year in 2017 (Deloitte, 2017). These cyber-attacks also affect consumers. A survey held among 1000 American citizens highlighted that 69% of consumers believe that businesses are vulnerable to hacks and cyber-attacks (PwC, 2017). Moreover, the study revealed that only 25% of consumers believe that businesses handle their personal information responsibly.

An increasing number of these cyber-attacks are targeted at Internet of Things (IoT) devices (Netscout, 2019). This is exceptionally problematic, since the market penetration of IoT devices is expected to rise rapidly in coming years, due to the improved affordability of the devices and the increased amount of use cases. This trend is supported by the development of the fifth generation of mobile networks (5G). This network technology could lower the latency of network connections as well as support enormous increases in data traffic over IoT networks. In this manner, 5G network technology would be able to vastly improve the quality of IoT networks and introduce new use cases for IoT devices. The increased amount of use cases introduces new opportunities for malicious parties to target IoT devices.

Consumers are expected to reap the benefits from IoT devices. To illustrate this, smart home technology can provide an extensive improvement of the quality of life for consumers. In smart homes, everyday objects are connected to the internet as well as other devices on the network in order to provide innovative functionalities to the occupants. For example, smart thermostats are able to detect whether the occupant is present and adjust the heating schedule of the home accordingly.

Although IoT devices improve the quality of life for consumers, they also introduce severe security and privacy risks. Firstly, the devices that are connected to the IoT often lack computational power. This makes the implementation of complex encryption protocols problematic. Moreover, the diversity of communication protocols on IoT networks gives rise to vulnerabilities which can be exploited by malicious third parties. Thirdly, passwords are often excessively simple or hardcoded into firmware, which makes it simple for malicious third parties to gain control over the device or access sensitive information. The lacking security of the devices also enables attackers to gain access to the home network of the occupants. This poses the risk of an attacker gaining access to other confidential information on the network. Privacy is also an important topic for IoT devices. IoT devices often collect a large amount of data in order to provide services to their users. In many cases, highly sensitive data is collected regarding the users of the devices. For example, smart thermostats collect data regarding energy use and the occupancy of the home. This poses the risk of privacy infringements when such information is shared with third parties.

Consumers can contribute to the security of their devices and their privacy when using the devices. In the case of smart homes, purchasing smart home devices with more sophisticated security controls strongly improves the security of their smart home system and their network as a whole. However, it seems sensible that a large part of consumers does not possess any knowledge related to security and are therefore not aware of the security of their devices. Even for consumers with a strong technical background, protecting the devices and against cyber-

attacks would consume a large amount of time and effort. This causes a disincentive for producers to invest into the security of their devices, since it can be expected that consumers who lack security awareness do not take security into account when purchasing devices. On the contrary, producers are inclined to focus their efforts on improving the functionality, ease of use, price, for the reason that these attributes are more easily understandable for consumers and are therefore expected to have a higher impact on the purchase decision of the consumer. Empirical evidence supports this argument, as many IoT devices on the market lack in crucial security controls, such as encryption or authentication schemes. With regard to privacy, consumers can evaluate the privacy notices of their devices to assess to what extent their privacy is safeguarded properly. However, privacy notices are often lengthy and complex, which makes it complicated for consumers to assess the level of privacy they can expect in a timely manner.

On the other side, it can be argued that producers are motivated to communicate the security and privacy related information to consumers, thereby stimulating sales under the assumption that consumers do value security and privacy. However, it can be argued that communicating security controls or risks triggers consumers to think about the risks of the device, thus lowering the probability of a consumer buying the device. Moreover, it is unclear which aspects security and privacy are most effective in improving the probability that a consumer purchases a certain device and how the security level should be framed in order to improve the attractiveness of a device.

Therefore, the question arises whether governmental bodies, such as the Ministry of Justice and Safety, should provide incentives for producers to communicate information regarding security and privacy towards consumers. One of the main goals of this organisation is to ensure the safety of its citizens. Cybersecurity plays a large role in reaching this goal, as the increasing impact of cyber-attacks puts the safety of citizens at risk. From a governmental perspective, such incentives would contribute to the resilience of cyberspace as well as the safety of the citizens who operate in cyberspace. Thus, one could argue that governmental bodies could play a significant role in stimulating consumers to buy secure products and taking privacy into account when purchasing devices. However, influencing consumer behaviour in such a way requires broad and detailed insights into the decision-making process of consumers when purchasing IoT devices.

First of all, it is important to know whether consumers actually take security and privacy into consideration when making a purchase decision. Acquiring this knowledge would allow us to assess whether the behaviour of consumers can be influenced by communicating security and privacy related information. Additionally, it is key to assess the relative importance of security and privacy when compared to other attributes such as functionality or price. This provides an insight into the trade-offs that consumers make when purchasing IoT devices. Finally, the way in which security or privacy are framed is also expected to influence the decision-making process of consumers. To illustrate this, a positively framed message regarding the security level of the device might trigger more behavioural changes than a negatively framed message. Therefore, it is desirable to examine which frames are most effective in nudging consumers towards buying more secure devices and taking privacy into account.

The main objective of this study is to provide governmental bodies and other entities with initial policy recommendations for nudging consumers towards buying more secure devices and taking privacy into account when purchasing devices. In order to reach this goal, the study aims to research the decision-making process of consumers when purchasing IoT devices. The insights which have been gained from this analysis will be used to develop the policy recommendations.

The study is societally relevant in the way it speaks to the issue of IoT security. By pursuing the abovementioned objectives, the study aims to assist governmental bodies in nudging consumers towards buying more secure devices and taking privacy into account when purchasing devices. If governmental bodies work towards this goal successfully, they are able to battle the ever-increasing costs of cybercrime as well as ameliorate the safety of their citizens.

The effect on security and privacy on choice behaviour has been investigated by researchers in the Technology Acceptance Modelling (TAM) field. The studies in this field concluded that the security of an online service has an influence on the choice of end-users to buy such devices or services. However, it is unclear if this is also the case for IoT devices. Furthermore, the existing studies only observe the result of one choice task per respondent. Presenting the respondent with a set of multiple choices allows the researcher to vary device attributes in order to gain a deeper insight into the relationships between the attributes and the choice behaviour. Moreover, existing research has not considered interaction effects between device attributes, framing and personal factors of the respondents, such as attitudes, opinions, and beliefs. Including the interactions with personal factors allows for the assessment of discrepancies in the effect of security and privacy between subgroups of consumers. The sensitivity of the effects to framing shows whether some frames are more effective in nudging consumers towards buying more secure devices and taking privacy into account when purchasing devices.

This study aims to fill these knowledge gaps by answering the following research question:

RQ: *"How do security and privacy influence the choice of consumers to buy an IoT device? And how sensitive is the effect of security and privacy to framing and personal factors?"*

The main research question can be decomposed into the following sub-questions:

RQ1: "What are the most prevalent issues in IoT security and privacy?"

RQ2: What are the most important frames in message framing literature? And how can these be applied to IoT security and privacy?"

RQ3: "Which hypotheses regarding the effect of device attributes and framing on the choice of consumers to purchase can be drawn from literature?

RQ4: "How does the security level of IoT influence the choice of consumers to buy a certain device, when comparing to other device attributes?"

RQ5: "To what extent is the influence of the security level of IoT devices on the choice of consumers to buy a certain device sensitive to framing and personal factors?"

RQ6: "What underlying rationales determine how security and privacy affect the decision of consumers to buy an IoT device?

The study aims to generate policy recommendations by analysing the decision-making process of consumers when buying IoT devices. In order to do so, complete and detailed insights regarding this decision-making process should be attained. Firstly, the effect of security and privacy on the decisions of consumers should be investigated. Additionally, the underlying rationales that determine how security and privacy affect the decisions should be analysed. A quantitative study is most suited to study the effects of security on the decisions, as this approach allows us to validly demonstrate these effects and compare them to effects of other device attributes. On the contrary, the analysis of the underlying rationales that determine these effects asks for a qualitative approach. A qualitative approach is more suited to gain more in

depth insights regarding the decision-making process. Thus, a mixed methods approach is most suited to reach the main objective of the study.

The Complex Systems Engineering and Management programme challenges its students to analyse complex systems or processes with a strong technical component and design technological, institutional or process interventions to improve these systems or processes on multiple facets. For this study, the complex system or process takes the form of a decision-making process in the market for IoT devices. The technical component of this system lies in the field of IoT security. Research in this field focuses strongly on the technical security controls and risk mitigation measures regarding IoT devices. From the analysis, insights are gained that are used to develop a set of policy recommendations for governmental bodies. These policy recommendations are aimed at improving the market for IoT devices by nudging consumers towards buying more secure devices and taking their privacy into account when purchasing the devices. In this manner, the study analyses a complex system with a strong technical component and provides an initial institutional design to improve the system in the form of policy recommendations. Thus, this study provides a good fit with the Complex Systems Engineering and Management programme.

The study will reach its objective by following a clear structure. The figure below describes the outline of the research in various phases. For each phase, the figure depicts which sub-questions are answered as well as which deliverable has to be developed during the phase. In the first phase, a literature study is conducted. The literature study aims to develop an oversight of the current knowledge on the research topic. These insights are used to identify the main knowledge gaps in the existing body of knowledge and develop hypotheses for the quantitative study. The second phase is comprised of a quantitative study. The main goal of this phase is to determine the effect of security on purchase behaviour and assess whether these effects are sensitive to personal factors and framing. The final phase of the study consists of a qualitative study, in which the decision-making process of consumers when buying an IoT device is evaluated in depth. More specifically, this phase is targeted at examining what underlying rationales determine how security and privacy affect the purchase behaviour.

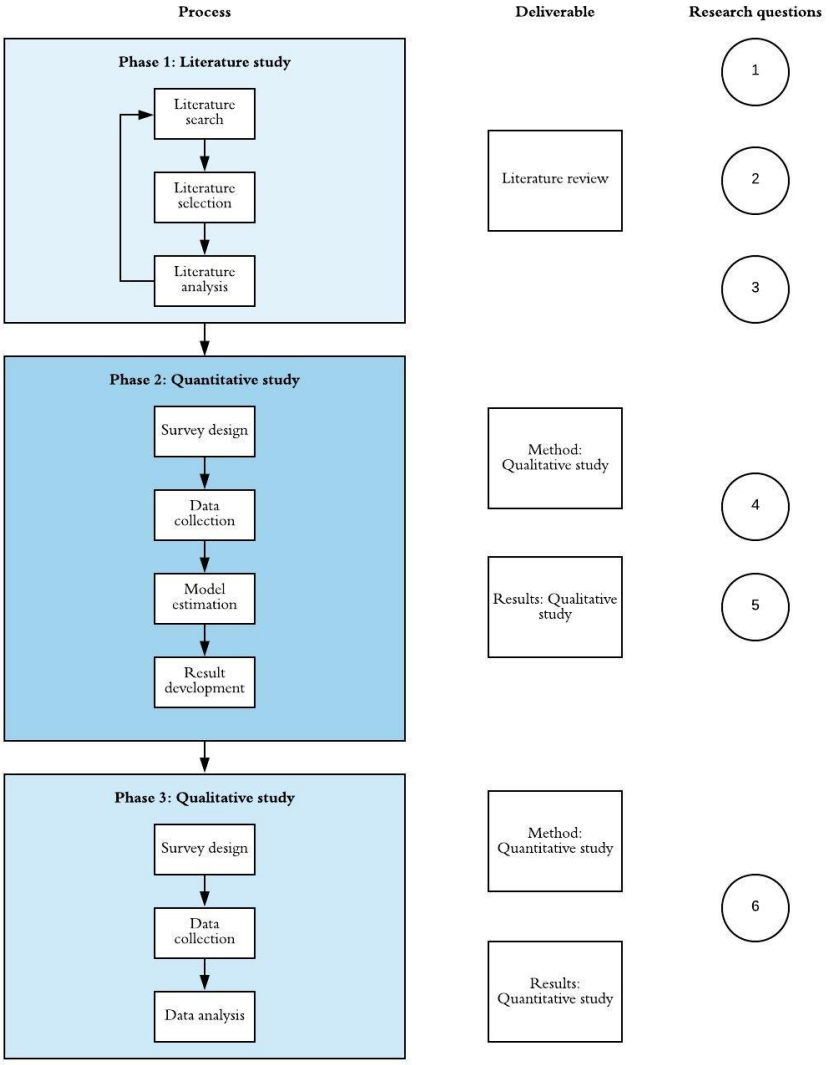


Figure 1.1: Research outline

The remainder of this thesis is divided into five chapters. In chapter 2, the results from the literature review are presented. The literature review highlights knowledge gaps in the existing body of literature and is used to construct a set of hypotheses that form the basis of the quantitative study. Chapter 3 contains a detailed description of the methodologies of the qualitative and quantitative study. The results of the quantitative study are presented in chapter 4. Subsequently, the results of the qualitative study are discussed in chapter 5. In chapter 6, the results from both studies will be combined to draw conclusions that answer the research question. Moreover, the chapter contains the discussion of the results, including the scientific and practical implications, limitations and possibilities for further research.

Chapter 2 Background

In this chapter, the existing literature regarding the subject of this study is discussed. Firstly, a general overview of the IoT security and privacy field is be presented to highlight a significant knowledge gap that can be filled by means of the main research question of this study. Subsequently, the conceptual frameworks for both the quantitative and the qualitative part of the study are constructed.

2.1 IoT Security & privacy

The Internet of Things (IoT) is a concept characterised by a network of connected heterogeneous objects or systems (Singh & Kapoor, 2017). Each of these objects, often including everyday objects such as common house holding equipment, is connected to the network via internet connections. The objects in IoT systems engage in interactions to serve a certain goal, such as to provide a service to a user group (Sicari, Rizzardi, Grieco, & Coen-Porisini, 2015). In order accomplish this, the objects collect, analyse and communicate data (Arias, Wurm, Hoang, & Jin, 2015). In many cases, the objects are able to actuate in the real world when they receive information which states that actuation is needed (Gubbi, Buyya, Marusic, & Palaniswami, 2013).

The main goal of this section is to describe the issues surrounding IoT security and privacy. In order to reach this goal, the terms privacy and security should first be conceptualised. Security from a cyber perspective is often defined by the concept of information security. Information security is most commonly defined by three general security requirements (Cherdantseva & Hilton, 2013):

- Confidentiality: Assuring that personal information is not accessible by non-authorised entities

- Integrity: Assuring that the accuracy, completeness and trustworthiness of personal information is maintained.

- Availability: Assuring that personal information is available to authorised entities at all times

Although this model of information security is widely used, scholars have challenged the model in recent years by stating that cybersecurity and information security cannot be used interchangeably, while the two concepts do have some strong overlaps (Von Solms & Van Niekerk, 2013). Von Solms & Van Niekerk argue that cybersecurity is not only about securing the cyberspace itself, but also the entities which function via cyberspace and the assets which can be reached through it.

A distinction should be made between corporate- and product security. The main goal of corporate security is to keep an organisation secure by protecting its primary resources and secrets, while product security is defined as the security of products developed by organisations (SANS, 2013). To illustrate this, product security controls protect the product itself against harmful adversaries. Corporate security controls protect the organisation itself against attacks. This study researches the effect of privacy and security on the purchase decision of consumers. Product security is more relevant for consumers, since it more directly affects their protection against attacks.

As IoT devices often collect a large amount of highly sensitive personal information, privacy is an important topic for IoT devices. Finding an agreed-upon general definition for privacy is problematic, since the concept is used in multiple fields to describe a variety of topics. A well-known privacy definition has been given by Westin (1968), who defined privacy as "The claim of individuals, groups, or institutions, to determine for themselves when, how and to what extent information about them is communicated to others". Gurtwirth (2002) gave a similar definition, stating that privacy can be seen as "the control over information about oneself". Although these definitions seem fair at first inspection, they miss out on privacy-related topics that have arisen during the early stages of the information age. Burgoon (1982) addressed this issue by identifying the following categories of privacy:

1. Informational privacy: Control of entities over personal information.

2. Interactional privacy: Control of entities over interactions with other entities.

3. Psychological privacy: Control over when and with whom personal information is shared.

4. Physical privacy: The right to a physical personal space.

This section has provided an extensive conceptual description of security and privacy. The conceptualisation of privacy and security forms the basis of the following sections, in which the main privacy and security related issues of IoT devices are described.

2.1.1 Security controls

Security is amongst the largest factors that limit the development and societal acceptation of IoT technology. One of the driving factors for these issues is the lack of security controls of IoT devices. Firstly, many low resource devices are being connected to IoT networks. These devices are limited in their memory, power use, battery size, and computational power. This makes the implementation of computationally demanding or mathematically complex AES, 3DES, RSA and elliptic curve problematic (Zhao & Ge, 2013). As a resolution for this issue, several lightweight encryption algorithms have been developed that are especially suited for devices with limited resources. However, the strength and ease of use of these algorithms still have to improve significantly (Singh, Sharma, Moon, & Park, 2017). Moreover, recent contributions have put forward Cipthertext-Policy Attribute Based Encryption (CP-ABE) as a solution to manage access control for a large number of IoT devices. CP-ABE allows the encryption of data with an access policy that only allows users that have been authorised and are in line with a set of predefined attributes to unencrypt the data (Oualha & Nguyen, 2017; Odelu, Das, Khan, Choo & Jo, 2017). The authors recognise some significant issues for the implementation of CP-ABE for IoT devices, such as costly bilinear maps, long decryption keys, cyphertexts, and significant computation costs and provide initial solutions to overcome these hurdles.

In many cases, the passwords which protect IoT devices lack strength, which makes them an easy target for brute force attacks (Pasha, Shah, & Pasha, 2016). The lack of password strength may be caused by the inability of end-users to change the credentials of the device. Moreover, it can be expected that end users who are not aware of the security risks of their devices prefer a weak password which is relatively easy to understand over a more complex and secure password. There have also been reports of cases where the firmware of the devices contained

a backdoor, which allowed malicious third parties to control the device and access the information stored on it (Xie, Jiang, Tang, Ding, & Gao, 2017).

Thirdly, patching is one of the largest issues for the security of IoT devices (OWASP, 2018). If the software on an IoT device is not patched frequently, it is very simple for attackers to exploit known vulnerabilities in the older versions of the software. Often, it is very difficult to update the firmware on the devices, allowing attackers to make use of these vulnerabilities to attack the devices.

Finally, access control poses a significant security challenge for IoT networks (Mahmoud, Yousuf, Aloul, & Zualkernan, 2015). In IoT networks, objects should be able to continuously identify and authenticate other objects on the network. Given the heterogeneity of the objects on the IoT networks, this process is often highly complex. Thus, sophisticated access control schemes should be put in place on the device level. The absence of such controls enables attackers to gain unwarranted access to devices or information which is stored on those devices.

2.1.2 Attack types

Lacking security controls adversaries to target the IoT devices with physical- or cyber-attacks. Security researchers have developed an understanding of which attack strategies compose the largest threats for IoT devices and networks. In a Distributed Denial of Service (DDoS) attack, The attacker attempts to limit the availability of the system by overloading it with traffic from a large amount of sources (Abomhara & Køien, 2014; Wang et al., 2016; Mineraud, Mazhelis, Su, & Tarkoma, 2016; Roman, Zhou, & Lopez, 2013). Because of such an attack, legitimate users may not be able to use the service or access their data.

Eavesdropping occurs when a malicious party is able to intercept communication between legitimate parties. In a man-in-the-middle attack, the attacker is able to read, alter and insert communication between parties on the system. During such an attack, the legitimate parties think they are directly communicating with each other, while in reality the attacker is in control over the communication channel (Abomhara & Køien, 2014; Bohli, Langendörfer, & Skarmeta, 2013; Mineraud et al., 2016; Sadeghi, Wachsmann, & Waidner, 2015).

In a phishing attack, the attacker attempts to attain personal information from legitimate users by disguising as a trustworthy organisation in digital communications. If legitimate users believe they are actually communicating with a trusted organisation, they might provide them with sensitive personal information, such as usernames or passwords (Sadeghi et al., 2015).

Node capture allows the attacker to access the data at the node as well as cloning and deploying malicious nodes in the network. Node capture is a significant risk for IoT systems, since a large amount of physical nodes exist in the network which are accessible to anyone (Roman et al., 2013; Arias et al., 2015)

2.1.3 Consequences

If the adversaries are able to execute one of the abovementioned attack strategies successfully, the consequences for the owner of the IoT device or network are substantial. IoT devices are physical objects that have been connected to a network via internet connectivity. Thus, if an attack targeting an IoT device is successful, this could have severe consequences in the physical world. For example, if a smart traffic system has been compromised and is now under the control of a cyber-terrorist group, the safety of the citizens participating in traffic is at stake.

This example suggests that the consequences of cyber-attacks on IoT systems are not only limited to cyberspace, but also pose significant risks for harm in the physical world.

The data collected by the devices also introduce severe risks. The various connected objects collect a large amount of data in order to provide functionalities to its end users. For example, smart home devices collect a large amount of information regarding the occupant of the smart home. Similarly, IoT devices in the medical world collect highly sensitive medical information. If the security of these devices is lacking, malicious third parties are able to access the information which is stored on the devices as well as messages between devices on the network.

Producers of IoT devices collect a vast amount of data from the devices. This data is used for a wide variety of purposes, such as the improvement of the devices or services the producer offers to its customers. To illustrate this, Google and Amazon have openly admitted that they hired contractors to annotate recordings collected by virtual home assistants (CNET, 2019). Producers might share the information they collect regarding the device owner with external parties, without the knowledge or approval of the device owner. In some cases, this information might even be used against the device owner in legal cases. Thus, the sharing of highly sensitive personal information by the producer or other parties might inflict harm upon the device owner.

To conclude, the IoT security and privacy literature has primarily focused on discovering the main risks, attack strategies, and consequences related to the security and privacy of IoT devices However, little is known about the perception of consumers on these issues. More specifically, it is not clear how security and privacy affect the decision of consumers to buy IoT devices. The main research question of this study targets this exact research domain. In the remainder of this chapter, the conceptual framework of both the quantitative and the qualitative part of this study is constructed.

2.2 Conceptual framework quantitative study

In this section, related work is evaluated in order to construct the conceptual framework for the quantitative part of the study. Per research subject, the main knowledge gaps are identified. These knowledge gaps support the need for the answering of the research questions of this study. In addition, hypotheses are postulated that describe the relations between the concepts in the framework. For methodological reasons, the framework for the quantitative study focuses on the effect of security on choice behaviour, thus excluding the effect of privacy. This decision is motivated in section 3.2.4.

2.2.1 Security & Technology Acceptance

IT security researchers have not yet investigated to what extent security and privacy influence the choice of consumers to buy a certain IoT device. However, researchers have developed an understanding of how various factors, such as price, ease of use and usefulness affect the acceptance of other innovative products or services, such as online banking. The basis of this field, commonly known as Technology Acceptance Modelling (TAM) has been formed by Davis (1989), who concluded that there exist clear relationships among ease of use, usefulness and acceptance of innovative technologies. Davis defined acceptance as the usage of a technology or system by its end users.

In the following years, IT researchers have extended this model by adding perceived security, risk and trust-related factors and applying it to digital products. For example, Gu, Lee & Suh (2009) applied TAM to mobile banking. From this study, the authors concluded that trust, ease of use and the acceptance of mobile banking are closely interrelated. Furthermore, a study by

Salisbury, Pearson, Pearson & Miller (2001) evaluated which factors affect the willingness to engage in web-based shopping. The results of this study showed that Web security perception plays a large role in determining purchase intent. Even more, it has a stronger effect than ease of use and usefulness of technology. The authors defined Web security perception as "the extent to which one believes that the Web is secure for transmitting sensitive information" (Salisbury, Pearson, Pearson & Miller, 2001, p.3). Their measurement of this concept did not take into account any framing effects. On the contrary, positive and negative frames were used additively to determine the security perception of respondents. In line with this thinking, a study by Crespo, del Bosque & de Los Salmones Sanches (2009) has led to the conclusion that various risk factors such as security, strongly limit the acceptance of e-commerce. The researchers framed the risk factors as potential losses, thus negating the possible effect of framing in the communication of these risks.

Knowledge gaps

The research in the TAM field suggests that clear relationships exist between the perceived security and the acceptance of innovative technologies or services. The researchers in this field set their experiments up in such a way that every decision-maker only faces one specific decision. This limits the quality of the estimated models, as increasing the amount of decisions made by each subject increases the validity and precision of the models. Presenting the individuals with a single decision also limits the suitability of the model for drawing conclusions about the trade-offs between the various attributes of the devices.

Moreover, researchers have not investigated whether the effect of security differs within various subgroups of consumers. It is highly likely that the effect of security and privacy is sensitive to certain personal factors that describe such subgroups. For example, consumers who have a strong sense of awareness with regard to security and privacy are expected to take security and privacy more strongly into account when purchasing devices. Finally, the researchers have not assessed whether the effect of security on the choices of consumers is sensitive to framing. The manner in which security is framed could have a significant effect on the decisions made by consumers. In section 2.3, the concept of framing is defined. Moreover, an overview is presented of the various frames which have been applied in the message framing field.

Hypotheses

The results of studies in the TAM field show that the functionality, price and perceived security of online services influences the choice of individuals to use the services. More specifically, if a product is perceived to be more secure, the probability that an individual uses it increases. At the time these studies were published, these online services were still innovative and not yet fully accepted by society. In this manner, they resemble the current position of IoT devices. Moreover, it is assumed that there exists a clear interrelation between the actual security and perceived security of a product or service. This leads to the following hypotheses:

H1: The price of an IoT device negatively influences the probability that the device is purchased.

H2: The number of functionalities of an IoT device positively influences the probability that the device is purchased.

H3: The security level of an IoT device positively influences the probability that the device is purchased.

2.2.2 Framing

The previous section has highlighted a significant knowledge gap. It is still unclear to what extent the effect of security on the choices of consumers is sensitive to framing. Entman (1993) defined framing as "the selection of some aspects of a perceived reality and making them more salient in a communicating text, in such a way as to promote a particular problem definition, causal interpretation, moral evaluation, and/or treatment recommendation for the item described" (p. 2). Moreover, according to Entman, frames describe problems, diagnose causes, make moral judgements and select the most suited remedies. Chong & Druckman (2002) provide a more high-level definition of framing, defining the concept as "the process by which people develop a particular conceptualisation of an issue or reorient their thinking about an issue".

Gain/loss

Gain/loss framing is one of the most prevalent frames in message framing literature. In the gain frame, the message focuses on the gains the decision-maker can acquire when opting for a certain alternative. On the contrary, the loss frame communicates the possible losses of an outcome situation. According to Prospect Theory, people tend to be risk-averse when being presented with sure gains and risk-seeking when facing sure losses (Kahneman & Tversky, 2013). This goes against classical utility theory, in which similar outcomes provide the same amount of value to the decision-maker. Kahneman & Tversky developed a different choice model, in which value is attained from gains and losses rather than net outcomes and the probabilities in the utility function are replaced by decision weights.

Researchers in the medical field have applied the concept of gain/loss framing in order to assess effect of gain/loss framing on the choice of patients to opt for treatment. In these studies, the frame was applied to the communication of treatment information to patients who face the decision to opt for treatment. Armstrong, Schwartz, Fitzgerald, Putt, & Ubel (2002) presented a group of 451 individuals with treatment information. The individuals were randomly divided into three groups. The first group only received the survival rates of the treatment, while the second group received the mortality rates and the third group received both the mortality rates and the survival rates. Upon receiving the information, the individuals were asked to make the decision whether to opt for preventative surgery. The results suggested that individuals who received the mortality rates were less likely to prefer the surgery. These results are in line with the hypotheses of Prospect Theory, as individuals who are presented with the loss frame are risk-seeking and vice versa.

Many studies that follow a similar procedure have been published during the years. A study by Detweiler, Bedell, Salovey, Pronin, & Rothman (1999) concluded that beachgoers who received a message which focused on the gains of using sunscreen were more likely to buy and use sunscreen. Similarly, Schneider et al. (2001) concluded that a message describing the benefits of stopping had a stronger effect on the willingness of the smokers to stop smoking than a message which contained the negative effects of smoking. Kühberger (1998) conducted a meta-analysis of the early contributions in message framing literature. From a sample set of 136 empirical analyses, Kühberger calculated a set of 230 effect sizes. The results were in line with the original hypothesis of Tversky and Kahneman, as messages in the gain frame generally led to risk-averse behaviour and messages in the loss frame caused more risk seeking behaviour.

Relative/absolute

A message can also differ in the way the loss or gain is framed. For example, a gain or loss can be specified in terms of its absolute value or its relative value to a certain threshold or historical value. A study by Malenka, Baron, Johansen, Wahrenberger, & Ross (1993) concluded that a message in which an outcome is framed in relative terms causes more risk-averse behavior than a message which communicates the outcome in absolute terms. However, the authors did only assess the effect of relative/absolute framing for a message which focuses on gains.

Relevancy

The relevancy of the information in the message can also be expected to affect the choices made by decision-makers. A research by De Vries, Terwel, & Ellemers (2014) concluded that adding moderately irrelevant or completely irrelevant cues to a message regarding carbon dioxide capture and storage lowers the persuasiveness of the message and weakens the beliefs of people around the issue.

Hypotheses

The analysis of message framing literature has shown that decisions under risk are expected to be affected by the way in which risks are framed to decision-makers. More specifically, studies that investigate the effects of gain/loss framing have concluded that messages that communicate gains are more effective in nudging people to take preventative measures to mitigate risks. In this line of thinking, buying a secure product can also be seen as a preventative measure to mitigate the risk of cyber threats. Therefore, it can be expected that messages that focus on gains are more successful in nudging users towards buying more secure devices and taking privacy into account when purchasing devices. This leads to the following hypothesis.

H4: Security has a stronger effect on the probability that a device is purchased when it is framed in terms of gains rather than losses.

2.2.3 Personal factors

The second part of the main research question targets the sensitivity of the effect of security and privacy to personal factors. In this section, three personal factors are conceptualised that are expected to moderate the effect of security on the choice behaviour of consumers.

Firstly, the consciousness of consumers with regard to the security risks of IoT devices is expected to moderate the effect of security. According to Khan, Alghathbar, Nabi & Khan (2011), consciousness consists of two key elements: awareness and action. Choi, Kim, Goo & Whitmore (2008) defined information security awareness as the passive association and interest of an individual with regard to security risks. In other words, awareness entails knowledge of security and privacy risks without actuating upon this knowledge. The second component, action, is related to the actual behavioural change that results from the attainment of novel knowledge. The well-known KAB model provides a highly similar view upon behavioural change. According to the model, an individual's knowledge influences his/her attitude towards a certain topic or object. The attitude directly affects the behaviour of the individual, thus creating behavioural change.

Secondly, it seems sensible that the effect of security is sensitive to the acceptance of innovative IT technologies. To illustrate this, security and privacy might have a weaker effect on choice behaviour for consumers who want to use the newest innovative technologies. Davis (1989)

defined the acceptance of technologies as the intention to purchase and use an innovative technology. Moreover, the frequency of use is included in the definition of acceptance.

Thirdly, innovativeness is a plausible personal factor that moderates the effect of security on choice behaviour. Innovativeness is often defined in terms of the willingness of a user to adopt a new technology, thus strongly resembling our definition of technology acceptance (Midlgey & Dowling, 1978). However, for this study, the term resembles an individual's perception of the perceived value of innovative technologies to the individual and society as a whole.

As no previous research has researched the sensitivity of the effect of security to these personal factors, no hypotheses will be drawn. However, an exploratory analysis is conducted to evaluate whether these personal factors moderate the effect of security.

2.3 Conceptual framework qualitative study

For the qualitative part of this study, security and privacy risks are presented to consumers to assess their risk evaluation decision. In order to so, it is necessary to first provide a conceptualisation for the description of risks. Risks are often described by the canonical "Bowtie" model (de Ruijter & Guldenmund, 2016). This theoretical framework describes a risk in terms of a threat, event and consequences.

For this study, a similar conceptual framework of risk will be applied. First of all, a security risk can be described in terms of the actor who poses a threat to the target. For example, if the threat actor is a criminal, the risk may be perceived to be significantly more severe than a similar case where the threat actor is a governmental body. The threat actor poses a threat by harming the Confidentiality, Integrity and/or Availability (CIA) of an asset which is owned by the target. The type of asset can also play a role in determining the severity of the risk. Inflicting harm upon the asset triggers a consequence for the threat actor and the target. For the target, the consequence is negative and can take the forms of privacy infringements, financial damage, reputational damage and other harm. The introduction of a consequence that harms the target deals with the critique of Van Solms & Niekerk on the canonical CIA model. In our theoretical risk framework, the impact of the risk goes further than just the confidentiality, integrity, or availability of the asset. When the confidentiality, integrity, and/or availability of an asset are harmed, this has a consequence for the owner of these assets. The consequence often consists of an impact on assets outside of cyberspace that can be reached via cyberspace. For example, when a smart thermostat is controlled by a harmful adversary, the adversary can damage the home by setting the temperature of the home extremely low or high. Finally, the consequence results in financial, political or other gains for the threat actor.

2.4 Conclusion

The IoT security and privacy literature has primarily focused on discovering the most prevalent risks of IoT devices and networks. The researchers have not yet investigated to what extent consumers are aware of these risks and take these risks into account when purchasing IoT devices. However, the field of Technology Acceptance Modelling (TAM) does provide insights into the effect of security on the decisions of consumers. The results of research in the TAM field suggest that functionality, price and perceived security and privacy influence the acceptance of innovative technologies. From these results, four hypotheses have been postulated regarding the effect of these device attributes on the probability that a device is purchased. These hypotheses will be used to construct a conceptual framework that forms the basis of the quantitative part of this study.

H1: The price of an IoT device negatively influences the probability that the device is purchased.

H2: The number of functionalities of an IoT device positively influences the probability that the device is purchased.

H3: The security level of an IoT device positively influences the probability that the device is purchased.

In the studies, the individuals were presented with a single choice situation. This strongly limits the validity and precision of the resulting model. Moreover, it limits the suitability of the model for analysing the trade-offs between various attributes of products. Finally, the researchers did not consider whether the effect of security on the choices of consumers is sensitive to framing.

The message framing literature clearly suggests that the manner in which risks are framed to individuals affect their choices. For example, according to Prospect Theory, people are more risk-averse when presented with the gains of a situation. On the contrary, people are more risk-seeking when the losses of a situation are communicated. This hypothesis has been validated by numerous studies in the medical field. In these studies, the effect of framing the outcomes of a treatment on the decision of patients to opt for a treatment is examined. The results are in line with the initial hypothesis of Prospect Theory; individuals to whom the survival rates of a treatment are communicated are more likely to opt for the treatment than individuals who obtained a message containing the mortality rates of the treatment. This leads to the following hypothesis:

H4: Security has a stronger effect on the probability that a device is purchased when it is framed in terms of gains rather than losses.

Nonetheless, the effect of framing in communications regarding the security and privacy of IoT devices has not yet been investigated. The effects of framing and other characteristics of communications have been examined for a similar type of content, viz privacy policies of web shops. The results of this analysis propose that the length, complexity and contextuality of privacy policies on web shops have an effect on the willingness of consumers to share information, their rating of the privacy policy and their decisions to purchase products via the web shop. The studies also found that this effect may be sensitive to gain/loss framing. However, it is not yet clear to what extent these conclusions hold for the communication of the security and privacy level of an IoT device towards consumers.

2.4.1 Conceptual framework quantitative study

In this chapter, various hypotheses have been postulated regarding the effect of device attributes, framing and personal actors on the choice behaviour of consumers. The A visual overview of the various hypotheses that have been developed from the literature review has been displayed in the figure below. The hypotheses will be tested by means of a quantitative study. A detailed description of the methodology of this study is be presented in chapter 3.

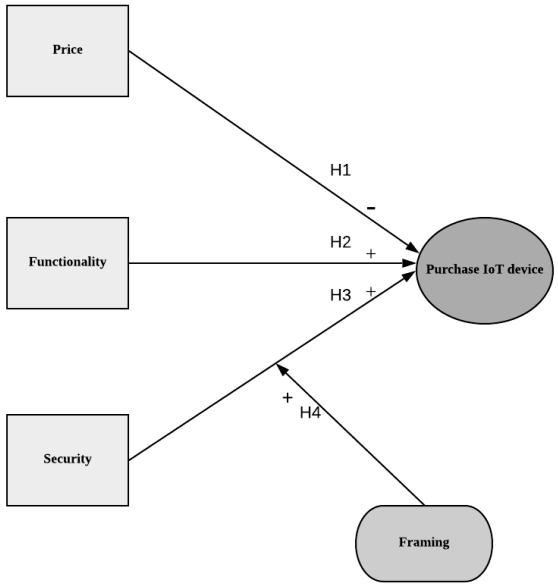


Figure 2.1: Visualisation conceptual framework quantitative study

2.4.2 Conceptual framework qualitative study

In addition, this chapter has provided a conceptualisation of risks in terms of the related threat actor, asset, consequence and target. This conceptualisation will be used to assess the risk assessment decision of consumers for the qualitative part of the study. The factors in the conceptual framework are expected to have a significant influence on the risk assessment of consumers

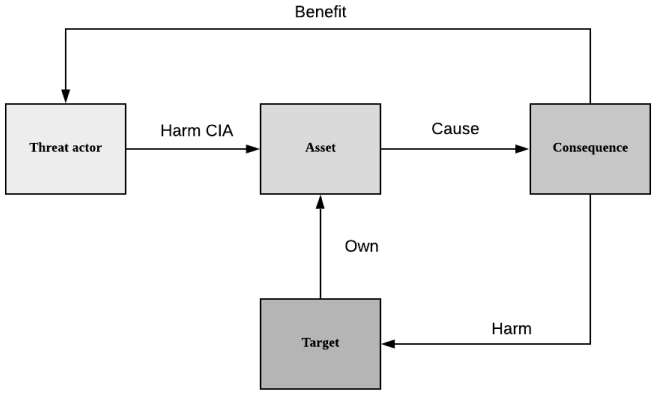


Figure 2.2: Visualisation conceptual framework qualitative study

Chapter 3 Method

This chapter describes the methods that have been applied to answer the research questions. Firstly, the methodology of the quantitative study is discussed. Subsequently, the methods that have been used for the qualitative study are presented and motivated.

3.1 Quantitative study: Stated choice experiment

The main goal of the quantitative study is to investigate the effects of the price, functionality, security and privacy on the choice of consumers to buy IoT devices. Moreover, the research aims to discover whether the effect of security and privacy is sensitive to personal factors and framing. Stated choice experiments are especially suited for such an analysis, as the method allows for the researcher to collect data that describes the effects of attributes on choice behaviour and the trade-offs that are made between these attributes. In a stated choice experiment, respondents are asked to make choices between a set of alternatives that vary in a defined set of attributes. Stated choice data is preferred over revealed choice data in this case, since many consumers have not yet purchased a smart thermostat. The collection of stated preference data allows for the inclusion of potential end-users of smart thermostats. Moreover, this data collection method enables the researcher to control the set of alternatives a decision-maker chooses from. With revealed preference data, it is often unclear from which set of alternatives the decision-maker has chosen. In addition to this, revealed preference data often requires more data to be collected, since only one observation can be collected per respondent and the correlations between attributes are high. This multicollinearity may also cause issues in terms of biased and unreliable parameters. The main downside of collecting stated choice data is the limited validity of the models that are created from such data. It can be argued that the validity of models that have been developed from stated preference data is limited, as people might exhibit significantly different choice behaviour in case of real-world purchase decisions.

In addition to the choices, a set of personal factors has been measured in the experiment, in order to identify whether the effects of the attributes vary between sub groups of consumers. The personal factors have been measured by including a set of indicators in the experiment. These indicators are described in section 3.3.2.

The experiment has been distributed by means of a web-based survey. A web-based survey allows for the collection of a large amount of data in a timely manner without incurring any additional monetary costs. The data collection process has been executed by a group of BSc students from the faculty of Technology Policy and Management (TPM). These students have shared the survey in their personal network, mostly among their social contacts and family members. In most cases, the data collection has an impact on the representativity of the collected sample. In the following chapter, the representativity of the sample and the implications for the validity of the results is discussed.

3.2 Quantitative study: Survey design

In this section, the design of the distributed survey is presented. A full overview of the survey questions can be found in appendix A.1. Firstly, the questions regarding demographic variables are discussed. These questions allow for the assessment of the representativity of the collected sample. Moreover, the design of the stated choice experiment is evaluated. Subsequently, the questions relating to the indicators are considered. These indicators are used to construct values on factors which are expected to influence the choices made by respondents.

3.2.1 Demographics

The survey contained several questions targeted at the age, gender, and education level of the respondents. These questions are included in order to test the representativity of the collected sample. These specific demographics have been chosen, since the Central Bureau for Statistics (CBS) has collected the value of these demographics for the target population of the study.

3.2.2 Indicators

The survey contained questions aimed at operationalising the personal factors that have been defined in section 2.2.3. The indicators were measured by asking the respondents the extent to which they agree with a statement on a five-point Likert scale. Detailed scales for measuring these personal have been developed. For example, detailed scales exist for the operationalisation of factors such as cybersecurity awareness and innovativeness. (Egelman & Peer, 2015; Goldsmith & Hofacker, 1991). Applying these scales to measure the personal factors in the survey would vastly increase the amount of questions in the survey. As the stated choice experiment is the most crucial part of the survey and asks a lot of time and effort from the respondents, the decision has been made to limit the number of questions targeted at measuring factors. This decision makes the use of previously defined and tested scales for the measurement of factors infeasible. Therefore, a small set of indicators have been included into the survey to operationalise the factors that have been defined in section 2.2.3.

Principal Axis Factoring (PAF) has been used to construct the factors. PAF is especially suited to measure the values of non-measurable constructs such as views, opinions, and beliefs. The factor scores have been calculated by means of linear regression. This method assumes that the relationship between the indicators and factors is linear. The axes have been rotated by making use of oblique rotation, which allows the axes to be correlated if this provides a better fit for the data. The number of factors is determined by the threshold value for the eigenvalue of the solution. An extra factor is accepted into the model if the initial eigenvalue of the solution containing this extra factor is larger than 1.

The indicators, their respective statements, and motivation have been listed in the table below.

Indicator	Statement	Motivation
I1:	“I keep up with technological developments”	Positive interaction with functionality, negative interaction with price, positive interaction with security
I2	“I read the technology section when reading newspapers or visiting news websites”	Positive interaction with functionality, negative interaction with price, positive interaction with security
I3	“I find it interesting to follow the development of new IT products”	Positive interaction with functionality, negative interaction with price, positive interaction with security
I4	“Innovation is important for economic development”	Positive interaction with functionality

I5	"Investments in innovative technologies are important for society"	Positive interaction with functionality
I6	"If a new IT product has been developed, I want to buy the first version"	Positive interaction with functionality, negative interaction with price, positive interaction with security
I7	"I pay attention to the security risks of my IT devices"	Positive interaction with security
I8	"When purchasing an IT device, I consider the security risks of the device"	Positive interaction with security
I9	"The security of my IT devices is important to me"	Positive interaction with security
I10	"My personal information should be protected sufficiently"	Positive interaction with security
I11	"I keep track of which information is collected when using online services"	Positive interaction with security
I12	"I am concerned with the security risks of my IT devices"	Positive interaction with security
I13	"When using IT devices, I am concerned with the use of my personal data by external parties"	Positive interaction with security
I14	"When using online services, I am concerned with the use of my personal data by external parties"	Positive interaction with security
I15	"I undertook action to improve the security of my IT devices"	Positive interaction with security

Table 3.1: Indicators

3.2.3 Case introduction

Before entering the section of the survey that contains the stated choice experiment, the respondents were asked to read a short description of the concept of smart thermostats. This description is shown below.

"A smart thermostat is a thermostat that is connected to the internet. The internet connectivity allows the smart thermostat to enable new functionalities which traditional thermostats are not

able to provide. For example, smart thermostats can be controlled from a distance by making use of specially designed mobile applications."

This specific case introduction has been chosen as it very clearly reflects the IoT concept. The IoT concept is characterised by everyday objects being connected via internet connections, in order to enable novel and innovative functionalities. Moreover, the remote-control functionality is included, since even the most basic smart thermostats enable this functionality and it is the simplest to understand.

3.2.4 Stated choice experiment

Attributes

The alternatives in the experiment varied in three attributes; Price, functionality, and security. Adding more attributes to the design would strongly increase the needed amount of choice sets per respondent to develop valid models. Therefore, privacy is excluded as an attribute from the analysis. The attributes are selected because they are expected to have a strong influence on the choice of consumers to buy a smart thermostat. The price of smart thermostats varies roughly between 100 and 250 euros. A list of the various smart thermostats in the market with their features and price when purchased via Amazon has been listed below.

Name	**Features**	**Price**
Nest learning thermostat	Voice control (Alexa) Learning Geofencing Google home compatibility	€230-250
Ecobee4	Voice control (Alexa) Geofencing Apple homekit integration	€180-220
Sensitouch Wi-Fi Thermostat	Large user interface Scheduling	€130-175
Wiser Air	Voice control (Alexa) Learning	€145-175
Lyric T6	Scheduling Geofencing	€145-175
ZEN thermostat	Scheduling Remote control	€90-110

Table 3.2: Smart thermostat prices

The level of the price attribute in the survey varied on four levels in order to represent the variance of this attribute in actual devices. The distance between the four levels is kept constant in order to ensure orthogonality between attributes. The following levels are used to vary the price of the smart thermostats in the experiment. For the analysis, the price attribute has been scaled by dividing it by 100. This allows for the estimation of similar parameters for the attributes.

Price attribute level	Price
1	€100
2	€150
3	€200
4	€250

Table 3.3: Price attribute levels

The various smart thermostats provide a diverse set of functionalities. The functionality attribute has been represented by four levels. With each extra point on the functionality attribute, the smart thermostat possesses an extra functionality. The four functionalities have been described in the table below.

Functionality	Description
Remote control	The user is able to remotely access the device in order to adjust the temperature, scheduling or make use of other functionalities
Geofencing	The geofencing capability of the user's smartphone is used to assess whether the users has left his/her house and adjust temperatures accordingly
Sensing	The home is equipped with sensors, which assess whether the occupants are awake, sleeping or outside of the house. The temperature is adjusted according to the data collected by the sensors
Learning	The user inputs basic schedule parameters. The device makes use of algorithms in order to learn the schedule of the occupants and collects data from sensing to detect changes in the schedule and respond to them

Table 3.4: Functionalities

Each of these functionalities is related to a level of the functionality attribute. To illustrate this, the various levels of the functionality attribute as well as the related functionalities have been displayed below. As the functionality attribute level increases, a more complex functionality is added to the device. The functionality attribute could also have been coded by four binary variables representing each of the functionalities. However, this would create unrealistic alternatives. To illustrate this, it is unrealistic that a smart thermostat with learning capabilities does not allow for remote control.

Functionality attribute level	Functionalities
1	Remote control
2	Remote control, Geofencing
3	Remote control, Geofencing, Sensing
4	Remote control, Geofencing, Sensing, Learning

Table 3.5: Functionality attribute levels

Finally, the security level is varied on two levels, since the operationalisation of security on more levels is problematic. Although there exist frameworks that assess the security level of (IoT) devices, these frameworks are far too complex to use in the context of a stated choice experiment. Therefore, the description is kept very basic and simple to understand from a consumer perspective.

Chapter two concluded that gain/loss framing is the most prevalent method in framing literature. According to existing studies, gain/loss framing is expected to have the most significant effect on the choices of decision-makers. Therefore, gain/loss framing has been chosen as the method to frame the security level of the security level. In the gain frame, the description of the security level focuses on the measures which have been taken in order to protect the device from adversaries. It is assumed that most consumers are not aware of the specific security controls on the devices. Thus, a description is chosen which is also understandable for consumers who do not possess any technical or security-related knowledge. The main aspect of these controls lies in their ability to secure the device. This leads to the following description: “this device is/is not secured properly. In the loss frame, the description of the security level should be target at the risks of possible cyber-attacks. The number of risks related to smart thermostat is immense. Including each of these risks would create an excessively large description of the security level, which is not suited for stated choice experiments. Thus, an overarching term is needed which includes a large set of the risks of smart thermostats. This leads to the following description: “This device can/cannot be hacked”. The description of the security level for each group is presented in the table below.

Frame	**Security attribute level**	**Description**
Gain	-1	“This device is not secured properly”
Gain	1	“This device is secured properly”
Loss	-1	“This device can be hacked”
Loss	1	“This device cannot be hacked”

Table 3.6: Security attribute levels

In order to assess the effect of framing on the choices made by respondents, the respondents were split up into two groups. The first group was given the gain frame, while the second group received the loss frame message.

Alternatives and choice sets

When the attributes and their levels are identified, it is key to construct the choice sets for the stated choice experiment. Firstly, choice sets can be constructed by means of a full factorial design. In such a design, respondents are asked to make choices between every possible alternative. A full factorial design is simple to implement and allows for the estimation of each of the direct and interaction effects. However, the number of alternatives increases vastly when introducing more attributes and more levels in these attributes.

For this reason, fractional factorial designs are often preferred over full factorial designs. A frequently used type of fractional factorial design is the orthogonal design. Orthogonal designs are constructed in such a way that the various attributes of the alternatives in the design are not correlated. This allows for a precise estimation of direct effects. Moreover, it is important that attribute level balance is ensured. Attribute level balance implies that each level of the attributes has the same amount of observations in order to ensure that the attributes have the same standard errors.

A commonly used method to construct orthogonal designs is the basic plan. Basic plans are published fractional factorial designs which allow for the constructing of alternatives using a given set of attributes. If the attributes are assigned correctly, attribute level balance is ensured. For this experiment, a basic plan is a sensible experiment design method, since it allows for the constructional of an orthogonal design in a simple and timely manner. Basic plan 3 has been used to construct the alternatives for the stated choice experiment, as this basic plan is especially suited for a design which includes two variables with four levels and one variable with two levels.

A description of basic plan 3 and the set of alternatives that has been constructed by making use of this basic plan can be found in Appendix B.1. The alternatives have been randomly assigned to a set of 16 choice sets. The choice sets have been divided into two blocks in order to limit the amount of choices a respondent faces in the experiment. The choice sets are presented in the table below. A drawback of blocking is that attribute level balance is not ensured fully. As a result, it is possible that the coefficients within an attribute are not estimated with an equal precision.

Choice questions

Per choice set, the respondents were presented with three questions. Firstly, the respondents were asked whether they would purchase each individual device in the case their (regular) thermostat had broken and were faced with the decision to purchase a smart thermostat. Asking this question allows for the inclusion of respondents who have not purchased a smart thermostat yet.

Secondly, the respondents were asked to make a choice between the two smart thermostats. Their choice indicates which of the two smart thermostats they would buy in the case they had to buy one of the two alternatives. The first two questions focus more on the general willingness of the respondents to buy a certain smart thermostat, while the third question measures a choice given that the respondent is willing to buy a smart thermostat in general. An example of a choice question has been displayed below. A set of example choice questions can be found in appendix A.1.3.

	Product A	Product B
Price	150	250
Functionality	Remote control, Geofencing	Remote control
Security	This device can be hacked	This device cannot be hacked

Imagine that your (smart) thermostat has broken and you are faced with the decision to buy a new smart thermostat, would you

1: Buy product A? (Yes/No)
2: Buy product B? (Yes/No)

3: In this case, if you had to choose one of the options, would you buy product A or product B? (A/B)

Figure 3.1: Example choice question

In this case, the first two questions are labelled as the "single alternative" responses, while the third question is labelled as the "cross alternative" response

3.2.5 Model groups

The resulting choice models can be divided into two model groups. The models in the first group are constructed with the cross alternative responses as the dependent choice variable, while the models in the second group are developed with the single alternative response as the dependent choice variable**.**

Model group	Choice input
1: Cross alternative choice models	1: Would you purchase option A or option B?
2: Single alternative choice models	1: Would you purchase option A? 2: Would you purchase option B?

Table 3.7: Model groups

3.3 Quantitative study: Discrete choice modelling

The data that has been collected from the stated choice experiment is analysed by means of discrete choice modelling. Discrete choice modelling is appropriate for this goal, since the method is primarily focused at quantitatively determining the influence of attributes and factors on choices. With this method, models are developed that describe the choices of a decision-maker from a set of two or more discrete alternatives.

3.3.1 RUM and MNL

More specifically, Random Utility Maximization (RUM) based models will be developed to assess the direct effects of device attributes as well as their interactions with framing and personal factors. RUM based models describe the probability that a certain decision-maker chooses an alternative from a given set of alternatives which vary on a set of criteria. This probability is based upon the utility that an alternative provides to the decision-maker. It is assumed that a decision-maker chooses the alternative that has the maximum utility of the choice set.

Specification

For this study, a specific type of RUM based model will be developed, the Multinomial Logit (MNL) model. MNL models assume that the error term in the utility function is independently and identically distributed across all alternatives, which implies that they have the same probability distribution and are mutually independent. The utility of an alternative is calculated by the sum of the product of the criteria scores and a set of linear parameters. Thus, the utility is calculated by the following formula:

$$U(a_i) = \sum_{j=1}^{m} w_j * E(a_i, c_j) + \varepsilon \quad (1)$$

Where
w_x = The parameter or weight of attribute x
$E(a_x, c_y)$ = The expected effect of alternative x on attribute y
ε = Error term

For MNL models, the probability that an alternative is chosen from a set of alternatives is calculated by the following formula:

$$P(X = a_i) = \frac{e^{U(a_i)}}{\sum_{j=0}^{n} e^{U(a_j)}} \quad (2)$$

Where
$P(X = a_x)$ = The probability that alternative x is chosen from the choice set
$U(a_x)$ = Utility of a_x
n = The number of alternatives in the choice set

3.3.2 Model statistics

Various model statistics will be calculated to assess the goodness of fit of the models. These statistics show to what extent the developed models fit the collected sample.

The log-likelihood of a model is equal to the log of the likelihood of the calculated parameters for the dataset. The log-likelihood of the model can be calculated by the following formula:

$$LL(\beta) = \ln \left(\prod_{n} \prod_{i} y_n(i) * \ln\left(P_n(i|\beta)\right) \right) \quad (3)$$

Where
β = Estimated parameters
n = The set of observations
i = The set of alternatives
y_i = 1 When an alternative is chosen and
y_i = 0 When an alternative is not chosen

The Rho squared statistic of a model determines the amount of the total variance which can be explained by the model. The statistic can be calculated by the following formula:

$$\rho^2 = 1 - \frac{\mathcal{L}^*}{\mathcal{L}^i} \quad (4)$$

Where
$\mathcal{L}^*$ = The log-likelihood of the estimated model
$\mathcal{L}^i$ = The log-likelihood of the initial model, in which each parameter is fixed at 0

Comparing various models is of high importance for model selection. The Likelihood ratio test is especially suited for this purpose. The Likelihood ratio statistic is calculated by the following formula:

The Likelihood Ratio Test (LRT) can be used to compare a set of two models. The test describes the extent the better fit of a model can be explained by peculiarities in the sample. This test is based upon the value of the Likelihood Ratio Statistic (LRS). The value of this statistic is calculated by the following formula:

$$\text{LRS} = -2 * (\text{LL}_\text{A} - \text{LL}_\text{B}) \quad (5)$$

Where
LL_x = The Log-Likelihood of model x

The hypotheses of this test are as follows:

H0: “Model B is not a more valid model than model A”

H1: “Model B is a more valid model than model B”

If the number of observations in the collected sample is large, the LRS is distributed X^2 with q degrees of freedom. If a high LRS is found, the probability of finding this LRS value (or a higher LRS value) is relatively small, which implies the null hypothesis should be rejected. In that case, we can conclude that model B is a better fit than model A.

3.3.3 Indicators

In order to analyse the trade-offs that have been made by respondents, the values for two indicators are calculated: The Willingness to Pay (WtP) and market shares. These indicators are calculated from the parameters of the resulting models and provide further insights into the decision-making process of the respondents.

The Willingness to pay (WtP) compares the extent to which a certain variable impacts the utility of an alternative compared to the impact of price on the utility of an alternative. Therefore, the indicator gives insight into the willingness of the respondent to pay for an improvement of an alternative with to regard to an attribute.

For linear models, the WtP can be calculated by the following formula:

$$WtP_x = \frac{\frac{\delta V}{\delta X}}{\frac{\delta V}{\delta P}} = \frac{\beta_x}{B_P} \quad (6)$$

Where
X = The variable of which the WtP is calculated
V = The value or utility of an alternative
P = The price of an alternative
β_x = The estimated parameter of attribute x

3.4 Qualitative study

In order to investigate which underlying rationales determine the effects of security and privacy on the choices, a qualitative study was conducted. The qualitative study provides a more detailed and in-depth analysis of the decision-making process of consumers purchasing smart home devices. The study focused on how consumers perceive the security and privacy risks of smart home devices as well as whether they take these risks into account when purchasing smart home devices.

The qualitative study was conducted by means of a web-based survey. In this survey, respondents were asked which security and privacy risks of smart thermostats were known to them and whether these risks played a role in their decision to buy a smart thermostat. As it is assumed that many consumers are not aware of the privacy and security risks of smart thermostats, a set of hypothetical scenarios were included in the survey. Each of these scenarios described a privacy or security related risk of smart thermostats. The respondents were asked to rate the severity of the risk that has been described in the scenario and provide a motivation for their rating. The aim of the hypothetical scenarios is to trigger respondents who are not aware of any security or privacy risks to consider their perception on these risks.

3.5 Qualitative study: Survey design

In this section, the survey is described which has been used to conduct the qualitative study. A complete overview of the survey questions can be found in appendix A.2. The survey contains a set of demographic questions, questions addressing risk awareness and a series of questions which instruct the respondent to assess risk scenarios.

3.5.1 Demographics

The qualitative study contained a set of questions focused at determining the values of demographic variables. Similarly to the stated choice experiment, the survey contained questions that ask for the age, gender, and education level of the respondents.

3.5.2 Purchase decision

Firstly, it is necessary to make a distinction between respondents who own a smart thermostat and those who do not own a smart thermostat. Both groups are included in the survey, as the underlying rationales for consumers who chose not to purchase a smart thermostat are equally

important to the rationales for consumers who did purchase the device. The following section describes the questions that have been included for each group.

For device owners, it is important to know why the respondent chose to buy a smart thermostat in general, as this might highlight security or privacy considerations without explicitly mentioning security or privacy in the questionnaire. Additionally, the device owners were asked for their motivation to buy a specific smart thermostat, as security considerations might have played a role in this decision as well. For example, a respondent might have chosen not to buy a certain smart thermostat as she does not trust the manufacturer to handle her personal information. As for non-device owners, the decision not to buy a smart thermostat might also have been induced by security or privacy considerations. Subsequently both respondent groups were asked to assess the effect of security and privacy on their decision to buy or not to buy a smart thermostat. Device owners might have considered the security and privacy risks after purchasing the device. These considerations might highlight other rationales than those that have been involved in the purchase decision itself. Moreover, it is possible that device owners undertook actions to improve the security of the device or their privacy. The actions of device owners demonstrate their willingness to invest resources into security and privacy.

3.5.3 Risk awareness

Subsequently, a set of questions are added to assess whether the respondents are aware of the security and privacy risks of smart thermostats. The questions regarding risk awareness are placed after the purchase decision section, since it is not desirable to trigger respondents to contemplate security or privacy risks in detail before asking them to evaluate the effect of security and privacy on their purchase decision. This might cause the respondents to overestimate the effect of security and privacy on their purchase decision.

3.5.4 Scenarios

Finally, the respondents were presented with a set of scenarios. Each of the scenarios describes a security or privacy risk of smart thermostats. For each scenario, the respondents were asked to assess the severity of this risk on a five-point scale and provide a motivation for their assessment. Additionally, the respondents were asked which scenario describes the most severe risk in their opinion and whether the risks would influence their choice to purchase a smart thermostat. In the case respondents do not possess any knowledge relating to the risks of smart thermostats, these scenarios allow us to trigger the respondents to consider these risks and provide an indication of how they perceive these risks. Furthermore, the results of these questions allow for the generation of policy recommendations regarding risk awareness campaigns by governmental bodies. For such campaigns, it is necessary which risks should be communicated in order to effectively nudge users towards buying more secure devices and securing the purchased devices properly. The scenarios have been developed by making use of the conceptual framework that has been presented in section 2.1.

Scenario nr.	Threat actor	Asset	Consequence	Scenario
1	Criminal	Energy use and location data	Privacy infringement Increased risk of burglary	The smart thermostat collects data about your energy use and keeps track of your location. A criminal gains access to this information to determine the right moment for a burglary

2	Producer Energy supplier Insurer Tax authorities	Energy use and location data	Privacy infringement Financial damage Legal consequences	The smart thermostat collects data about your energy use and keeps track of your location. The producer of your thermostat collects this data and may be obligated to share it with external parties, such as insurers or tax authorities.
3	Energy supplier Producer Marketing bureaus	Energy use and location data	Privacy infringement	The smart thermostat collects data about your energy use and keeps track of your location. The producer of your thermostat collects this data and shares it with marketing bureaus, which use it to develop personalised advertisements.
4	Criminal	Smart thermostat	Physical damage	A criminal gains access to your smart thermostat, allowing him/her to control the heating in your house.
5	Criminal	Smart thermostat Home network Confidential information on home network	Privacy infringement Financial damage	A criminal gains access to your home network via your smart thermostat, allowing the criminal to gain access to personal information on the network, such as passwords or browsing data.
6	Criminal	Smart thermostat	Lower performance Legal consequences	Your smart thermostat is part of a large network of devices which is being used to execute cyber-attacks on large organisations.

Table 3.8: Hypothetical scenario's qualitative study

3.6 Tools

A variety of tools have been used to conduct the study. First of all, the surveys have been developed and spread by making use of Collector. Collector is a software package which allows researchers to construct a survey, spread it to respondents and export files containing the responses of the survey. SPSS has been used to analyse the representativity of the collected sample. SPSS is a statistical software package which allows researchers to execute statistical tests with an easily understandable GUI. Finally, the discrete choice models have been developed with Biogeme. Biogeme is an open-source Python module, which allows for the estimation of a diverse set of discrete choice models (Bierlaire, 2018). Although the software package does require some basic programming skills, it provides a high level of flexibility for the modeller.

Chapter 4 Results: Quantitative study

In this chapter, the results of the quantitative part of this study are discussed. Firstly, the response of the survey is examined. Secondly, the representativity of the collected sample is analysed. Subsequently, the personal factors are constructed from the collected indicator values. Furthermore, the model selection process is described. In addition, the parameters and indicator values from the selected model are presented. Finally, conclusions are drawn to develop the resulting causal model of the study.

4.1 Response

The students collected a dataset containing 709 respondents. A subset of 93 respondents who did not provide an answer to the questions related to the choice experiment were removed from the dataset. Moreover, a set of 35 responses were collected from the same IP address within a distinctly small time frame. These responses were removed from the data set as it is unlikely for such a large amount of valid responses to be collected within a small time frame from the same IP address. It is likely that these responses consist of students who filled the survey in themselves multiple times.

The response rate of the survey can be calculated by the following formula:

$$\text{RR (Response Rate)} = \frac{\text{Number of completed survey respones}}{\text{Total number of survey responses}} * 100\% \quad (1)$$

A total of 581 respondents submitted a complete response to the choice experiment. Thus, the response rate of the survey is 81.95%. The total number of survey responses is underestimated, since the respondents who received the invitation to fill in the survey and chose not to open this link are not included in the measurement of this variable. It is not clear exactly how much respondents did receive the link but decided not to open the survey. The statistic shows that a significant number of respondents did not provide a complete response to the survey, which can be explained by the complexity of the subject or the questions in the survey. Although a short description was provided, it is likely that respondents who were not familiar with the concept of smart thermostats had issues with answering the questions related to the choice experiment and therefore were not able or willing to provide a complete response.

4.2 Representativity analysis

The research is targeted at determining the choices made by (potential) consumers in the Netherlands. Only consumers older than 18 years are included in the target population of the study, as consumers below 18 years generally do not live independently and therefore are not likely to make the decision to purchase and use a smart thermostat. Therefore, the target population of the research can be defined as “Dutch citizens older than 18 years”. In order to generalise conclusions found in the sample to the population, the collected sample should represent this population. The survey contained a number of questions regarding demographics of the respondents. The demographics included in the survey are age, gender, education level, and working situation. In this section, the demographics of the respondents in the sample are compared to population data in order to assess whether the collected sample represents the target population with regard to the selected demographics. The CBS has published multiple datasets that contain information regarding these demographics for Dutch citizens (CBS, 2019; CBS, 2018). Moreover, the problems that arise from such overrepresentations are assessed by estimating choice models that contain the direct effects of the demographic variables and the

interaction effects of demographic variables and device attributes. Furthermore, the correlations between the demographics and various other explanatory variables are analysed. Detailed results of the analysis with regard to the representativity of the collected sample can be found in appendix C.

4.2.1 Results

The main goal of the representativity analysis is to test whether the distribution of the demographic variables in the sample is significantly different from the distribution of these variables in the target population. This analysis has been conducted by making use of a Chi-Squared tests. The table below displays the frequencies of the categories in the demographic variables for the sample and the population. Moreover, the value of the Chi-squared statistic, the number of degrees of freedom and the resulting p-value are presented. The most significant overrepresentations are highlighted in blue. The most significant underrepresentations are highlighted in green.

		Sample (%)	Population (%)	Difference	Chi2	df	p-value
Gender	Man	49.85	49.84	-3.5	2.994	1	0.084
	Vrouw	50.15	50.15	3.5			
Age	18-24 years	33.27	11.06	22.21	449.053	7	0.000
	25-29 years	8.58	7.74	0.84			
	30-39 years	6.48	15.42	-8.94			
	40-49 years	7.71	19.37	-11.66			
	50-59 years	31.87	17.76	14.11			
	60-69 years	8.93	15.07	-6.14			
	70-79 years	2.1	8.92	-6.82			
	80 years and older	1.05	4.66	-3.61			
Education level	Elementary	0.34	10.3	-9.96	601.820	4	0.000
	Vocational	1.03	8.9	-7.87			
	MBO	9.81	41.2	-31.39			
	HAVO/VWO	18.24	9.5	8.74			
	WO	70.56	30	40.56			
Working situation	Student	33.3	18.6	14.7	600.152	4	0.000
	Paid job	54.6	33.6	21			
	Unemployed	3.7	11.4	-7.7			
	Retired	4.2	8	-3.8			
	Other	3.4	6.3	-2.9			

Table 4.1: Results representativity analysis

Firstly, the age group 18-24 years and 50-59 years are highly overrepresented, while the age groups 30-39 years and 40-49 years are underrepresented. This result can be explained by the data collection process. The BSc students most likely shared the survey with their friends, siblings or mature family members. These groups are expected to belong to the overrepresented age categories. With regard to education level, the lower education levels are underrepresented, while the higher education levels are overrepresented. As the students who spread the survey belong to the higher education level category, it is likely that this education level is overrepresented in their collected responses. Finally, the sample mostly contained respondents who are either students or have a paid job. Again, this is most likely due to the fact that the students shared the survey with their friends, siblings or mature family members.

4.2.2 Problems due to over- and underrepresentations

The analysis has highlighted some clear overrepresentations of certain groups in the sample in comparison with the target population. In order to assess whether these overrepresentations are problematic, a MNL model has been estimated that assesses the effect of demographic variables on the choices made by the respondents. The results show that the demographics do not have a significant direct effect on the choices of the respondents. However, two interaction effects of the demographics with the device attributes have been found. For example, the interaction between gender and functionality has a significant negative parameter. Secondly, the interaction between gender and security has a positive parameter. The parameter for this interaction effect is not statistically significant.

Overrepresentations of certain demographics might also cause some issues because of their correlations with the personal factors. The complete set of correlations between the demographics and other explanatory factors has been displayed below. The development of these factors is discussed in section 4.3. The results show that the correlations between the demographic variables and the factors are low, which indicates that the effect of under- or overrepresentations with regard to these demographic variables is unproblematic.

	Privacy/Security Consciousness	**Technology Acceptance**	**Conservativeness**
Gender	0.10	0.10	0.10
Education	-0.09	-0.09	-0.09
Age	0.10	0.10	0,10

Table 4.2: Correlations

4.2.3 Conclusions

The analysis of the representativity has shown that there exist some significant overrepresentations of certain groups in the sample. These overrepresentations might cause under- or overestimation of relations between factors, attributes, demographics and the choice behaviour of consumers. This might limit the precision of the estimated relationships between these variables as well as the ability of the resulting model to predict values of variables in observations outside of the collected dataset. However, the main goal of this research is to show that certain relations exist between these variables. The overrepresentations do not limit the

ability of the developed models to reach this goal. Thus, the overrepresentations in the sample are deemed not to be problematic.

4.3 Factor analysis

An important prerequisite for factor analysis is that the communality of every indicator should be higher than 0.25. The communality of an indicator is equal to the shared variance with the other indicators. The communalities of each indicator have been listed below. The communality of I6 is slightly below the threshold but has still been included in the analysis because of its high load on one of the factors.

The factors, indicators and their relative loads have been displayed below. In order to simplify the interpretation of the results, loads below a threshold value of 0.3 are not displayed. Moreover, indicators that do not load high on any factors or load on multiple indicators have been excluded from the analysis.

The final step of the factor analysis is the labelling of factors. The indicators are developed to measure the factors that have been defined in section 2.2.3. The first factor is defined by indicators that relate to the attitude of the respondents towards privacy/security issues of IT/IoT devices and the extent to which the respondent has taken action to tackle these issues. Therefore, the first factor corresponds to the "Privacy/Security Consciousness" factor. The second factor relates to the respondent's interest in the development of technology as well as their adoption of new technology. Therefore, the second factor can be labelled as "Technology Acceptance". Finally, the third factor is determined by the two indicators that measure the importance of innovation. Thus, this factor is in line with our definition of "Innovativeness". The two indicators load negatively on the factor, which implies that the indicators measure the pole opposite of this construct. Consequently, this factor can be labelled as "Conservativeness". The indicators that have been removed from the factor analysis are excluded from the analysis completely, since they do not possess a significantly different meaning than the factors.

Indicator name	**Communality**	**Factor 1: Privacy/security Consciousness**	**Factor 2: Technology acceptance**	**Factor 3: Conservativeness**
I1: IT Trends	0.609	-	-	-
I2: IT News	0.513		.785	
I3: Development IT products	0.588		.733	
I4: Economic importance innovation	0.612			-.888
I5: Societal importance innovation	0.592			-.830
I6: Technology adoption	0.245		.536	

I7: Consideration Security risks	0.620	-	-	-
I8: Security concern purchase	0.569	.556		
I9: Importance security	0.653	-	-	-
I10: Importance privacy	0.616	-	-	-
I11: Privacy awareness	0.354	.534		
I12: Concern security	0.517	.755		
I13: Concern privacy IT devices	0.690	.897		
I14: Concern privacy online services	0.649	.833		
I15: Action Security	0.326	.407		

Table 4.3: Indicators with communalities and factor loads

4.3.1 Factor correlations

As described in chapter 3, oblique rotation has been used to construct the various factors. This rotation method allows for the existence of correlations between the factors. The correlations between the factors have been listed in the table below. The results show that some low to moderate correlations exist between the factors. Firstly, the correlation between the Privacy/Security Consciousness factor and the technology acceptance factor is positive. This seems sensible, as it can be expected that people who are among the first adopters of technology generally are interested in technological developments and therefore have attained some knowledge regarding privacy and security. Secondly, the correlation between the Privacy/Security Consciousness factor and the conservativeness factor is negative. A possible explanation for this result is that people who do not value innovative technologies are not interested in subjects related to privacy and security. Finally, the correlation of the technology acceptance factor with the conservativeness factor is negative, as people who do not value innovation strongly are not likely to be among the first adopters of innovative technology.

	Privacy/Security Consciousness	**Technology Acceptance**	**Conservativeness**
Privacy/Security Conciousness	1,000	0,242	-0,294
Technology Acceptance	0,242	1,000	-0,279
Conservativeness	-0,294	-0,279	1,000

Table 4.4: Factor correlations

4.4 Model estimation process

4.4.1 Model group 1

A summary of the first group of models with their respective loglikelihood, rho squared value and LRT can be found below. The corresponding parameters of the models are displayed in appendix D. The LRT values indicate that a significantly better fitting model is found from adding a set of new parameters to the model in each iteration. Thus, the final iteration (model 1.4) has the highest quality.

Model	Description	Final log likelihood	Rsquare	LRT (Critical value)
1.1	MNL: Device attributes	-2556,545	0,156	-
1.2	MNL: Device attributes + interaction effects factors and framing with security attribute	-2514,620	0,17	83,85(9,488)
1.3	MNL: Device attributes + interaction effects factors and framing with security and functionality attribute	-2496.602	0.176	36.036 (9.488)
1.4	MNL: Device attributes + interaction effects factors and framing with security. functionality and price attribute	-2488,602	0,179	16 (9,488)

Table 4.5: MNL models cross alternative choices

4.4.2 Model group 2

A similar model development process has been applied for the models that have been developed from the single alternative choices. A description of the models, including their respective log-likelihood, rho squared value, and LRT value has been displayed below. The values of the parameters for each model can be found in appendix E.

Model	Description	Log likelihood	R-square	LRT (Critical value)
1.1	MNL: Device attributes	-5054,914	0,265	-
1.2	MNL: Device attributes + interaction effects factors and framing with security attribute	-4580,136	0,297	949,556 (9,488)
1.3	MNL: Device attributes + interaction effects factors and framing with security and functionality attribute	-4544,435	0,306	71,402 (9.488)
1.4	MNL: Device attributes + interaction effects factors and framing with security. functionality and price attribute	-4541,340	0,307	6,19 (9,488)

Table 4.6: MNL models single alternative choices

4.5 Model selection

The main difference between both model groups lies in the semantics of the dependent choice variable. The models in model group 1 are developed from the choices between two smart thermostats. In this choice situation, it is assumed that the respondents are willing to buy a smart thermostat. This assumption might be incorrect, as it is possible that respondents do not want to buy any of the smart thermostats in the choice set. In other words, he choice observation measures the effect of device attributes on the choice for a specific smart thermostat, rather than the effect of security to buy a smart thermostat in general. On the contrary, the models in model group 2 have been developed from the single alternative choices. In this case, it is possible for the respondent to indicate that he/she does not want to buy any of the smart thermostats from the choice set. The models in model group two measure the effect of device attributes on the acceptance of smart thermostats. Therefore, the models in model group 2 are more suited for answering the research question of this study.

Moreover, the comparison of the model statistics for both model groups show that the models in model group 2 have a higher rho squared value. This implies that the models in model group

2 provide a better fit to the data in comparison with model group 1. On the contrary, the models in model group 2 have a lower log likelihood value than the models in model group 1. This can be explained by the amount of choice questions in each model group. In model group 2, the dependent variable consists of two choices per choice set, while the dependent variable for model group 1 entails a single choice per choice set. Thus, the log likelihood of model group 2 is expected to be lower, regardless of the goodness of fit of the developed models.

One possible downside of drawing conclusion from the models in model group 2 is the possibility of bias in the dependent variable. If respondents have a negative perception of smart thermostats in general, it can be expected that they reply "No" to each choice question. In order to check for such a bias, the distribution of "Yes" and "No" responses in the dataset should be investigated. The results of this analysis have been presented in the table below. The results show that the amount of "Yes" and "No" replies in the dataset is relatively balanced.

Choice	Frequency (#)	Percentage (%)
Yes	4652	53,4
No	4060	46,6

Table 4.7: Choice distribution model group 2

The LRT test concluded that the final model, model 1.3, provided the best fit for the collected data sample. However, the LRT value of model 1.4 is relatively close to the threshold value of 9,488 and contains a significant interaction of the price attribute with the Technology Acceptance factor. For these reasons, model 1.4 will be used to draw conclusions for the remainder of this chapter.

4.6 Model results

The resulting parameters of model 1.4 from model group 2 are displayed in the table below.

Attributes	**Parameter**	**p**
Price(€/100)	-0,656	0,000
Functionality	0,108	0,000
Security	1,041	0,000
Constant	0,771	0,000
Framing interactions		
Framing * Security	0,041	0,000
Framing * Functionality	0,025	0,264
Framing * Price	-0,025	0,315
Factor interactions		
Technology Acceptance * Security	-0,054	0,092
Privacy/Security Consciousness * Security	0,162	0,000
Conservativeness * Security	-0,098	0,001
Technology Acceptance * Functionality	0,095	0,000
Privacy/Security Consciousness * Functionality	-0,126	0,525
Conservativeness * Functionality	-0,037	0,152
Technology Acceptance * Price	-0,059	0,045

Privacy/Security Consciousness * Price	0,022	0,429
Conservativeness * Price	-0,042	0,132

Table 4.8: Parameters model 1.4 from model group 2

Firstly, the model shows the direct effects of the three device attributes on the purchase decision. In line with our hypothesis, the probability that a device is purchased increases as the price of the device lowers. On the contrary, secure devices or devices that provide a high number of functionalities are more likely to be purchased. The effect of security on the purchase decision is notably high.

In order to illustrate this, the impact of varying the security attribute on the probability that a device is purchased is calculated for a sample set of 4 alternatives. For each alternative, the values of the price and functionality are kept constant. The values for the personal factors are set to their mean value of 0. Then, the value for ΔP is calculated by subtracting the purchase probability when the alternative is unsecure from the purchase probability when the device is secure. The resulting values are displayed in the table below. The results show that security has a strong impact on the purchase decision under constant price and functionality. Over all alternatives in the design, the average value for ΔP is equal to 0,44. This implies that the probability that a secure device is purchased is 44% higher on average when compared to an unsecure device with the same price and number of functionalities.

Alternative	Price	Funct	ΔP
1	250	0	0.39
2	200	1	0.44
3	150	2	0.45
4	100	3	0.43

Table 4.9: Direct effect security on purchase probability

The second part of the main research question aims to discover whether the manner in which security is framed affects the purchase decision of consumers. The interaction effect of the framing variable and the security attribute is positive and significant. This result suggests that security has a stronger effect for respondents who received a gain focused description of security. In other words, communicating the gains of buying a secure device is more effective in nudging consumers towards buying more secure devices. This finding is in line with the findings of message framing literature that have been discussed in section 2.2.2. The studies found that people are more risk averse and thus more likely to opt for a preventative measure when faced with possible gains. In this case, buying a secure device is seen as this preventative measure. In line with the hypothesis, respondents are more likely to buy a secure device when faced with possible gains.

In order to illustrate the result, the effect of framing on the purchase probability is calculated for a sample set of four alternatives. For each alternative, ΔP is calculated by subtracting the purchase probability in the loss frame from the purchase probability in the gain frame. The values of the personal factors are set to their mean value of 0. The results are displayed in the table below. The calculations show that the framing of the security attribute has a relatively weak effect on the purchase probability. However, it is clear that the respondents who are receive a gain focused security description are more likely to buy a secure product, and less likely to buy an unsecure product.

Alternative	Price	Funct	Sec	ΔP
1	250	0	1	.020
2	150	2	1	.016
3	100	3	-1	-.019
4	200	1	-1	-.013

Table 4.10: Effect of framing on purchase probability

Moreover, the question targeted the sensitivity of the effect of security to personal factors. In section 4.3, three factors have been constructed: Security/Privacy Consciousness, Technology Acceptance and Conservativeness. The Technology Acceptance factor negatively moderates the effect of the price and security attribute and positively moderates the effect of the functionality attribute. This implies that respondents with a high score on this factor are willing to make concessions on price and security in order to make use of devices that provide them with innovative functionalities. The Privacy/Security Consciousness factor positively moderates the effect of security, which indicates that security contributes more strongly to the value of an alternative for respondents who are more aware of privacy and security risks of smart thermostats and actuate upon those risks. This result suggests that consumers can be nudged towards buying more secure devices when by improving their risk awareness and ensuring that they act upon this awareness. Finally, the role of security in the decision-making process is relatively small for respondents with a high score on the conservativeness factor. A possible explanation for this result is that respondents who do not value innovation strongly possess less technical knowledge and therefore are less likely to take security into account when purchasing devices.

4.7 Willingness to Pay (WtP)

The effect of security in the resulting model is notably strong when compared to the other device attributes. To illustrate this, the WtP value for security is calculated. According to (4), the willingness to pay for a certain variable can be calculated by dividing the derivative of the utility function with respect to this variable with the derivative of the utility function with respect to the price variable.

When making use of model 1.1 from model group 2, the WtP for security and functionality attributes can be calculated by dividing the parameter of these attributes by the parameter of the price attribute. The results of the analysis show an WtP for a secure product of €308,22. This figure is remarkably high, since it is higher than the price of most smart thermostats. A possible explanation for this result lies in the description of the security attribute. Due to the simple description of the security level, respondents strongly value this attribute. It is likely that respondents simply would never purchase a device that “can be hacked” or “is not secured properly”.

4.9 Conclusion

The causal diagram displayed below illustrates the various effects of the device attributes and their interactions with personal factors and framing on the choices of consumers to buy smart thermostats.

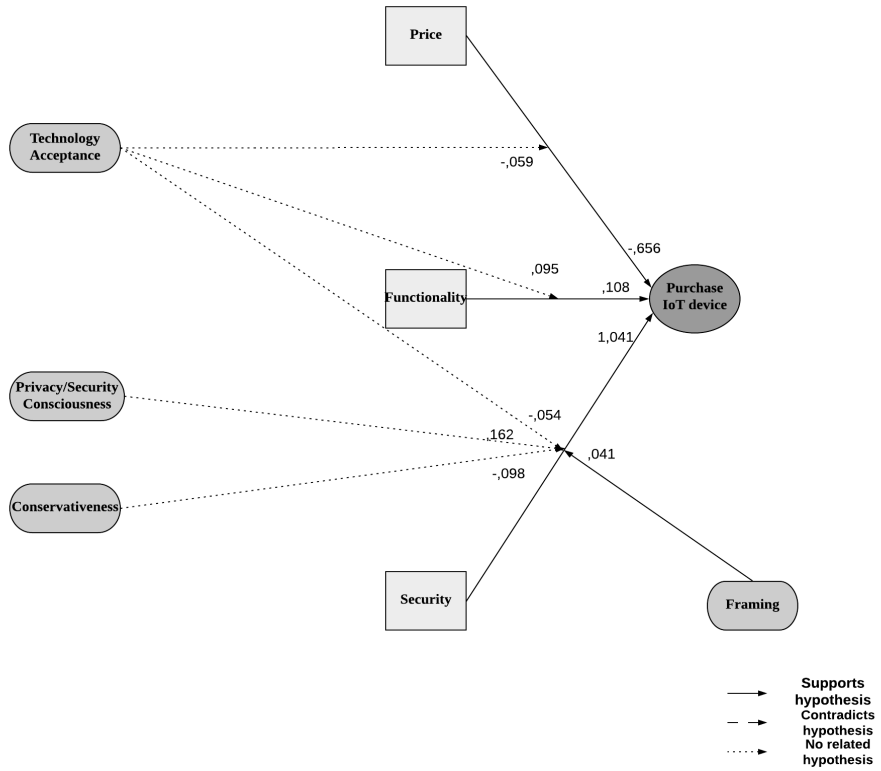


Figure 4.1: Causal diagram

In line with our hypotheses, the functionality and security of a device have a positive effect on the purchase probability, while the price of a device has a negative effect on this probability. The effect of security surprisingly high, possibly due to the binary coding of the attribute. It is likely that respondents generally do not want to buy a device that "is not secured properly" or "can be hacked", regardless of the functionality and price of the device. With regard to framing, the results found that a description of security that focuses on gains is more effective in nudging the respondents towards buying more secure devices. This finding is in line with the findings of message framing literature, which concluded that people are more risk averse when faced with gains and are therefore more likely to opt for a preventative measure. In this case, buying a secure product is the preventative measure. Finally, the study investigated whether the effect of the device attributes differed within the subgroups of the population that are characterised by the values of personal factors. Firstly, respondents are among the first adopters of innovative technologies are willing to make concessions on price and security in order to make use of innovative functionalities. Secondly, respondents who are aware of security and privacy risks and actuate upon this knowledge are more likely to purchase secure devices. Finally, security has a less strong effect for respondents who do not value innovative technologies strongly.

Chapter 5 Results: Qualitative study

In this chapter, the results of the qualitative study are presented. The main goal of the study is to investigate the underlying rationales that determine the effect of security and privacy on the choices of consumers. Firstly, the results of the questions regarding the purchase decision of consumers are analysed. Secondly, the risk awareness of the respondents is dealt with. Finally, the assessment of the scenarios by the respondents is investigated.

5.1 Response

A total of 27 responses have been provided to the survey. The aim of this study is to explore which underlying rationales affect the role of security and privacy in IoT device purchase behaviour. For this reason, the conclusions from the study are not generalised to a larger population and no statistical tests are executed to test the representativity of the collected sample. However, it is desirable to briefly discuss the demographic characteristics of the collected sample. The descriptive statistics of the demographic variables have been visualised in the figure below.

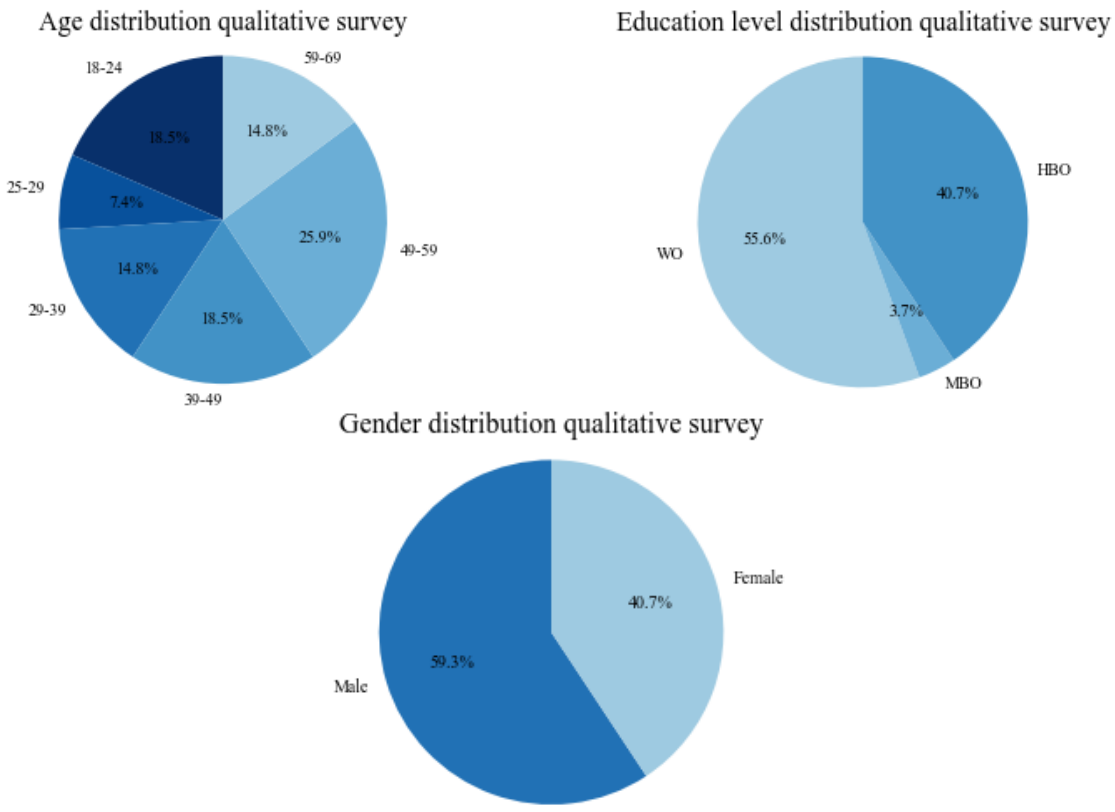


Figure 5.1: Demographics qualitative survey

The figure shows that the sample contains a substantive amount of responses from each age group and gender. However, the collected sample mainly consists of respondents with a high education level. This overrepresentation is deemed not to be problematic due to the qualitative nature of the study.

5.2 Purchase decision

Security and privacy were only mentioned twice as a reason not to purchase a smart thermostat. P8 did not purchase a smart thermostat due to concerns regarding the "storage of personal data". The other non-device owners indicated that they did not purchase the device because of the limited added value of the device or practical reasons such as the age of their home. When asked whether security and privacy played a role in their decision not to buy a smart thermostat, 15 of the 19 non-device owners answered "no". Most respondents did not actively think about privacy or security issues of the devices. P15 explained that he/she did not perceive smart thermostats as valuable in general, and therefore did not consider any privacy or security risks. However, security and privacy did play a role in the decision not to buy a smart thermostat for P22 and P23. P22 "wants to limit the sharing of data with third parties as much as possible, while P23 indicated that "we are being followed enough already, I do not want to increase my data footprint".

For device owners, the reasons to purchase a smart thermostat were mainly focused around the functionalities the device provides, ease of use and energy cost reductions P1 and P9 specifically mentioned the "remote control" functionality as the reason to purchase a smart thermostat. With regard to the decision to purchase a specific smart thermostat, the compatibility with other devices such as the boiler, voice assistants and smart home devices were indicated frequently. In addition to this, P9 mentioned that the reliability of the brand played a significant role in his/her decision to purchase a specific smart thermostat. 3 of the 7 device owners indicated that security did not play a role to purchase a smart thermostat. P9 and P15 indicated that they did not think about privacy and security when purchasing the device.

Device owners for whom security and privacy did play a role in the decision to buy a smart thermostat highlighted security and privacy risks such as sharing of personal data with third parties or adversaries gaining control over the device. The device owners were also asked to reflect on their considerations of privacy and security after they had purchased the device and whether they had undertaken any actions to mitigate these risks. 2 of the 7 device owners pointed out that they had considered security and privacy risks after the purchase. P1 described the development of Google as a tech giant as an important risk, while P3 mentioned the possibilities for burglary and the dependency of uptime. Finally, 4 device owners indicated that they undertook actions to mitigate security and privacy risks, the main actions were "safe passwords" and "selective permission to personal data".

From the results, it becomes clear that the respondents only started thinking about the role of security and or privacy in their decision to purchase or not to purchase a smart thermostat after being prompted with a question which specifically asks for the effect of security and privacy on their purchase decision. This result suggests that without receiving any information related to privacy and security, the respondents tend to focus on the value the device can bring to them, how it affects their costs and its compatibility with other devices. However, after being triggered to think about security and privacy, respondents are able to address some high-level concerns and risks related to these topics. Moreover, some device owners assert that they did have some concerns regarding privacy and security and undertook actions to mitigate the risks. However, the actions that are mentioned are not likely sufficient protection against harmful adversaries.

5.3 Risk awareness

12 of the 19 non-device owners indicated that they were aware of the security- and privacy risks of smart thermostats. The main risks that the respondents mentioned were lacking security controls, the use of data on the device for burglary purposes, insecure storage of personal data, sharing of personal data with third parties and targeted advertising. 3 of the 8 device owners were able to list some security and privacy risks of smart thermostats. The respondents mentioned high level description of risks such as “hacking” or “personal data going public”. The main takeaway regarding risk awareness is that respondents who do have some understanding of security and privacy risks are only able to give a high-level description of these risks using basic keywords such as “hacking”, “data leaks” or “lacking security”. However, only a small amount of the respondents provided a detailed description of a risk in terms of threat actors, probability and impact.

This shows that the respondents have some basic understanding of the security and privacy risks of smart thermostats and know that security and privacy are important topics in relation to innovative technology. However, their descriptions of these risks strongly lack any detail or reference to realistic threat events.

5.4 Scenarios

The assessment of scenarios allows for the generation of insights regarding the risk assessment of the respondents. The main goal of the analysis is to determine the underlying factors that influence this process rather than quantifying the effects of these factors. For this reason, the focus lies on analysing the motivations that the respondents have provided for their rating rather than quantitatively assessing the ratings per scenario.

5.4.1 Perception of security/privacy level

Firstly, the perceived security or privacy level is a frequently mentioned factor in the risk assessment process. A small number of respondents assume that the privacy or security risks are limited due to sufficient technical or institutional protections. For example, P15 scored the severity of the risk in scenario 2 as “very low” as he/she presumed that “my data is being handled fairly and in compliance with GDPR”. P15 added that he/she had “nothing to hide” and therefore the severity of the risk was very low in his/her opinion. In a similar fashion, P12 rated the severity of scenario 3 as “very low”, with the motivation that “GDPR has been put in place to mitigate this risk”. P13 rated the severity of scenario 1 as “very low” and gave the following motivation: “I may be a bit naïve, but I assume that my information is protected adequately”. These responses indicate an overestimation of security and privacy with regard to smart thermostats. In the case of the privacy risks, some respondents assume that their privacy is sufficiently safeguarded by a regulation which has been put in place. However, practice has taught us that this is certainly not always the case, as many privacy infringements have occurred, and many large fines have been issued since the introduction of GDPR. On the contrary, most respondents indicated that they perceived the security level of smart thermostats or IoT/IT devices in general to be low. P11 gave a “very high” rating to scenario 1, with the motivation that “information is mostly secured inadequately”. P12 also rated the scenario as “very high”, stating that “everything can be hacked these days”. P8 rated scenario 3 as “very high and expressed that “when money is involved, people do not care about privacy”

5.4.2 Probability of occurrence

Secondly, the probability of occurrence of a scenario is mentioned by the respondents as a motivation for the rating of a scenario. P13 provided a rating of "very low" for scenario 1, as the scenario is "too far-fetched" in his/her opinion. P9 stated that he/she regarded the severity of scenario 4 to be "low", since he/she "did not see this happening". The probability of occurrence was also mentioned by respondents to motivate a high score regarding the severity of a scenario. P6 and P12 both rated the risk in scenario 1 to be of high severity, as "everything can be hacked these days". Similarly, P10 responded "it seems to be a likely scenario" when asked for a motivation regarding his/her high rating of scenario 1. Often, very realistic scenarios are judged to be unlikely to occur. This indicates that some respondents assume that these risks are not likely to affect them, which makes them unlikely to value security strongly when purchasing devices. Another explanation for this result is limited knowledge regarding security and privacy. If a respondent does not possess knowledge related to security and privacy, it seems logical that the respondent would not assess many of the security and privacy related risks as likely to occur.

5.4.3 Third party benefits

Thirdly, the benefits that can be achieved by the parties involved in the scenario play a role in determining the rating of a scenario. More specifically, potential financial benefits are mentioned frequently as a motivation to rate the severity of a scenario as "high". P8 rated scenario 1 as "high", for the reason that "third parties are able to earn money". In a similar fashion, P10 rated the severity of scenario 3 as high, with the motivation that "this could be a potential revenue model". For scenario 4, some respondents stated that the severity was "low", as they deemed that gaining control over the smart thermostat in order to adjust temperatures would not be of high benefit to an adversary. P1 rated the scenario as "very low", as "a heating system is not interesting for adversaries". P15 stated the following: "Why would an outsider want to regulate the temperatures in my house?". This shows that the potential benefits for third parties play a role in determining the perceived severity of a risk. If respondents are not aware of these benefits, they might underestimate the risk that is presented to them.

5.4.4 Impact

When asked what scenario describes the most severe risks, many respondents mentioned a new factor, the consequences or impact of a risk. For P4, P5, and P16 scenario 1 describes the most severe risk, providing motivations such as "this is a personal impact with the possible consequence that objects will be stolen which are related to valuable memories" or "this has the highest personal impact". Scenario 1 is the only scenario in which actual physical assets can be at stake, rather than informational or financial assets. It seems that for some respondents, the type of asset is that can be harmed is of high importance when assessing risks. In this case, the harm of physical objects is deemed more severe than the harm of privacy or financial assets. P4 indicated that these physical objects are more valuable to him/her because of their irreplaceability.

On the contrary, other respondents indicate that scenarios in which the confidentiality of certain information assets is at stake describe the most severe risk. For 7 respondents, scenario 5 describes the most severe risk. Scenario 5 is the only scenario in which a third party is able to gain access to other information than the information that is collected and stored in relation to the use of the smart thermostat. Many of these respondents mention the consequences of the

attack as a motivation for this assessment. P7 says that "By doing so, attackers can gain access to banking information and personal cloud information". Similarly, P15 mentioned that "access to information can have direct personal far reaching consequences such as identity fraud and theft of banking credentials".

Finally, 2 respondents were of the of the opinion that scenario 6 describes the most severe risk. Scenario 6 is the only scenario which describes a risk that has a clear societal impact rather than a personal impact. P6 recognised this and indicated that this scenario describes the most severe risk in his/her opinion, with the motivation that "shutting down large organisations has disastrous consequences for the economy". For P6, the societal impact of scenario 6 outweighs the personal impact of the other scenarios.

5.5 Conclusion

When asked about their decision to buy or not to buy a smart thermostat, the respondents do not mention privacy or security directly. From their responses, it seems that the functionalities, ease of use and cost reductions are the most prevalent factors that affect their decision to buy smart thermostats. However, when triggered to contemplate the role of privacy and security risks of their devices, some device owners did indicate that they had thought about some of these risks before purchasing the device. Furthermore, some respondents addressed concerns regarding privacy and security after the purchase decision had been made. Such concerns include the development of Google as a tech giant and the storage of personal information. Many respondents who own a smart thermostat undertook actions to mitigate the privacy and security risks of their device. Password protection is the most frequently mentioned mitigation measure. Although this measure does protect against simple attacks, it certainly does not provide sufficient protection against more sophisticated attacks. With regard to risk awareness, many respondents are able to list a few security and privacy risks of smart thermostats. However, their knowledge is often limited to basic keywords such as "hacking", "data leaks" and "personal information".

The risk assessment analysis has highlighted which factors determine how the respondents perceive the severity of a risk in a threat scenario. First of all, the perceived level of security and privacy plays a role in determining the perceived severity of the risk. Many respondents assume that they are sufficiently protected against certain risks because mitigation measures are in place. Secondly, respondents often mention the probability that a scenario occurs as an important factor. If a scenario seems to be unlikely for a respondent, he/she is likely to rate the severity of the risk is rated as "low". Furthermore, the possible benefits for third parties can be a deciding factor for the rating. The scenarios in which third parties are able to gain financial benefits are often rated as "high" or "very high", as financial gain is perceived a strong motivation. Finally, the impact of a scenario can be a determining factor. Three types of impact are relevant: Physical impact, information impact and societal impact.

Chapter 6 Discussion

In this chapter, the results of the study are discussed. Firstly, the results from the quantitative and qualitative study are used to answer the main research question. Subsequently, the practical and scientific implications of the results are evaluated. Finally, the limitations of the study and possibilities for further research are highlighted.

6.1 Conclusions

The adoption of IoT devices by consumers is ever-increasing. Given the significant security and privacy risks of these devices, it is desirable to nudge users towards buying more secure devices and taking their privacy into account when purchasing such devices. In order to do so, insights should be attained regarding the decision-making process of consumers when purchasing the devices. Currently, little is known about how security and privacy affect the decision of consumers to purchase an IoT device. This study has targeted this knowledge gap by answering the following research question:

"How do security and privacy influence the choice of consumers to buy an IoT device? And how sensitive is the effect of security and privacy to framing and personal factors?"\

The study has taken a mixed methods approach towards answering this research question. The first part of the study focused on quantitatively determining the effect of security on the choice of consumers to buy an IoT device, while the second part targeted the underlying rationales that determine the influence of security and privacy on this decision-making process.

A stated choice experiment was conducted for the quantitative study. In this experiment, the respondents were asked to make purchase decisions from predefined choice sets. From the collected data, MNL choice models have been developed. In line with the hypotheses, positive parameters are found for the effects of the functionality and security attribute. This implies that the probability that a device is purchased increases as the number of functionalities grows or the security of the device improves. On the contrary, this probability is lower for devices with a higher price. The effect of security on the choice behaviour is remarkably high when compared to the effects of the other device attributes.

The qualitative study took a significantly different approach, by targeting the underlying rationales that determine how security and privacy affect the choices of consumers. The results of this study strongly differ from the quantitative study. When asked which factors had influenced their decision to buy or not to buy a smart thermostat, close to none of the respondents mentioned privacy or security concerns. Ease of use, functionality, energy cost reduction and compatibility with other devices were most frequently mentioned as a motivation to buy the devices.

The main difference between the methodology of both studies is that the respondents in the quantitative study are triggered to think about security, while this is not the case in the qualitative study. Moreover, the respondents in the quantitative study are presented with an easily understandable description of the security level of the device, which allows them to easily compare alternatives with regard to security. It is likely that this is not the case in real world situations. To conclude, it seems that security and privacy do play a large role in the decision of consumers to buy an IoT devices, under the condition that information regarding security and privacy is accessible, understandable and allows for simple and timely comparison of devices.

The second part of the research question aims to assess whether framing effects moderate the effect of security and privacy on the purchase decision of consumers. The results show that the respondents are more likely to purchase a secure device when faced with the possible gains of buying a secure device. This finding is in line the well-established hypothesis of Prospect Theory, which postulates that people are more risk averse when faced with the gains of outcomes in a choice situation. This result suggests that communicating the gains of buying a secure product is more effective in nudging users towards buying more secure devices.

In addition, the second part of the research question targets the sensitivity of the effect of privacy and security on consumer purchase behaviour with regard to personal factors. For this reason, the quantitative study has evaluated interaction effects of security with personal factors. A positive parameter was found for the interaction of the security attribute with the Privacy/Security Consciousness factor. This suggests that security has a stronger effect on the choice behaviour of respondents who are more aware of privacy and security risks and act upon this knowledge. Thus, improving the security and privacy awareness of consumers and ensuring that they act upon their awareness might lead towards consumers buying more secure devices.

The qualitative study also investigated the risk awareness of consumers. The results indicated that some consumers are able to list some of the security and privacy risks of smart thermostats. However, the descriptions of these risks strongly lack detail and are not specific for smart thermostats. Furthermore, the qualitative study examined the risk assessment process of consumers. From this analysis, a set of factors have been derived that were frequently mentioned as a motivation to assess the severity of a privacy or security related risk of smart thermostats. The following factors were found to be relevant: Perceived security and privacy level, probability of occurrence, third party benefits, and impact.

Finally, the quantitative study found a negative interaction effect of the technology acceptance factor with the price and security attributes and a positive interaction effect with the attribute functionality. This suggests that people who score high on this factor can be seen as the “first adopters” of innovative technologies and are more willing to buy less secure and more expensive products that do provide them with new functionalities and improve their quality of life.

To conclude, the study has found that security and privacy can have a strong effect on the purchase decision of consumers, under the specific circumstances that privacy and security related information is easily available and communicated in an understandable manner that allows for comparison of alternative devices in a simple and timely manner. The effect of security is moderated by the Privacy/Security Consciousness, Technology Acceptance and conservativeness of consumers. Finally, the results show that security related information that focuses on the gains of security is more effective in nudging consumers towards buying more secure devices.

6.2 Implications

6.2.1 Practical implications

This study has significant practical implications for stakeholders in the field of IoT security and privacy. The results of the quantitative study show that security has a strong effect on the probability that an IoT device is purchased. A possible explanation for this result lies in the methodology of the study. In the stated choice experiment that was conducted for the quantitative study, the description of security in the was kept extremely simple. For example, an unsecure device in the gain frame was given the following description: "This device is not protected properly". Similarly, an unsecure device in the loss frame was given the following description: "The device can be hacked". It is likely that these descriptions had such a strong effect on the choices of the respondents because of their simplicity and understandability. In this line of reasoning, security has a strong effect on the purchase decision of consumers under the condition that the security related information is communicated in such a way that consumers are able to easily understand the risks they face when purchasing the device. Another explanation for this result is that the simplicity of the security attribute enabled the respondents to easily compare two alternatives in the stated choice experiment with regard to security. To illustrate this, the difference between a device that "can be hacked" and a device that "cannot be hacked" or a device that "is secured properly and a device that "is not secured properly "in terms of security is evident. On the contrary, the respondents in the qualitative study did not mention security or privacy when asked to contemplate their decision to buy or not to buy a smart thermostat. The respondents mentioned other factors, such as the functionalities the device is able to provide, cost reductions and ease of use.

From these results, it can be argued that governmental bodies could effectively nudge consumers towards buying more secure devices by ensuring that security or privacy related information is communicated towards consumers. Furthermore, it is crucial the information is communicated in an understandable manner which allows consumers to easily compare devices with regard to security and privacy. Governmental bodies could work towards this goal by defining standards or legislation that describe what security and privacy related information should be provided to consumers and how this information should be communicated.

Since IoT security is a complex topic from a technical, institutional and organisational point of view, it is challenging to convey a message regarding security or privacy in such a way that it is understandable for consumers and allows for comparison of devices in a simple and timely manner. Moreover, the threat-landscape of IoT devices is ever-changing, as adversaries are continuously searching for novel vulnerabilities in the devices themselves or the software that is used to control the devices. Even if producers are able to structurally improve their devices with regard to security by implementing security controls such as remote or automatic patching, this provides no guarantee that the device is actually secure, since adversaries are often able to find new ways to inflict harm. For this reason, it seems sensible for governmental bodies to stimulate the involvement of market parties, such as manufacturers or retailers. Through collaboration between private- and public parties, knowledge can be shared in order to adequately deal with the complexity of IoT security and privacy. Existing frameworks that conceptualise security and privacy of IoT devices can leveraged as a starting point for such collaborations.

Furthermore, the results of the indicate that consumers who are more aware of privacy and security risks are more likely to consider security and privacy when purchasing IoT devices. Thus, improving the risk awareness of consumers supports the goal of nudging users towards

buying more secure devices and taking their privacy into account in their purchase decision. In order to reach this goal, governmental bodies could initiate awareness programs that specifically focus on communicating security and privacy risks of IoT devices to consumers. The results of the qualitative study have provided initial insights into what factors of these risks might have an influence on the risk perception of consumers. The respondents mentioned perceived security and privacy, probability of occurrence, third party benefits, and impact as a motivation for their assessment of the severity of a risk. These factors could form the basis of governmental efforts to improve the privacy and security risk awareness of consumers with regard to IoT devices. Finally, the results of the quantitative study revealed that security has a relatively weak effect on the purchase decision for the first adopters of innovative technologies. Moreover, this group of consumers is willing to make concessions on price and security for novel and innovative functionalities. Thus, the first adopters of innovative technologies can be identified as a focus group for security and privacy awareness programs.

6.2.2 Scientific implications

The results of the study also have implications from a scientific point of view. Firstly, the study has implications with regard to its methodology. By purposely varying the framing of attributes in a stated choice experiment, the study contributes to the discussion whether framing effects are relevant for studies using this methodology**.** Most studies that contribute to this discussion have found that framing has a significant effect on the model parameters and indicators. For example, Howard & Salkeld (2009) investigated the effects of attribute framing in the medical sector and found that framing had a significant influence on the model parameters and indicators such as the Willingness-to-Pay. Similarly, Veldwijk et al. (2016) investigated the effect of gain/loss framing in a stated choice experiment regarding participation in genetic screening for colorectal cancer. The results showed that framing had a notably strong effect of the decision-making process of respondents. 56% of the respondents who were faced with survival rates of the respondents ranked survival as the most important attribute, while only 8% of the respondents who faced mortality rates ranked mortality as the most important attribute. In line with this thinking, a study by Kragt & Bennett (2012) concluded that gain/loss framing has a significant effect on the model parameters that resulted from a stated choice experiment targeted at catchment management in Tasmania, Australia. The results of this study support the hypothesis that framing, more specifically gain/loss framing, has a significant effect on choice model parameters and indicators. This further strengthens the argument that studies using stated choice experiments should take framing effects into account. Moreover, the study shows that stated choice experiments can be used as a method to investigate the effects of framing on choice behaviour. In current studies, framing effects are often evaluated by presenting research subjects with a single choice task. By means of stated choice experiments, the research subject can be presented with multiple choice tasks. This lowers the standard errors of the estimated parameters, thus improving the validity of the developed models. Additionally, the lower standard errors limit the required amount of responses that have to be collected in order to estimate valid models. Finally, stated choice experiment allows for the effect of framing on various model parameters.

In addition, this study contributes to the TAM field. Existing studies in this field have found that security and privacy shape the intention of consumers to make use of innovative products and services, such as online banking. This study confirms the findings of these studies, as the results indicate that security and privacy play a role in the decision to purchase IoT devices. However, this study strongly differs from the studies in the TAM field with regard to the dependent variable in its causal model. The acceptance of a technology is the dependent

variable for studies in the TAM field. In many cases, this variable is measured by asking respondents for their use or intention to use the technology. For this study, more specifically the quantitative study, the dependent variable is the choice for a specific device. Thus, this study observes choices under the assumption that the respondents are willing to use and thus accept the device, rather than measuring acceptance of technologies.

The measurement of the dependent variable also differs from existing studies. The studies in the TAM field collect one choice observation per respondent to validate their causal model. In this study, the dependent variable is measured by means of a stated choice experiment in which the respondents are asked to make several choices from predefined choice sets. This allows for a more precise measurement of the effects of variables on the acceptance or purchase of innovative technologies. Moreover, it allows the researcher to estimate interaction effects of the device attributes with various explanatory variables such as personal factors, demographic variables or frames.

6.3 Limitations

6.3.1 Stated choice experiment

A stated choice experiment has been used to collect information regarding the choice behaviour of consumers. A limitation of stated choice experiments is that stated choices are observed, rather than actual choices. It is possible that respondents exhibit significantly different choice behaviour in the "real world" than during a stated choice experiment. For example, security might not play such a strong role in the decision-making process of consumers in the case of real-world purchases.

Moreover, the number of attributes that vary per alternative is strongly limited in a stated choice experiment, as the required sample size that is needed to draw valid conclusion increases for each added attribute. A high number of attributes also introduces more complexity into the survey, which might limit the response and validity of the results. For this reason, the alternatives in the stated choice experiment varied on a small set of three attributes: Price, Functionality and Security. Although these attributes are found to have a strong direct effect on the choice behaviour of respondents, it is likely that there exist other device attributes which might have a strong direct effect, such as energy cost reduction, ease of use, or usefulness. The effects of these device attributes have not been investigated by the study. Finally, the respondents are asked to make choices between a set of two alternatives. These choice situations do not resemble real-world situations, in which consumers make choices between a larger set of alternatives. Moreover, it is possible that some alternatives are not available to consumers in real life because of limited budgets or because some alternatives are not known to them.

Limitations can also arise from the specific coding of the device attributes. In this case, the operationalisation of the security attribute has its drawbacks. The security attribute has been varied on two levels. It is possible that this coding has led to an overestimation of the effect of security on the choice behaviour. To illustrate this, it seems sensible that consumers simply not willing to purchase a device that "can be hacked", regardless of the functionalities it provides to them or its price. The calculated WtP value of €308,22 supports this argument. The value of the indicator suggests that security contributes extremely strong to the utility of alternative when compared to the contribution of the price attribute. It is likely that a more feasible value for this indicator would have been found if the security attribute was coded on three or four levels. However, it is challenging to describe security on more than two levels in such a way

that the respondents in the stated choice experiment are able to easily compare the security level of two alternatives. For example, security could have been coded on three levels by using the following description: “The security level of this device is weak/moderate/strong”. In this case, it is questionable for respondents what the difference is between a weak, moderate or strong security level.

6.3.2 Discrete Choice Modelling: MNL

MNL models have been used to construct the causal model resulting from the quantitative study. These models can be used to analyse the direct effects of attribute- and interaction effects in a simple and elegant manner. However, MNL models do have some significant limitations. MNL models assume that the error terms in the utility function are i.i.d., which implies that these error terms have the same distribution and are mutually independent. This assumption is based upon the underlying assumption that the complete variation in utility across alternatives and alternatives is captured by the utility function. These assumptions become problematic when two alternatives have some aspects in common which are not captured by the attributes in the utility function. In such cases, their error terms are correlated. Thus, MNL models incorrectly assume that the error terms are i.i.d. This results in incorrect choice probabilities and counterintuitive substitution patterns, which create bias in the estimated parameters. MNL models assume that tastes for certain attributes (e.g. Price, Security level, Functionality) are equal within the target population. However, it seems highly sensible that tastes would differ across people. MNL models are able to deal with this issue to some extent by introducing interaction effects attributes with demographics and personal factors. Even so, MNL models still fail to capture taste heterogeneity within these subgroups of the target population. Thirdly, MNL models assume that every observed choice is independent of the other choices in the dataset. This seems to be illogic, since it can be expected that the choices made by the same individual are correlated.

Another important caveat of MNL models is the notion of linear-additive utility maximisation. This notion might limit the validity of the resulting models, since it is likely that people use different decision rules than utility maximisation. Moreover, the notion assumes that people exhibit fully compensatory behaviour, which implies that the willingness to trade-off between two attributes is independent of the relative performance of alternatives in terms of these attributes. The notion also ignores choice set effects, which implies that evaluation of an alternative is independent on the presence and performance of other alternatives in the choice set. These two assumptions may be unrealistic.

Other modelling techniques are able to deal with the limitations of MNL models. For example, mixed logit (ML) models introduce an extra error term into the utility function, allowing tastes to vary within the population and taking panel effects into account. Random Regret Minimisation (RRM) models introduce a new decision rule than utility maximisation. RRM models assume that people desire to minimise the regret after a decision has been made. Finally, Latent Class (LC) models allow for differentiation of decision rules by defining a set of latent classes. LC models introduce variation in decision rules by allowing the decision rules to vary between the classes. However, the main goal of this study is to show that certain relationships exist between device attributes, personal factors and framing rather than generating models with a high prediction power. MNL models are suited for this goal, as they allow for an easy and elegant analysis of direct effects of device attributes and interaction effects between the device attributes, personal factors and framing.

6.3.3 Qualitative study

For the qualitative study, a survey was used to reveal the underlying rationales that determine how security affects the choice behaviour of consumers. A survey allows for the generation of responses in a timely and costless manner. However, a using a survey for this goal has its limitations. When using a survey, the researcher is not able to ask to follow up questions when needed. In a semi-structured interview, the researcher can trigger the respondents to provide more in-depth responses by means of follow up questions. Thus, semi-structured interviews might have been more successful in reaching the main goal of the qualitative study. However, conducting semi-structured interviews was deemed unfeasible for the purpose of this MSc thesis due to time restrictions.

6.3.4 Literature study

The literature study concluded that no previous research exists that investigates the effect of security and privacy on the choices of consumers to buy IoT devices. However, during the course of this study, a research has been published which targets this exact research domain.

Emami-Naeini, Dixon, Agarwal, & Cranor (2019) researched how privacy and security factor into IoT device purchase behaviour by conducting a set of 24 semi-structured interviews and spreading a follow up survey to which 200 participants provided a response. The authors found that the respondents did not consider privacy and security before the purchase. However, the respondents indicated that they had become concerned with the privacy and security risks due to news reports, friends, or unexpected device behaviour. The respondents who had searched for privacy and security related information, explained that this information was often hard to find and excessively complex. The respondents were asked to rank certain factors they would take into consideration when purchasing IoT devices. Security was ranked the as the third most important factor, after the price and features. Finally, the respondents were asked to evaluate a set of privacy and security labels. Most of the respondents found the labels to be informative and useful.

Although the research is similar to this study, there exist some significant differences. Firstly, only consumers who actually purchased an IoT device were selected for the research. Consumers who chose not to buy an IoT device can be seen as equally interesting, since they might have decided not to purchase such devices for security or privacy reasons. These consumers can also be seen as potential end users of the devices, which makes it sensible to include them in the research. Secondly, the methodology of the research strongly differs from this study. In the research, the respondents were asked to rank a set of factors or attributes, while actual choices between alternatives that vary on these attributes are observed in this study. Thus, it can be argued that the validity of the results regarding the effect of the attributes on the choice behaviour are more valid in this study.

6.3.5 Framing

The application of framing to the security level in this study has its limitations. As became clear from the literature study, a large set of frames can be applied to messages. This study has only investigated the effect of the most prevalent frames in message framing literature: gain/loss framing. However, it is likely that the effect of security on choice behaviour is also sensitive to other types of framing, such as absolute/relative and relevancy. Moreover, it is questionable whether security and privacy can be framed as a pure gain. In the gain frame, the security level of the device was framed as “this device is/is not secured properly”. The concept

of "secured" still suggests that there exists some external threat against which the device should be secured. This external threat can be conceptualised as potential losses. Therefore, it can be argued that the applicability of gain/loss framing to security and privacy is relatively limited.

6.4 Further research

Although this study has contributed to filling the identified knowledge gap, additional efforts are needed to generate novel insights. Firstly, this study only investigated the effect of a limited set of three device attributes. Privacy was not included as a device attribute in this study. In order to assess whether similar conclusions hold for privacy, future research could build upon this study by including privacy and other device attributes. Secondly, the security attribute was coded as a binary variable in this study. In order to gain more insights into how security affects the purchase decision, other operationalisations could form the basis of future research. This allows us to examine what operationalisation of security has the strongest effect on choice behaviour and is most effective in nudging consumers towards buying more secure devices. Thirdly, this study has observed stated choices rather than real-world choices. It can be argued that the validity of the models that have been developed from this data is limited, as real-world decisions may strongly differ from decisions that have been made in the context of a stated choice experiment. Therefore, it is key to assess whether similar conclusions are drawn when developing models from real-world choice data. In order to reach this goal, further research could use revealed choice data as an input for the development of choice models. For example, activity on web shops could be monitored to collect data regarding the purchase behaviour of consumers. Finally, the recommendation was given to develop standards or legislation that describe how security and privacy related information should be communicated. However, the development process of such legislation or standards is highly complex and involves various stakeholders with conflicting interests. In order to successfully create such legislation or standards, these challenges should be overcome. Future research could work towards this goal by taking a design science approach that aims to design a process for the involvement of multiple stakeholders in this development process.

References

Abomhara, M., & Køien, G. M. (2014). Security and privacy in the Internet of Things: Current status and open issues. *2014 International Conference on Privacy and Security in Mobile Systems (PRISMS)*, 1–8. IEEE.

Arias, O., Wurm, J., Hoang, K., & Jin, Y. (2015). Privacy and security in internet of things and wearable devices. *IEEE Transactions on Multi-Scale Computing Systems*, *1*(2), 99–109.

Bierlaire, M. (2018). *PandasBiogeme: a short introduction*.

Bohli, J.-M., Langendörfer, P., & Skarmeta, A. F. (2013). Security and privacy challenge in data aggregation for the iot in smart cities. *Internet of Things: Converging Technologies for Smart Environments and Integrated Ecosystems*, 225–244.

Cherdantseva, Y., & Hilton, J. (2013). A reference model of information assurance & security. *2013 International Conference on Availability, Reliability and Security*, 546–555.

Choe, E. K., Jung, J., Lee, B., & Fisher, K. (2013). Nudging people away from privacy-invasive mobile apps through visual framing. *IFIP Conference on Human-Computer Interaction*, 74–91. Springer.

Choi, N., Kim, D., Goo, J., & Whitmore, A. (2008). Knowing is doing: An empirical validation of the relationship between managerial information security awareness and action. Information Management & Computer Security, 16(5), 484-501.

CNET. (2019). Amazon and Google are listening to your voice recordings. Here's what we know about thate. Retrieved from https://www.cnet.com/how-to/amazon-and-google-are-listening-to-your-voice-recordings-heres-what-we-know/

de Ruijter, A., & Guldenmund, F. (2016). The bowtie method: A review. *Safety Science*, *88*, 211–218.

de Vries, G., Terwel, B. W., & Ellemers, N. (2014). Spare the details, share the relevance: The dilution effect in communications about carbon dioxide capture and storage. *Journal of Environmental Psychology*, *38*, 116–123.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS quarterly, 319-340.

Deloitte. (2017). Cyber crime costs Dutch SME sector €1 billion each year. Retrieved from https://www2.deloitte.com/nl/nl/pages/over-deloitte/articles/cyber-crime-costs-dutch-sme-sector-1-billion-each-year.html

Egelman, S., & Peer, E. (2015). Scaling the security wall: Developing a security behavior intentions scale (sebis). *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 2873–2882. ACM.

Emami-Naeini, P., Dixon, H., Agarwal, Y., & Cranor, L. F. (2019). Exploring how privacy and security factor into IoT device purchase behavior. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 534. ACM.

Fabian, B., Ermakova, T., & Lentz, T. (2017). Large-scale readability analysis of privacy policies. *Proceedings of the International Conference on Web Intelligence*, 18–25. ACM.

Gluck, J., Schaub, F., Friedman, A., Habib, H., Sadeh, N., Cranor, L. F., & Agarwal, Y. (2016). How short is too short? implications of length and framing on the effectiveness of privacy notices. *Twelfth Symposium on Usable Privacy and Security ({SOUPS} 2016)*, 321–340.

Goldsmith, R. E., & Hofacker, C. F. (1991). Measuring consumer innovativeness. *Journal of the Academy of Marketing Science*, *19*(3), 209–221.

Gubbi, J., Buyya, R., Marusic, S., & Palaniswami, M. (2013). Internet of Things (IoT): A vision, architectural elements, and future directions. *Future Generation Computer Systems*, *29*(7), 1645–1660.

Howard, K., & Salkeld, G. (2009). Does attribute framing in discrete choice experiments influence willingness to pay? Results from a discrete choice experiment in screening for colorectal cancer. Value in Health, 12(2), 354-363.

Kahneman, D., & Tversky, A. (2013). Prospect theory: An analysis of decision under risk. In Handbook of the fundamentals of financial decision making: Part I (pp. 99–127). World Scientific.Kobsa, A., & Teltzrow, M. (2004). Contextualized communication of privacy practices and personalization benefits: Impacts on users' data sharing and purchase behavior. *International Workshop on Privacy Enhancing Technologies*, 329–343. Springer.

Khan, B., Alghathbar, K. S., Nabi, S. I., & Khan, M. K. (2011). Effectiveness of information security awareness methods based on psychological theories. African Journal of Business Management, 5(26), 10862-10868.

Kragt, M. E., & Bennett, J. W. (2012). Attribute framing in choice experiments: how do attribute level descriptions affect value estimates?. Environmental and Resource Economics, 51(1), 43-59.

Kühberger, A. (1998). The influence of framing on risky decisions: A meta-analysis. *Organizational Behavior and Human Decision Processes*, *75*(1), 23–55.

Mahmoud, R., Yousuf, T., Aloul, F., & Zualkernan, I. (2015). Internet of things (IoT) security: Current status, challenges and prospective measures. *2015 10th International Conference for Internet Technology and Secured Transactions (ICITST)*, 336–341. IEEE.

Malenka, D. J., Baron, J. A., Johansen, S., Wahrenberger, J. W., & Ross, J. M. (1993). The framing effect of relative and absolute risk. *Journal of General Internal Medicine*, *8*(10), 543–548.

Midgley, D. F., & Dowling, G. R. (1978). Innovativeness: The concept and its measurement. Journal of consumer research, 4(4), 229-242.

Mineraud, J., Mazhelis, O., Su, X., & Tarkoma, S. (2016). A gap analysis of Internet-of-Things platforms. *Computer Communications*, *89*, 5–16.

Netscout. (2019). Netscout Threat Report. Retrieved from https://www.netscout.com/sites/default/files/2019-07/SECR_010_EN-1901 – NETSCOUT Threat Report 1H 2019 – Web.pdf

Odelu, V., Das, A. K., Khan, M. K., Choo, K. K. R., & Jo, M. (2017). Expressive CP-ABE scheme for mobile devices in IoT satisfying constant-size keys and ciphertexts. IEEE Access, 5, 3273-3283.

Oualha, N., & Nguyen, K. T. (2016). Lightweight attribute-based encryption for the internet of things. In 2016 25th International Conference on Computer Communication and Networks (ICCCN) (pp. 1-6). IEEE.

OWASP. (2018). OWASP Internet of Things project. Retrieved from https://www.owasp.org/index.php/OWASP_Internet_of_Things_Project%09

Pasha, M., Shah, S. M. W., & Pasha, U. (2016). Security framework for IoT systems. *International Journal of Computer Science and Information Security*, *14*(11), 99.

PwC. (2017). Consumer Intelligence Series: Protect.me. Retrieved from https://www.pwc.com/us/en/advisory-services/publications/consumer-intelligence-series/protect-me/cis-protect-me-findings.pdf

PwC. (2018). The Global State of Information Security Survey 2018. Retrieved from https://www.pwc.com/us/en/services/consulting/cybersecurity/library/information-security-survey.html

Renaud, K., & Shepherd, L. A. (2018). *How to Make Privacy Policies both GDPR-Compliant and Usable*. https://doi.org/10.1111/j.1365-2923.2008.03183.x

Roman, R., Zhou, J., & Lopez, J. (2013). On the features and challenges of security and privacy in distributed internet of things. *Computer Networks*, *57*(10), 2266–2279.

Sadeghi, A.-R., Wachsmann, C., & Waidner, M. (2015). Security and privacy challenges in industrial internet of things. 2015 52nd ACM/EDAC/IEEE Design Automation Conference (DAC), 1–6. IEEE.SANS. (2013). Corporate vs. Product Security. Retrieved from https://www.sans.org/reading-room/whitepapers/bestprac/paper/34237

Schaub, F., Balebako, R., & Cranor, L. F. (2017). Designing effective privacy notices and controls. *IEEE Internet Computing*.

Schneider, T. R., Salovey, P., Pallonen, U., Mundorf, N., Smith, N. F., & Steward, W. T. (2001). Visual and Auditory Message Framing Effects on Tobacco Smoking 1. *Journal of Applied Social Psychology*, *31*(4), 667–682.

Sicari, S., Rizzardi, A., Grieco, L. A., & Coen-Porisini, A. (2015). Security, privacy and trust in Internet of Things: The road ahead. *Computer Networks*, *76*, 146–164.

Singh, K. J., & Kapoor, D. S. (2017). Create Your Own Internet of Things: A survey of IoT platforms. *IEEE Consumer Electronics Magazine*, *6*(2), 57–68.

Singh, S., Sharma, P. K., Moon, S. Y., & Park, J. H. (2017). Advanced lightweight encryption algorithms for IoT devices: survey, challenges and solutions. *Journal of Ambient Intelligence and Humanized Computing*, 1–18.

Veldwijk, J., Essers, B. A., Lambooij, M. S., Dirksen, C. D., Smit, H. A., & de Wit, G. A. (2016). Survival or mortality: does risk attribute framing influence decision-making behavior in a discrete choice experiment?. Value in Health, 19(2), 202-209

Von Solms, R., & Van Niekerk, J. (2013). From information security to cyber security. *Computers & Security*, *38*, 97–102.

Wang, P., Chaudhry, S., Li, L., Li, S., Tryfonas, T., & Li, H. (2016). The Internet of Things: a security point of view. *Internet Research*.

Wilson, S., Schaub, F., Liu, F., Sathyendra, K. M., Smullen, D., Zimmeck, S., … Sadeh, N. (2018). Analyzing privacy policies at scale: From crowdsourcing to automated annotations. *ACM Transactions on the Web (TWEB)*, *13*(1), 1.

Xie, W., Jiang, Y., Tang, Y., Ding, N., & Gao, Y. (2017). Vulnerability detection in iot firmware: A survey. *2017 IEEE 23rd International Conference on Parallel and Distributed Systems (ICPADS)*, 769–772. IEEE.

Zhao, K., & Ge, L. (2013). A survey on the internet of things security. *2013 Ninth International Conference on Computational Intelligence and Security*, 663–667.

Appendix A:Survey design

In this appendix, the questions in the survey for both the quantitative- and qualitative study are displayed.

A.1 Quantitative survey

A.1.1 Demographics

Firstly, the respondents were asked a set of questions to measure the values of a set of demographic variables. This allows for the assessment of the representativity of the collected sample.

Demographic	Question	Categories
Age	"What is your year of birth?"	Open question
Gender	"What is your gender"?	Male/Female
Education level	"What is your highest level of education?"	Elementary/Basic/MBO/Havo/VWO/WO
Working situation	"What is your main daily occupation?"	Student/Paid work/Unemployed/Homemaker/Retired/Other

Table A.1: Demographic questions quantitative survey

A.1.2 Indicators

Secondly, the survey contained a set of indicators. The values for each indicator are measured by proposing asking the respondents to what extent they agree with a statement on a five-point Likert scale. The indicators and respective statements have been displayed below.

Indicator	Statement	Categories
I1: IT Trends	"I keep up with technological developments"	5-Scale Likert (1=completely agree, 5 = completely disagree)
I2: IT News	"I read the technology section when reading newspapers or visiting news websites"	5-Scale Likert (1=completely agree, 5 = completely disagree)
I3: Development IT products	"I find it interesting to follow the development of new IT products"	5-Scale Likert (1=completely agree, 5 = completely disagree)
I4: Economic importance innovation	"Innovation is important for economic development"	5-Scale Likert (1=completely agree, 5 = completely disagree)

I5: Societal importance innovation	“Investments in innovative technologies are important for society”	5-Scale Likert (1=completely agree, 5 = completely disagree)
I6: Technology adoption	“If a new IT product has been developed, I want to buy the first version”	5-Scale Likert (1=completely agree, 5 = completely disagree)
I7: Consideration Security risks	“I pay attention to the security risks of my IT devices”	5-Scale Likert (1=completely agree, 5 = completely disagree)
I8: Security concern purchase	“When purchasing an IT device, I consider the security risks of the device”	5-Scale Likert (1=completely agree, 5 = completely disagree)
I9: Importance security	“The security of my IT devices is important to me”	5-Scale Likert (1=completely agree, 5 = completely disagree)
I10: Importance privacy	“My personal information should be protected sufficiently”	5-Scale Likert (1=completely agree, 5 = completely disagree)
I11: Privacy awareness	“I keep track of which information is collected when using online services”	5-Scale Likert (1=completely agree, 5 = completely disagree)
I12: Concern security	“I am concerned with the security risks of my IT devices”	5-Scale Likert (1=completely agree, 5 = completely disagree)
I13: Concern privacy IT devices	“When using IT devices, I am concerned with the use of my personal data by external parties”	5-Scale Likert (1=completely agree, 5 = completely disagree)
I14: Concern privacy online services	“When using online services, I am concerned with the use of my personal data by external parties”	5-Scale Likert (1=completely agree, 5 = completely disagree)
I15: Action Security	“I undertook action to improve the security of my IT devices”	5-Scale Likert (1=completely agree, 5 = completely disagree)

Table A.2: Indicators quantitative study

A.1.3 Example choice sets

The survey included a set of choice sets to measures the effects of device attributes on the choice behaviour of the respondents. 8 example choice sets have been displayed below.

Choice set 1

Stelt u zich voor dat uw thermostaat kapot gaat en u genoodzaakt bent een slimme thermostaat te kopen en u de volgende twee opties heeft voor de aanschaf van dit product

	Optie A
Prijs	€150
Functionaliteit	Op afstand bestuurbaar
Beveiligingsniveau	Dit apparaat is goed beveiligd

	Optie B
Prijs	€150
Functionaliteit	Op afstand bestuurbaar, geofencing en sensoren
Beveiligingsniveau	Dit apparaat is niet goed beveiligd

14. Zou u optie A aanschaffen? ○ Ja ○ Nee

15. Zou u optie B aanschaffen? ○ Ja ○ Nee

16. Welke optie heeft uw voorkeur? ○ A ○ B

Figure A.1: Example choice set 1

Choice set 2

Stelt u zich voor dat uw thermostaat kapot gaat en u genoodzaakt bent een slimme thermostaat te kopen en u de volgende twee opties heeft voor de aanschaf van dit product

	Optie A
Prijs	€150
Functionaliteit	Op afstand bestuurbaar
Beveiligingsniveau	Dit apparaat is goed beveiligd

	Optie B
Prijs	€150
Functionaliteit	Op afstand bestuurbaar, geofencing en sensoren
Beveiligingsniveau	Dit apparaat is niet goed beveiligd

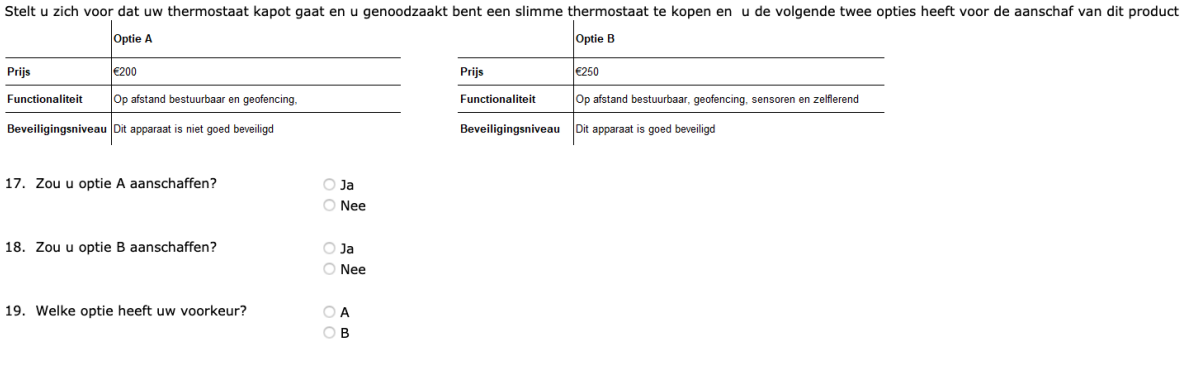
14. Zou u optie A aanschaffen? ○ Ja ○ Nee

15. Zou u optie B aanschaffen? ○ Ja ○ Nee

16. Welke optie heeft uw voorkeur? ○ A ○ B

Figure A.2: Example choice set 2

Choice set 3



Stelt u zich voor dat uw thermostaat kapot gaat en u genoodzaakt bent een slimme thermostaat te kopen en u de volgende twee opties heeft voor de aanschaf van dit product

	Optie A
Prijs	€200
Functionaliteit	Op afstand bestuurbaar en geofencing,
Beveiligingsniveau	Dit apparaat is niet goed beveiligd

	Optie B
Prijs	€250
Functionaliteit	Op afstand bestuurbaar, geofencing, sensoren en zelflerend
Beveiligingsniveau	Dit apparaat is goed beveiligd

17. Zou u optie A aanschaffen? ○ Ja ○ Nee

18. Zou u optie B aanschaffen? ○ Ja ○ Nee

19. Welke optie heeft uw voorkeur? ○ A ○ B

Figure A.3: Example choice set 3

Choice set 4

Stelt u zich voor dat uw thermostaat kapot gaat en u genoodzaakt bent een slimme thermostaat te kopen en u de volgende twee opties heeft voor de aanschaf van dit product

	Optie A
Prijs	€200
Functionaliteit	Op afstand bestuurbaar, geofencing en sensoren
Beveiligingsniveau	Dit apparaat is goed beveiligd

	Optie B
Prijs	€100
Functionaliteit	Op afstand bestuurbaar, geofencing, sensoren en zelflerend
Beveiligingsniveau	Dit apparaat is niet goed beveiligd

20. Zou u optie A aanschaffen? ○ Ja ○ Nee

21. Zou u optie B aanschaffen? ○ Ja ○ Nee

22. Welke optie heeft uw voorkeur? ○ A ○ B

Figure A.4: Example choice set 4

Choice set 5

Stelt u zich voor dat uw thermostaat kapot gaat en u genoodzaakt bent een slimme thermostaat te kopen en u de volgende twee opties heeft voor de aanschaf van dit product

	Optie A
Prijs	€100
Functionaliteit	Op afstand bestuurbaar
Beveiligingsniveau	Dit apparaat is goed beveiligd

	Optie B
Prijs	€250
Functionaliteit	Op afstand bestuurbaar en geofencing,
Beveiligingsniveau	Dit apparaat is niet goed beveiligd

23. Zou u optie A aanschaffen? ○ Ja ○ Nee

24. Zou u optie B aanschaffen? ○ Ja ○ Nee

25. Welke optie heeft uw voorkeur? ○ A ○ B

Figure A.5: Example choice set 5

Choice set 6

Stelt u zich voor dat uw thermostaat kapot gaat en u genoodzaakt bent een slimme thermostaat te kopen en u de volgende twee opties heeft voor de aanschaf van dit product

	Optie A
Prijs	€100
Functionaliteit	Op afstand bestuurbaar en geofencing,
Beveiligingsniveau	Dit apparaat is goed beveiligd

	Optie B
Prijs	€150
Functionaliteit	Op afstand bestuurbaar, geofencing, sensoren en zelflerend
Beveiligingsniveau	Dit apparaat is niet goed beveiligd

26. Zou u optie A aanschaffen? ○ Ja ○ Nee

27. Zou u optie B aanschaffen? ○ Ja ○ Nee

28. Welke optie heeft uw voorkeur? ○ A ○ B

Figure A.6: Example choice set 6

Choice set 7

Stelt u zich voor dat uw thermostaat kapot gaat en u genoodzaakt bent een slimme thermostaat te kopen en u de volgende twee opties heeft voor de aanschaf van dit product

	Optie A
Prijs	€150
Functionaliteit	Op afstand bestuurbaar, geofencing, sensoren en zelflerend
Beveiligingsniveau	Dit apparaat is niet goed beveiligd

	Optie B
Prijs	€150
Functionaliteit	Op afstand bestuurbaar
Beveiligingsniveau	Dit apparaat is goed beveiligd

29. Zou u optie A aanschaffen?
- ○ Ja
- ○ Nee

30. Zou u optie B aanschaffen?
- ○ Ja
- ○ Nee

31. Welke optie heeft uw voorkeur?
- ○ A
- ○ B

Figure A.7: Example choice set 7

Choice set 8

Stelt u zich voor dat uw thermostaat kapot gaat en u genoodzaakt bent een slimme thermostaat te kopen en u de volgende twee opties heeft voor de aanschaf van dit product

	Optie A
Prijs	€250
Functionaliteit	Op afstand bestuurbaar en geofencing,
Beveiligingsniveau	Dit apparaat is niet goed beveiligd

	Optie B
Prijs	€200
Functionaliteit	Op afstand bestuurbaar
Beveiligingsniveau	Dit apparaat is niet goed beveiligd

32. Zou u optie A aanschaffen?
- ○ Ja
- ○ Nee

33. Zou u optie B aanschaffen?
- ○ Ja
- ○ Nee

34. Welke optie heeft uw voorkeur?
- ○ A
- ○ B

Figure A.8: Example choice set 8

A.2 Qualitative study

A.2.1 Demographics

The qualitative survey contained a set of questions aimed at measuring the values for demographic variables. These questions have been presented below.

Demographic	Question	Category
Age	"What is your birth year?"	Open question
Gender	"What is your gender"?	Nominal (Male/Female)
Education level	"What is your highest level of education?"	Ordinal (Elementary/Basic/MBO/Havo/VWO/WO)

Table A.3: Demographic questions qualitative study

A.2.2 Purchase decision

Subsequently, the respondents were asked to contemplate their decision to purchase or not to purchase a smart thermostat. These questions have been listed in the table below.

Respondent group	Question	Category
Device owners	Why did you choose to purchase a smart thermostat?	Open
Device owners	Which smart thermostat do you own?	Open
Device owners	Did security or privacy play a role in your decision to buy a smart thermostat? Why (not)?	
Device owners	Did security or privacy play a role in your decision to buy this specific smart thermostat? Why (not)?	
Device owners	Why did you choose for this specific smart thermostat?	Open
Non-device owners	Why did you choose not to buy a smart thermostat?	Open
Device owners	Did you consider any privacy or security risks of your smart thermostat after purchasing the device? If yes, which ones?	Open
Device owners	Did you undertake any actions to mitigate the privacy or security risks of your device?	Open

Table A.4: Purchase decision questions qualitative survey

A.2.3 Risk awareness

In the survey, the respondents were asked whether they were aware of any security or privacy related risks of smart thermostats. This question is displayed in the table below.

Respondent group	Question	Category
Device owners/non-device owners	Do you own a smart thermostat?	Open

Table A.5: Risk awareness question qualitative survey

A.2.4 Scenarios

Finally, the respondents were asked to evaluate a set of hypothetical scenarios which describe a privacy or security risk of smart thermostats. Per scenario, the respondents were asked to assess the severity of this risk and provide a motivation for the assessment. Moreover, they were asked which of the scenarios describes the most severe risks in their opinion and whether the risks would have an influence on their decision to buy or not to buy a smart thermostat. The questions are presented in the table below.

Respondent group	Category
How high do you rate the severity of the risk that is described in the scenario? Why?	5-Scale Likert (1=completely agree, 5 = completely disagree) / Open
Which of these scenarios describes the most severe risk in your opinion? Why?	Nominal (Scenario 1/Scenario 2/Scenario 3/Scenario 4/ Scenario 5) / Open
Would these risks influence your decision to buy or not to buy a smart thermostat? Why?	Nominal (Yes,No) / Open

Table A.6: Scenario questions qualitative study

Appendix B: Experimental design

B.1 Basic plan 3

The basic plan which has been used to develop the choice sets in the stated choice experiment has been displayed below. For each attribute in the experiment, a column has to be chosen from the basic plan. The final rows of the basic plan contain rules for combining columns within a design. For this study, row "1*" is used for the price of an alternative. Row "3*" is used for the functionality attribute. Finally, row "04" is used for the security level of the alternatives.

BASIC PLAN 3: 4^5; 3^5; 2^{15}; 16 trials

1	2	3	4	5	1	2	3	4	5	0	0	0	0	0	0	0	0	0	1	1	1	1	1
*	*	*	*	*	*	*	*	*	*	1	2	3	4	5	6	7	8	9	0	2	3	4	5
										x	x	x	x										
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	1	1	2	3	0	1	1	2	1	0	0	0	0	1	1	0	1	1	1	1	1	1	0
0	2	2	3	1	0	2	2	1	1	0	0	0	1	0	1	1	0	1	1	0	0	1	1
0	3	3	1	2	0	1	1	1	2	0	0	0	1	1	0	1	1	0	0	1	1	0	1
1	0	1	1	1	1	0	1	1	1	0	1	1	0	0	0	0	1	1	0	1	0	1	1
1	1	0	3	2	1	1	0	1	2	0	1	1	0	1	1	0	0	0	1	0	1	0	1
1	2	3	2	0	1	2	1	2	0	0	1	1	1	0	1	1	1	0	1	1	0	0	0
1	3	2	0	3	1	1	2	0	1	0	1	1	1	1	0	1	0	1	0	0	1	1	0
2	0	2	2	2	2	0	2	2	2	1	0	1	0	0	0	1	0	1	1	1	1	0	1
2	1	3	0	1	2	1	1	0	1	1	0	1	0	1	1	1	1	0	0	0	0	1	1
2	2	0	1	3	2	2	0	1	1	1	0	1	1	0	1	0	0	0	0	1	1	1	0
2	3	1	3	0	2	1	1	1	0	1	0	1	1	1	0	0	1	1	1	0	0	0	0
3	0	3	3	3	1	0	1	1	1	1	1	0	0	0	0	1	1	0	1	0	1	1	0
3	1	2	1	0	1	1	2	1	0	1	1	0	0	1	1	1	0	1	0	1	0	0	0
3	2	1	0	2	1	2	1	0	2	1	1	0	1	0	1	0	1	1	0	0	1	0	1
3	3	0	2	1	1	1	0	2	1	1	1	0	1	1	0	0	0	0	1	1	0	1	1

1				2				3				4				5		
-	0	0	0	-	0	0	0	-	0	0	0	-	1	1	1	-	1	1
*				*				*				*				*		
-	1	2	3	-	4	5	6	-	7	8	9	-	0	1	2	-	3	5

Figure B.1: Basic plan 3

B.2 Profiles

By making use of basic plan 3, a set of 16 profiles has been constructed. The 16 profiles have been listed in the table below.

Profile Nr.	Price	Functionality	Security
1	0	0	-1
2	0	1	-1
3	0	2	1
4	0	3	1
5	1	0	-1
6	1	1	-1
7	1	2	1
8	1	3	1
9	2	0	1
10	2	1	1
11	2	2	-1
12	2	3	-1
13	3	0	1
14	3	1	1
15	3	2	-1
16	3	3	-1

Table B.1: Profiles

B.3 Choice sets

The profiles have been randomly assigned to a set of 16 choice sets. These choice sets have been divided into two blocks in order to limit the amount of choices per individual in the experiment. The choice sets and their corresponding blocks have been displayed in the table below.

Choice set Nr.	Block	Alternative 1	Alternative 1
1	1	8	10
2	1	9	11
3	1	6	4
4	1	7	16
5	1	13	2
6	1	14	12
7	1	12	9
8	1	2	5
9	2	4	15
10	2	11	13

11	2	16	3
12	2	3	14
13	2	10	7
14	2	15	8
15	2	1	6
16	2	5	1

Table B.2: Choice sets

Appendix C: Representativity

First of all, it is important to test whether the mean of the age in the sample is representative for the population. The hypotheses for this test are as follows:

H0: "The average age in the sample is equal to the average in the population"

H1: "The average age in the sample is not equal to the average age in the population"

Since we are evaluating differences between two groups on a single variable on an interval scale, the one sample t-test is used. The t-statistic in this test is calculated by the following formula:

$$T = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} \tag{7}$$

Where

$\bar{x}$ = The mean of the sample
μ = The mean of the population
s = The standard deviation of the sample
n = The sample size

The p-value resulting from this test is equal to 0.045, which indicates that H0 should be rejected for the 0.05 significance level. Thus, there exists a significant difference between the average age in the sample and the average age in the population. The difference is equal to -1.5 years, which indicates that the average age in the sample is 1.5 years lower than the average in the population.

In addition to testing the mean difference, the difference in age distribution between the sample and the population should be assessed. The hypotheses for this test are as follows:

H0: "The age distribution in the sample is equal to the age distribution in the population"

H1: "The age distribution in the sample is not equal to the age distribution in the population"

The distribution of age in the population is defined in terms of age groups. Therefore, the test involves comparing two groups with regard to a variable on an ordinal scale. The most suited test in this case is the Chi-Squared test. The Chi-Squared value is calculated by the following formula:

$$\chi^2_{df} = \sum \frac{(f_o - f_e)^2}{f_e} \tag{8}$$

Where

f_o = The observed frequency
f_e = The expected frequency
df = The degree of freedom

The p-value resulting from this test is equal to 0.000, which indicates that H0 should be rejected for the 0.05 significance level. The differences between the predicted and observed values for each age group have been displayed in the table below. We can conclude that the age distribution in the sample is significantly different from the age distribution in the population.

The largest overrepresentations are in the age groups 18-24 years and 50-59 years. The age group 18-24 most likely corresponds to the friends, fellow students and siblings of the BSc students who spread the survey, while the age group 50-59 is overrepresented due to the presence of many parents and other adult family members of the students. At the same time, the age groups 30-39 years and 40-49 years are underrepresented. This might be problematic, since this age group contains the potential customers of smart thermostats.

		Sample(%)	Population(%)	Difference	Chi2	df	p-value
Age	18-24 years	33,27	11,06	22,21	449,053	7	0,000
	25-29 years	8,58	7,74	0,84			
	30-39 years	6,48	15,42	-8,94			
	40-49 years	7,71	19,37	-11,66			
	50-59 years	31,87	17,76	14,11			
	60-69 years	8,93	15,07	-6,14			
	70-79 years	2,1	8,92	-6,82			
	80 years and older	1,05	4,66	-3,61			

Table C.1: Results Chi-Squared test age distribution

The second demographic which has been collected in the survey is the gender of the respondents. The following test will assess whether the gender distribution is different in the sample. The hypotheses for this test are as follows:

H0: “The gender distribution in the sample is equal to the gender distribution in the population”

H1: “The gender distribution in the sample is not equal to the gender distribution in the population”

Similarly to the test involving the distribution of age categories, this analysis makes use of the Chi-Squared test. The results of this test have been displayed below. The p-value resulting from this test is to 0.133, which indicates that H0 should be accepted given a 0.05 significance level. Thus, there exists no significant difference in gender distribution between the sample and the population for the 0.05 significance level. However, males were (slightly) overrepresented in the sample, which might have been caused by the overrepresentation of this gender in the group of BSc students which spread the survey.

		Sample(%)	**Population(%)**	**Difference**	**Chi2**	**df**	**p-value**
Gender	Man	49,85	49,84	-3,5	2,994	1	0,084
	Vrouw	50,15	50,15	3,5			

Table C.2: Results Chi-Squared test gender distribution

Thirdly, we should assess whether the education level in the sample is representative for the population. The hypotheses for this test are as follows:H0: “The education level distribution in the sample is equal to the education level distribution in the population”

H1: “The education level distribution in the sample is not equal to the education level distribution in the population”

Since this test involves the comparison of an ordinal variable among two groups, Chi-Squared test is used. The results of this test have been displayed in the table below. The p-value resulting from this test is equal to 0.000, which indicates that H0 should be accepted given a 0.05 significance level. The results show that there exists a significant difference between the education level distribution in the sample and the education level distribution in the population. More specifically, the higher education levels are strongly overrepresented, while the lower education level are underrepresented in the sample. This can be explained by the relation between the education level of the students who have spread the survey and the responses they collected. The students belong to the group with the highest education level, which makes it more likely that people in their network also belong to this group.

		Sample (%)	**Population (%)**	**Difference**	**Chi2**	**df**	**p-value**
Education level	Elementary	0,34	10,3	-9,96	601,820	4	0,000
	Vocational	1,03	8,9	-7,87			
	MBO	9,81	41,2	-31,39			
	HAVO/VWO	18,24	9,5	8,74			
	WO	70,56	30	40,56			

Table C.3: Results Chi-Squared test education level

Finally, the representativity of the sample with regard to the working situation is examined. The hypotheses for this test are as follows:

H0: “The working situation distribution in the sample is equal to the working situation distribution in the population”

H1: ““The working situation distribution in the sample is not equal to the working situation distribution in the population”

As with the education level, a Chi-squared test is used to conduct the analysis. The results of the test are presented in the table below. The resulting p-value is equal to 0,000. Thus, H1 can be accepted, which implies that the distribution of the working situation in the sample is not equal to the working situation in the population. Students and paid employees are strongly overrepresented in the sample, while unemployed and retired people are underrepresented. As with the other demographics, the overrepresentations are in line with expectations. The student category likely corresponds to the fellow students, roommates of friends of the BSc students who spread the survey, while the paid job category likely consists of the parents or other adult family members of the students.

		Sample (%)	**Population (%)**	**Difference**	**Chi2**	**df**	**p-value**
Working situation	Student	33,3	18,6	14,7	600,152	5	0,000
	Paid job	54,6	33,6	21			
	Unemployed	3,7	11,4	-7,7			
	Retired	4,2	8	-3,8			
	Other	3,4	6,3	-2,9			

Table C.4: Results Chi-Squared test working situations

Appendix D: Parameters cross alternative MNL models

	Model 1.1		Model 1.2		Model 1.3		Model 1.4	
Attributes	**Parameter**	**p**	**Parameter**	**p**	**Parameter**	**p**	**Parameter**	**p**
Price	-0.824	0.000	-0.833	0.000	-0.832	0.000	-0.895	0.000
Functionality	0.460	0.000	0.476	0.000	0.490	0.000	0.512	0.000
Security	0.980	0.000	1.040	0.000	1.050	0.000	1.080	0.000
Interactions								
Framing * Security	-	-	0.029	0.248	-0.006	0.843	0.041	0.302
Technology Acceptance * Security	-	-	-0.101	0.000	-0.047	0.121	-0.038	0.358
Privacy/Security Consciousness * Security	-	-	0.190	0.000	0.195	0.000	0.168	0.000

Conservativeness * Security	-	-	-0.089	0.000	-0.148	0.000	-0.131	0.000
Framing * Functionality	-	-	-	-	-0.057	0.031	-0.018	0.589
Technology Acceptance * Functionality	-	-	-	-	0.094	0.000	0.104	0.003
Privacy/Functionality Consciousness * Functionality	-	-	-	-	0.004	0.874	-0.022	0.525
Conservativeness * Functionality	-	-	-	-	-0.098	0.000	-0.082	0.001
Framing * Price	-	-	-	-	-	-	-0.096	0.0896
Technology Acceptance * Price	-	-	-	-	-	-	-0.0105	0.86
Privacy/Security Consciousness * Price	-	-	-	-	-	-	0.0505	0.383
Conservativeness * Price	-	-	-	-	-	-	-0.04	0.341

Table D.1: Parameters cross alternative MNL models

Appendix E: Parameters single alternative MNL models

	Model 1.1		Model 1.2		Model 1.3		Model 1.4	
Attributes	**Parameter**	**p**	**Parameter**	**p**	**Parameter**	**p**	**Parameter**	**p**
Price(€/100)	-0.632	0.000	-0.664	0.000	-0.658	0.000	-0.656	0.000
Functionality	0.115	0.000	0.110	0.000	0.108	0.000	0.108	0.000
Security	0.974	0.000	1.036	0.000	1.041	0.000	1.041	0.000
Constant	0.725	0.000	0.787	0.000	0.774	0.000	0.771	0.000
Interactions								
Framing * Security	-	-	0.106	0.000	0.111	0.000	0.041	0.000
Technology Acceptance * Security	-	-	-0.061	0.049	-0.066	0.121	-0.054	0.092
Privacy/Security Consciousness * Security	-	-	0.159	0.000	0.166	0.000	0.162	0.000
Conservativeness * Security	-	-	-0.113	0.000	-0.105	0.000	-0.098	0.001

Framing * Functionality	-	-	-	-	0.007	0.599	0.025	0.264
Technology Acceptance * Functionality	-	-	-	-	0.052	0.001	0.095	0.000
Privacy/Functionality Consciousness * Functionality	-	-	-	-	-0.110	0.874	-0.126	0.525
Conservativeness * Functionality	-	-	-	-	-0.068	0.000	-0.037	0.152
Framing * Price	-	-	-	-	-	-	-0.025	0.315
Technology Acceptance * Price	-	-	-	-	-	-	-0.059	0.045
Privacy/Security Consciousness * Price	-	-	-	-	-	-	0.022	0.429
Conservativeness * Price	-	-	-	-	-	-	-0.042	0.132

Table E.1: Parameters single alternative MNL models

Appendix F: Academic article

Investigating the effect of security and privacy on IoT device purchase behaviour

Nick Ho-Sam-Sooi
MSc Complex Systems Engineering and Management
Delft University of Technology: Faculty of Technology, Policy and Management
Email: N.D.Ho-Sam-Sooi@student.tudelft.nl
Phone: +31621430971

Abstract – Given the significant privacy and security risks of IoT devices, it seems desirable to nudge consumers towards buying more secure devices and taking privacy into account when purchasing these devices. In order to support this goal, this study has examined the effect of security and privacy on IoT device purchase behaviour and assessed whether these effects are sensitive to framing with a mixed methods approach. The first part of the study focuses on quantifying the effect of security and privacy compared to the effect of other device attributes such as the price or functionality by testing a causal model with choice models that have been developed from stated choice data. The second part aims to reveal the underlying mechanisms that determine the effect of privacy and security on purchase behaviour by means of a qualitative survey. The results suggest that security and privacy can strongly affect purchase behaviour, under the circumstances that privacy and security related information is available and communicated in an understandable manner that allows consumers to compare devices. Moreover, the results show that a description of security that focuses on gains is more effective in nudging consumers towards buying more secure devices. Future efforts could build upon this study by comparing the effect of security and privacy to more device attributes, such as ease of use or cost reduction or providing a technical, institutional or process design for a collaboratory effort to nudge users towards buying more secure devices and taking privacy into account when purchasing devices.

1 Introduction

In the IoT, physical objects are connected to a network via internet connectivity to deliver a service to a user [1-2]. The market penetration and societal acceptance of IoT devices is ever-increasing, as more and more use cases for the devices arise and the affordability of the devices improves. This trend is supported by the development of 5G network technology. 5G connectivity allows for lower latency connections of IoT devices and enables larger volume traffic, thus vastly improving the quality of services provided by IoT devices.

IoT devices can provide significant value to consumers by enabling new functionalities that improve their quality of life. For example, smart thermostats allow consumers enable consumers to remotely configure the heating in their home or even remove the need for manual adjustment of their heating system completely.

Although the adoption of IoT devices has significant benefits for consumers, it also introduces some notable risks with regard to security and privacy. In many cases, IoT devices are lacking with regard to basic security controls such as encryption or authentication schemes. Moreover, manufacturers collect large amounts of highly sensitive personal information, such as energy use data. When such data is shared with external third parties, an intentional or malicious infringement of the device owner's privacy might occur.

Consumers can play a large role in mitigating these risks, for example by purchasing secure devices and taking privacy into account when purchasing a device. However, consumers often do not have the required technical knowledge to assess the security level of a device. Moreover, communication of privacy information is often lengthy and overly complex [3].

Therefore, it seems desirable to nudge users towards buying more secure devices and taking their privacy into account when purchasing the devices. Governmental bodies could play an active role in reaching this goal, for example by designing legislation or standards that describe which security and privacy related information should be communicated towards consumers and how such information should be communicated.

However, undertaking such initiatives requires detailed and deep insights into the decision-making process of consumers when purchasing

IoT devices. More specifically, it is crucial to know how, and to what extent, privacy and security influence the choice of consumers to buy IoT devices. Moreover, the sensitivity of these effects with regard to personal factors should be investigated to evaluate whether the effect of privacy and security differs between various subgroups of consumers. Finally, the manner in which security and privacy are framed might determine their effect on purchase behaviour. To illustrate this, consumers might take security and privacy into account more strongly when receiving a description of the level of privacy and security that focuses on the gains that can be achieved from buying secure devices or taking privacy into account when purchasing devices. Thus, some frames might be more effective in nudging consumers towards buying more secure devices and taking privacy into account when purchasing devices. For this reason, the sensitivity of the effects of privacy and security to framing should be examined. This study aims to provide these insights by answering the following research question:

"How do security and privacy influence the choice of consumers to buy an IoT device? And how sensitive is the effect of security and privacy to framing and personal factors?"

The study takes a mixed methods approach towards answering this research question. The quantitative part of the study focuses on quantifying the effect of privacy and security on consumer choice behaviour by developing and testing a causal model that describes the effects of various explanatory factors on choice behaviour. This goal is reached by constructing choice models from data that is gathered from a stated choice experiment. The qualitative part of the study targets the underlying rationales that determine how privacy and security affects consumer choice behaviour by asking consumers open questions regarding the role of privacy and security in their decision to buy or not to buy an IoT device.

The remainder of this paper is structured as follows. Firstly, the background section provides a brief overview of the existing body of literature regarding the research topic. Section three describes the methods that have been used to conduct the analysis. In section 4, the results of the analysis are presented. Section 5 consists of the conclusions that answer the main research question of the study. Furthermore, the results of the study are discussed in terms of their implication and limitations and possibilities for further research are introduced.

2. Conceptual model

2.1 Technology Acceptance Modelling

Currently, the effect of security and privacy on the purchase behaviour of consumers is relatively understated. However, studies in the Technology Acceptance Modelling (TAM) field have investigated how the consumer perception of security and privacy with regard to innovative technologies influences their acceptance. The basis of this field, commonly known as Technology Acceptance Modelling (TAM) has been formed by Davis [2], who concluded that there exist clear relationships among ease of use, price, usefulness and acceptance of innovative technologies. Davis defined acceptance as the usage of a technology or system by its end users.

In the following years, IT researchers have extended this model by adding perceived security, risk and trust-related factors and applying it to digital products. For example, Gu, Lee & Suh [5] applied the Technology Acceptance Model (TAM) to mobile banking. From this study, the authors concluded that trust, ease of use and the acceptance of mobile banking are closely interrelated. Furthermore, a study by Salisbury, Pearson, Pearson & Miller [6] evaluated which factors affect the willingness to engage in web-based shopping. The results of this study showed that Web security perception plays a large role in determining purchase intent. Even more, it has a stronger effect than ease of use and usefulness of technology. The authors defined Web security perception as "the extent to which one believes that the Web is secure for transmitting sensitive information" [6, p.3]. Their measurement of this concept did not take into account any framing effects. On the contrary, positive and negative frames were used additively to determine the security perception of respondents. In line with this thinking, a study by Crespo, del Bosque & de Los Salmones Sanches (2009) has led to the conclusion that various risk factors such as security, strongly limit the acceptance of e-commerce. The researchers framed the risk factors as potential losses, thus negating the possible effect of framing in the communication of these risks. From the results of these studies, the following hypotheses can be construed:

- H1: The price of an IoT device negatively influences the probability that the device is purchased.

- H2: The number of functionalities of an IoT device positively influences the probability that the device is purchased.

- H3: The security level of an IoT device positively influences the probability that the device is purchased.

2.2 Framing

The previous section concluded that it is still unclear to what extent the effect of security on the choices of consumers is sensitive to framing. Entman [8, p.2] defined framing as "the selection of some aspects of a perceived reality and making them more salient in a communicating text, in such a way as to promote a particular problem definition, causal interpretation, moral evaluation, and/or treatment recommendation for the item described". Moreover, according to Entman, frames describe problems, diagnose causes, make moral judgements and select the most suited remedies. Chong & Druckman [9] provide a more high-level definition of framing, defining the concept as "the process by which people develop a particular conceptualisation of an issue or reorient their thinking about an issue"

Gain/loss framing is one of the most prevalent frames in message framing literature. In the gain frame, the message focuses on the gains the decision-maker can acquire when opting for a certain alternative. On the contrary, the loss frame communicates the possible losses of an outcome. According to Prospect Theory, people tend to be risk-averse when being presented with sure gains and risk-seeking when facing sure losses [10]. This goes against classical utility theory, in which similar outcomes provide the same amount of value to the decision-maker. Kahneman & Tversky developed a different choice model, in which value is attained from gains and losses rather than net outcomes and the probabilities in the utility function are replaced by decision weights.

Researchers in the medical field have applied the concept of gain/loss framing in order to assess effect of gain/loss framing on the choice of patients to opt for a certain treatment. In these studies, gain/loss framing was applied to the communication of treatment information to patients who face the decision to opt for a certain treatment. Armstrong, Schwartz, Fitzgerald, Putt, & Ubel (2002) presented a group of 451 individuals with treatment information. The individuals were randomly divided into three groups. The first group only received the survival rates of the treatment, while the second group received the mortality rates and the third group received both the mortality rates and the survival rates. Upon receiving the information, the individuals were asked to make the decision whether to opt for preventative surgery. The results suggested that individuals who received the mortality rates were less likely to prefer the surgery. These results are clearly in line with the hypotheses of Prospect Theory, as individuals who are presented with the loss frame are risk-seeking and vice versa.

Many studies following a similar procedure have been published during the years. A study by Detweiler, Bedell, Salovey, Pronin, & Rothman (1999) concluded that beachgoers who received a message which focused on the gains of using sunscreen were more likely to buy and use sunscreen. Similarly, Schneider et al. (2001) concluded that a message describing the benefits of stopping had a stronger effect on the willingness of the smokers to stop smoking than a message which contained the negative effects of smoking. Kühberger (1998) conducted a meta-analysis of the early contributions in message framing literature. From a sample set of 136 empirical analyses, Kühberger calculated a set of 230 effect sizes. The results were in line with the original hypothesis of Tversky and Kahneman, as messages in the gain frame generally led to risk-averse behaviour and messages in the loss frame caused more risk seeking behaviour.

Studies in the message framing literature have concluded that messages which focus on gains are more effective in nudging consumers to take preventative measures to mitigate risks. In this line of thinking, buying a secure product or taking privacy into account can also be seen as a preventative measure to mitigate the risk of cyber threats or privacy infringements. Therefore, it can be expected that messages focusing on the gains of buying more secure devices and taking privacy into account are more effective. This leads to the following hypothesis.

- H4: Messages that focus on the gains of security and privacy are more effective in nudging users to purchase more secure

devices and consider privacy when buying IoT devices

From the literature study, a set of 5 hypotheses have been developed regarding the effect of privacy and security on the purchase behaviour of consumers. These hypotheses are visualised in the causal model that has been displayed below.

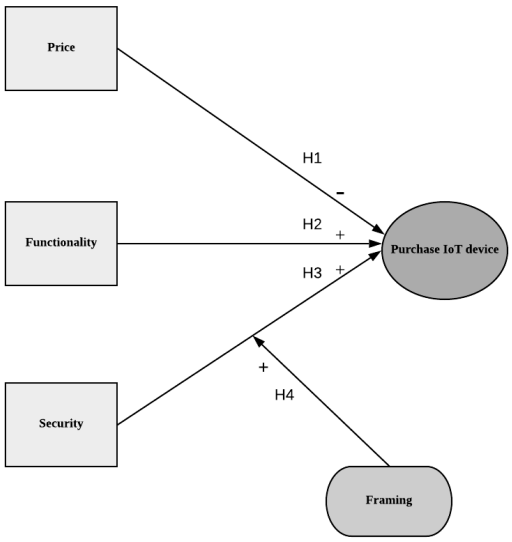


Figure 1: Causal model

3. Method

In this section, the methods of both the quantitative and qualitative study are presented.

3.1 Quantitative study: Stated choice experiment

The data for the quantitative study has been collected by means of a stated choice experiment. Stated choice experiments are especially suited to analyse the effect of device attributes, personal factors and framing on choice behaviour. In this experiment, the respondents were presented with various choice sets consisting of two smart thermostats. The alternatives in the choice set varied with regard to three attributes: Price, Functionality and Security. Privacy was not included as an attribute in order to limit the needed number of choice sets per respondent. In order to resemble real-world pricing, the price attribute varied on four levels: €100, €150, €200, and €250. The functionality attribute was coded additively, which implies that the number of functionalities increases as the value of the functionality attribute increases. The following functionalities were included as part of the attribute levels:

1. Remote control (F1): The user is able to remotely access the device in order to adjust the temperature, scheduling or make use of other functionalities.

2. Geofencing (F2): The geofencing capability of the user's smartphone is used to assess whether the users has left his/her house and adjust temperatures accordingly.

3. Sensing (F3): The home is equipped with sensors, which assess whether the occupants are awake, sleeping or outside of the house. The temperature is adjusted according to the data collected by the sensors.

4. Learning (F4): The user inputs basic schedule parameters. The device makes use of algorithms in order to learn the schedule of the occupants and collects data from sensing to detect changes in the schedule and respond to them.

The security level varied on two levels. Moreover, the respondents in the stated choice experiment were randomly divided into two groups. The descriptions of the security attribute for both levels are displayed in the table below. For the first group, the security level of the alternatives was framed in terms of gains, while the description of the security level focused on losses for the second group.

Frame	**Security description**
Gain	"This device is/is not secured properly"
Loss	"This device can/cannot be hacked"

Table 1: Security description

With these attribute levels, an orthogonal fractional factorial design was constructed by making use of basic plan 3. Each row of the design contains a profile. The choice sets were constructed by means of sequential construction.

Per choice set, the respondents were asked whether they would purchase each individual

smart thermostat in the choice set, given that their thermostat had broken and they were faced with the decision to buy a new smart thermostat. Moreover, they were asked which of the two smart thermostats had their preference in the case they had to choose between the two alternatives.

In addition, the respondents were asked questions regarding a set of demographical variables, in order to test the representativity of the collected sample. The following demographics were included in the survey: Age, gender, education level and working situation.

Finally, the survey measured a set of indicators that were expected to play a role in the choice behaviour of consumers purchasing IoT devices. These indicators function as input for a factor analysis, which aims to define a set of personal factors from the indicators. The factors have been constructed by means of Principal Axis Factoring (PAF). This method is especially suited to measure the values of non-measurable constructs such as views, opinions and beliefs. The axes have been rotated by means of oblique rotation, which allows for correlation between factors and simplifies the interpretation of factors.

In order to measure the values on the indicators, the respondents were asked to evaluate whether they agreed with a set of statements, the statements have been displayed in the table below.

Nr.	**Statement**
I1	"I keep up with technological developments"
I2:	"I read the technology section when reading newspapers or visiting news websites"
I3:	"I find it interesting to follow the development of new IT products"
I4:	"Innovation is important for economic development"
I5:	"Investments in innovative technologies are important for society"
I6	"If a new IT product has been developed, I want to buy the first version"
I7:	"I pay attention to the security risks of my IT devices"
I8	"When purchasing an IT device, I consider the security risks of the device"
I9:	"The security of my IT devices is important to me"
I10:	"My personal information should be protected sufficiently"
I11:	"I keep track of which information is collected when using online services"
I12:	"I am concerned with the security risks of my IT devices"
I13:	"When using IT devices, I am concerned with the use of my personal data by external parties"
I14:	"When using online services, I am concerned with the use of my personal data by external parties"
I15:	"I undertook action to improve the security of my IT devices"

Table 2: Indicators

The survey was spread by a group of BSc students from the faculty of Technology, Policy and Management of Delft University of Technology as part of a data analytics course. The students were asked to share the survey within their social network and collect 5 responses to the survey per person.

3.2 Quantitative study: Discrete choice modelling

From the collected data, Random Utility Maximisation (RUM) based discrete choice models have been developed. These models describe the probability that a certain-decision maker chooses an alternative from a given set of alternatives which vary on a set of criteria or attributes.

More specifically, Multinomial Logit (MNL) models are used to assess the effects of the attributes, personal factors and framing on choice behaviour. MNL models assume that the error terms in the utility function are independently and identically distributed across all alternatives, which implies that they have the same probability

distribution and are mutually independent. The utility of an alternative is calculated by the sum of the product of the criteria scores and a set of linear parameters. Thus, the utility is calculated by the following formula:

$$U(a_i) = \sum_{j=1}^{m} w_j * E(a_i, c_j) + \varepsilon \qquad (1)$$

Where w_x is the parameter or weight of attribute x, E (a_x,c_y) resembles the expected effect of alternative x on attribute y and ε is equal to the error term.

For MNL models, the probability that an alternative is chosen from a set of alternatives is calculated as follows:

$$P(X = a_i) = \frac{e^{U(a_i)}}{\sum_{j=0}^{n} e^{U(a_j)}} \qquad (2)$$

Where P(X = a_x) entails the probability that alternative X is chosen from a predefined choice set, U(a_x) is the utility of alternative x and n is equal to the number of alternatives in the choice set.

For the model selection process, various model statistics are calculated that measure the quality of the developed models. Firstly, the Likelihood Ratio Test (LRT) is used to compare the quality of two models. The statistic that relates to this test is calculated as follows:

$$LRS = -2 * (LL_A - LL_B) \qquad (3)$$

Where LL_x is the Log-Likelihood of model x.

Secondly, the R-squared value is calculated for each model by dividing the variance of the dependent variable that the model is able to explain by the total variance of the dependent variable.

The models have been divided in two groups. The models in the first group have been developed with the single alternative choices as the dependent variable, while the cross alternative choices functioned as the dependent variable for the models in the second group.

Finally, an iterative modelling process is applied, which implies that more explanatory variables are added to the model in each iteration to assess whether adding more variable to the model significantly improves the goodness of fit. The table below provides a description of the models that are developed in each iteration.

Model Nr.	Description
1.1	MNL: Device attributes
1.2	MNL: Device attributes + interaction factors and framing with security attribute
1.3	MNL: Device attributes + interaction factors and framing with security and functionality attribute
1.4	MNL: Device attributes + interaction factors and framing with security. functionality and price attribute

Table 3: Modelling process

3.3 Qualitative study

The qualitative study took a different approach by conducting a survey in which the respondents were asked open questions regarding their decision to purchase or not to purchase a smart thermostat. Firstly, the respondents were asked which factors had influenced their decision to buy or not to buy a smart thermostat. Subsequently, the respondents were triggered to contemplate the role of security and privacy in their decision to buy or not to buy a smart thermostat. Furthermore, the respondents were asked to rate the severity of a security or privacy risk that was described by means of a hypothetical scenario and provide a motivation for their rating on a five-point scale. An overview of the scenarios is presented in the table below. Finally, the respondents were requested to indicate which scenario described the most severe risk in their opinion.

Scenario Nr.	Description
1	The smart thermostat collects data about your energy use and keeps track of your location. A criminal gains access to this information to determine the right moment for a burglary
2	The smart thermostat collects data about your energy use and keeps track of your location. The producer of your thermostat collects this data and may be obligated to share it with external parties, such as insurers or tax authorities.

3	The smart thermostat collects data about your energy use and keeps track of your location. The producer of your thermostat collects this data and shares it with marketing bureaus, which use it to develop personalised advertisements.
4	A criminal gains access to your smart thermostat, allowing him/her to control the heating in your house.
5	A criminal gains access to your home network via your smart thermostat, allowing the criminal to gain access to personal information on the network, such as passwords or browsing data.
6	Your smart thermostat is part of a large network of devices which is being used to execute cyber-attacks on large organisations.

Table 4: Scenarios

4. Results quantitative study

4.1 Sample

For the quantitative study, the students collected a dataset containing 709 respondents. A subset of 93 respondents who did not provide an answer to the questions related to the choice experiment were removed from the dataset. Moreover, 35 responses were collected from the same IP address within a distinctly small time frame. These responses were removed from the data set as it is unlikely for such a large amount of valid responses to be collected within a small time frame from the same IP address. It is likely that these responses consist of students who filled the survey in themselves multiple times.

4.2 Representativity

In order to test the representativity of the collected sample, the values of the demographical variables in the sample are compared to the values of these demographical variables for the target population of the study. For this purpose, various Chi-Squared tests have been executed. The results show that the age groups 18-24 years and 50-59 years are overrepresented. Secondly, the sample mostly consists of respondents who have a high education level. Finally, the working situation categories “student” and “paid job” are strongly overrepresented in the sample. These overrepresentations can be explained by the data collection process. The BSc students who spread the survey most likely shared the survey with fellow students, housemates, siblings, parents and other mature family members. It seems sensible that these biases have caused the overrepresentations in the sample.

The overrepresentations in the sample might cause under- or overestimation of relations between factors, attributes, demographics and choice behaviour. However, the main aim of this research is to illustrate that certain relations exist. The overrepresentations do not limit the ability of the developed models to reach this goal.

4.3 Factor analysis

From the values of the indicators, personal factors are deduced by means of PAF. The resulting factors and resulting load have been displayed in the table below.

Nr.	Factor 1	Factor 2	Factor 3
I1	-	-	-
I2		.785	
I3		.733	
I4			-.888
I5			-.830
I6		.536	
I7	-	-	-
I8	.556		
I9	-	-	-
I10	-	-	-
I11	.534		
I12	.755		
I13	.897		
I14	.833		
I15	.407		

Table 5: Factor loads

The first factor is defined by indicators that relate to the attitude of the respondents towards privacy/security issues of IT devices. Thus, this first factor can be labelled as “privacy/security awareness”. The second factor relates to the respondent’s interest in the development of

technology as well as their adoption of new technology. Therefore, the second factor can be labelled as "Technology Acceptance". Finally, the third factor is determined by the two indicators that measure the perceived importance of innovation. The two indicators load negatively on the factor, which implies that the indicators measure the pole opposite of this construct. Consequently, this factor can be labelled as "Conservativeness". The indicators that have been removed from the factor analysis are excluded from the analysis completely, since they do not possess a significantly different meaning than the factors.

4.4 Model Selection

During the modelling process, various models have been developed and assessed by means of the model statistics that have been discussed in section 3. The models in model group 1 provide a significantly better fit to the data. To illustrate this, the R-squared value of model 1.4 in model group 1 is equal to 0,307, while the value for this statistic of model 1.4 in model group 2 is 0,179. For this reason, the remainder of this section will focus on the results of model group 1. The models and their respective R-Square value and LRT have been displayed in the table below.

Nr.	Log likelihood	R^2	LRT (critical value)
1.1	5054,914	0,265	-
1.2	4580,136	0,297	949,556 (9,488)
1.3	4544,435	0,306	71,402 (9,488)
1.4	4541,340	0,307	6,19 (9,488)

Table 6: Model selection

According to the LRT values, model 1.3 provides the best fit to the data. However, the LRT value of model 1.4 is relatively close to the critical value and the model contains a notable interaction effect of the price attribute with the technology acceptance factor. For this reason, model 1.4 is used to draw conclusions in the remainder of this paper.

4.5 *Model parameters*

The parameters of the resulting model, model 1.4 from model group 1, are displayed below.

Attributes	Parameter	p
Price	0,656	0,000
Functionality	0,108	0,000
Security	1,041	0,000
Constant	0,771	0,000
Framing interactions		
Framing * Security	0,041	0,000
Framing * Functionality	0,025	0,264
Framing * Price	-0,025	0,315
Factor interactions		
Technology Acceptance * Security	-0,054	0,092
Privacy/Security Awareness * Security	0,162	0,000
Conservativeness * Security	-0,098	0,001
Technology Acceptance * Functionality	0,095	0,000
Privacy/Security Awareness * Functionality	-0,126	0,525
Conservativeness * Functionality	-0,037	0,152
Technology Acceptance * Price	-0,059	0,045
Privacy/Security Awareness * Price	0,022	0,429
Conservativeness * Price	-0,042	0,132

Table 7: Model parameters

Firstly, the model contains the direct effects of the device attributes on the utility of the alternatives. Thus, three respective parameters have been calculated for each of these attributes; Functionality, Price and Security. The model also contains a constant that describes the expected value or utility of an alternative when each of the attributes is set to 0. Each of these effects is statistically and practically significant. In line with the hypotheses, the price attribute has a negative effect on the expected utility of an alternative. The security level and functionality of an alternative have a positive effect on its e utility.

The technology acceptance factor has significant interactions with the three device attributes. Respondents with a high score on this factor are willing to make concessions on security and price in order to buy the newest technology that provides them with innovative functionalities. Similarly, the privacy/security awareness factor positively moderates the effect of security on the purchase behavior, which implies that respondents who are more aware of security and privacy risks take security more strongly into account when purchasing a device. Finally, the conservativeness factor negatively interacts with the security attributes. This result suggests that security contributes less to the value of a device for respondents who do not value innovation.

With regard to framing, the results show that security has a stronger effect on the purchase decision for respondents who were faced with the gains of buying a secure device. This finding is in line with the hypothesis of Kahneman & Tversky, who postulated that people are more risk averse when faced with possible gains.

5. Results Qualitative study

5.1 Response

A total of 27 responses were provided to the survey for the qualitative study. In the collected sample, the higher education levels are highly overrepresented. This overrepresentation is deemed to be unproblematic due to the qualitative nature of the study.

5.2 Purchase decision

Firstly, the respondents were asked to evaluate what factors played a role in their decision to buy or not to buy a smart thermostat. Strikingly, security or privacy were only mentioned twice as a motivation for the purchase decision. For device owners, the reasons to purchase a smart thermostat were mainly focused around the functionalities the device provides, ease of use and energy cost reductions. With regard to the decision to buy a specific smart thermostat, the compatibility with other devices such as the boiler, voice assistants and smart home devices was mentioned frequently.

After being triggered to actively contemplate the role of security and privacy in their purchase decision, many respondents are able to address some high-level privacy and security related concerns regarding smart thermostats.

The results show that the respondents only start thinking about security and privacy concerns when being actively triggered to evaluate such topics. Without being prompted to think about privacy and security, the respondents focused mainly on other device attributes such as functionality and ease of use.

5.3 Risk awareness

15 out of the 27 respondents indicated that they were able to mention security and privacy risks of smart thermostats. The respondents mostly gave high level descriptions of security and privacy risks, using common terms such as “hacking” or “data going public”. It seems notable that the risk descriptions of the respondents strongly lack any detail and are not related to realistic threat scenarios.

5.4 Scenario’s

The assessment of scenarios allows for the generation of insights regarding the risk assessment of the respondents. The main goal of the analysis is to determine the underlying factors that influence this process rather than quantifying the effects of these factors. For this reason, the focus lies on analysing the motivations that the respondents have provided for their rating rather than quantitatively assessing the ratings per scenario.

Firstly, the perception of the level of security or privacy related to the device are often mentioned as a motivation to rate a scenario. Some respondents rate the severity of a risks scenario as “low” because they expect that sufficient controls have been put in place. For example, respondents rated the severity of risks in privacy related scenarios have been rated as “low” because

GDPR has been put in place and this regulation ought to be sufficient protection against privacy infringements. On the contrary, other respondents mentioned that they perceived the level of security and privacy with regard to IoT devices in general to be low.

Secondly, the probability of occurrence seems to play a role in the risk evaluation process of the respondents. Many respondents have rated the severity of a scenario to be low, as they thought that such a risk would be very unlikely to occur in real life. On the other side, probability of occurrence was also mentioned frequently as a motivation to rate the severity of a risk as "high".

Thirdly, the benefits for the third party are reported as a motivation for the assessment of a risk. If the respondent is of the opinion that the threat actor in the risk scenario is not able to achieve an attractive benefit, the respondent is likely to rate the severity of the risk as "low"

Finally, the respondents often mention the impact of a risk scenario as a crucial factor. To illustrate this, scenario 5 posed the most severe risk for many respondents, as this scenario has a further reaching impact than the other scenarios. In this scenario, the scope of the impact exceeds the information that is collected, stored and used with regard to the use of the smart thermostat.

6. Conclusions

This study has investigated the effect of security and privacy on the IoT device purchase decision of consumers by answering the following research question:

"How do security and privacy influence the choice of consumers to buy an IoT device? And how sensitive is the effect of security and privacy to framing and personal factors?"

The quantitative part of the study revealed that security has a notably strong effect on the purchase decision of respondents in a stated choice experiment. On the contrary, security and privacy were only mentioned once or twice as a motivation to buy or not to buy a smart thermostat by the respondents in the survey for the qualitative study. The main difference between both studies is that the respondents in the quantitative study are triggered to think about security, while this is not the case in the qualitative study. Moreover, the respondents in the quantitative study are presented with an easily understandable description of security, which allows them to easily compare alternatives with regard to the security level. It is likely that this is not is not the case in real world situations.

The second part of the research question targets the sensitivity of the effect of security and privacy to framing and personal factors. With regard to framing, the results show that security has a stronger effect for respondents who received a gain focused description of security. This finding is in line with the hypothesis of Prospect Theory, which postulates that people are more risk averse when faced with possible gains.

Furthermore, the results have illustrated that consumers who are more aware of the privacy/security risks of (IoT) devices, take security more strongly into account when purchasing IoT devices. The qualitative study also investigated the risk awareness of consumers. The results indicated that some consumers are able to list some of the security and privacy risks of smart thermostats. However, the descriptions of these risks strongly lack detail and are not specific for smart thermostats.

Next to this, the qualitative study examined the risk assessment process of consumers. From this analysis, a set of factors have been derived that were frequently mentioned as a motivation to assess the severity of a privacy or security related risk of smart thermostats. The following factors were found to be relevant: Perceived security and privacy level, probability of occurrence, third party benefits, and impact.

Finally, the quantitative study found a negative interaction effect of the technology acceptance factor with the price and security attributes and a positive interaction effect with the functionality attribute. This suggests that people who score high on this factor can be seen as the "first adopters" of innovative technologies and are more willing to buy less secure and more expensive products that do provide them with new functionalities and improve their quality of life.

To conclude, the study has found that security and privacy can have a strong effect on the purchase decision of consumers, under the specific circumstances that privacy and security related information is easily available and communicated in an understandable manner that allows for comparison of alternative devices in a simple and timely manner. The effect of security is moderated by the privacy/security awareness, technology acceptance and conservativeness of consumers. Finally, the results show that security related information that focuses on the gains of

security is more effective in nudging consumers towards buying more secure devices.

7. Discussion

The results of this study have several practical implications. The results have shown that security does affect the purchase behaviour of consumers under the condition that security or privacy related information is available and is communicated in a simple and understandable manner. This result suggests that governmental bodies could nudge users towards buying more secure devices and taking privacy into account by ensuring that by ensuring that security or privacy related information is communicated towards consumers in an understandable manner that allows for timely comparison of devices with regard to security and privacy. Governmental bodies could work towards this goal by defining standards or legislation that describe what security and privacy related information should be provided to consumers and how this information should be communicated. Due to the immense complexity of the IoT security and privacy topic, it is advised to include market parties, such as manufacturers and retailers, in the development process of such legislation or standards. Existing frameworks that conceptualise security and privacy of IoT devices can leveraged as a starting point for such collaborations.

Furthermore, the results of the indicate that consumers who are more aware of privacy and security risks are more likely to consider security and privacy when purchasing IoT devices. Thus, improving the risk awareness of consumers supports the goal of nudging users towards buying more secure devices and taking their privacy into account when purchasing devices. In order to reach this goal, governmental bodies could initiate awareness programs that specifically focus on communicating security and privacy risks of IoT devices to consumers. The results of the qualitative study have identified four potential factors that could form the basis of such efforts: Perceived security and privacy, probability of occurrence, third party benefits and impact. Finally, the results suggested that the first adopters of innovative technologies can be identified as a potential focus groups for awareness campaigns.

The study also has significant scientific implications. Firstly, it shows that stated choice experiments can be used as a method to estimate the framing effects. In current studies, framing effects are often evaluated by presenting research subjects with a single choice task. By means of stated choice experiments, the standard errors of the resulting parameters are lowered, thus improving the validity of the developed models. Additionally, the method allows researchers to compare the effects of various attributes on choice behaviour.

In addition, the study contributes to the TAM field by evaluating the effect of various explanatory factors on the purchase decision of consumers. The study differs from the studies in the TAM field with regard to the dependent variable in its causal model. The dependent variable in TAM studies is the acceptance of technologies, while the choice for a specific device functions as the dependent variable in this study. The measurement of the dependent variable also differs from existing studies. In this study, a stated choice experiment is used to measure the choices rather than observing the outcome of a single choice task.

8. Limitations

The quantitative study has observed stated choices rather than choices in real-world situations. It can be argued that this limits the validity of the developed models, as people might exhibit significantly different choice behaviour in the setting of a stated choice experiment. For example, the effect of security might be lower in the case of real-world purchases due to the limited availability of security related information.

Moreover, the alternatives in the stated choice experiment varied on a small set of three attributes. It can be expected that other device attributes, such as ease of use or compatibility with other devices, also have a strong effect on the purchase behaviour of consumers.

Limitations can also arise from the specific coding of the device attributes. In this case, the operationalisation of the security attribute has its drawbacks. The security attribute has been varied on two levels. It is possible that this coding has led to an overestimation of the effect of security on the choice behavior, as it seems sensible that most respondents would not purchase a device that “is not secured properly” or. “can be hacked”

Fourthly, it is questionable whether security and privacy can be framed as a pure gain. To illustrate this, the security attribute was framed as “this device is/is not secured properly”. The term “secured” still suggests that there exists some external threat. This external threat can be seen as

potential losses. However, the term “securing” seems a more positive term than “hacking” from a semantic point of view.

MNL models have been developed to assess the effect of security on choice behaviour in the quantitative study. MNL models assume that the error terms in the utility function are i.i.d. If this assumption is incorrect, this can result in biased parameter estimates.

For the qualitative study, a survey was used to reveal the underlying rationales that determine how security affects the choice behavior of consumers. A survey allows for the generation of responses in a timely and costless manner. However, a using a survey for this goal has its limitations. When using a survey, the researcher is not able to ask follow up questions when needed. However, an interactive survey design was applied that asked the respondents for more in depth answers in order to deal with this limitation.

9. Further research

More research is needed to further address the identified knowledge gaps. Firstly, this study only investigated the effect of a limited set of three device attributes. Privacy was not included as a device attribute in this study. In order to assess whether similar conclusions hold for privacy and compare the effects of security and privacy to other device attributes, future research could build upon this study by including privacy and other device attributes.

Secondly, the security attribute was coded as a binary variable, which might have led to the overestimation of the effect of security on the purchase decision of consumers. Future research could evaluate how other operationalisations of security affect choice behavior in order to determine what operationalisation is most suited to nudge consumers towards buying more secure devices.

Thirdly, this study has observed stated choices rather than real-world choices. Further research could use revealed choice data as an input for the development of choice models to assess whether real-world choice behavior resembles the choice behavior in a stated choice experimentFor example, activity on web shops could be monitored to collect data regarding the purchase behavior of consumers.

References

[1] Singh, K. J., & Kapoor, D. S. (2017). Create Your Own Internet of Things: A survey of IoT platforms. IEEE Consumer Electronics Magazine, 6(2), 57-68.

[2] Sicari, S., Rizzardi, A., Grieco, L. A., & Coen-Porisini, A. (2015). Security, privacy and trust in Internet of Things: The road ahead. *Computer networks*, *76*, 146-164.

[3] Schaub, F., Balebako, R., Durity, A. L., & Cranor, L. F. (2015). A design space for effective privacy notices. In *Eleventh Symposium On Usable Privacy and Security ({SOUPS} 2015)* (pp. 1-17).

[4] Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS quarterly, 319-340.

[5] Gu, J. C., Lee, S. C., & Suh, Y. H. (2009). Determinants of behavioral intention to mobile banking. Expert Systems with Applications, 36(9), 11605-11616.

[6] Salisbury, W. D., Pearson, R. A., Pearson, A. W., & Miller, D. W. (2001). Perceived security and World Wide Web purchase intention. Industrial Management & Data Systems, 101(4), 165-177.

[7] Crespo, A. H., del Bosque, I. R., & de los Salmones Sánchez, M. G. (2009). The influence of perceived risk on Internet shopping behavior: a multidimensional perspective. Journal of Risk Research, 12(2), 259-277.

[8] Entman, R. M. (1993). Framing: Toward clarification of a fractured paradigm. Journal of communication, 43(4), 51-58.

[9] Chong, D., & Druckman, J. N. (2007). Framing theory. Annu. Rev. Polit. Sci., 10, 103-126.

[10] Kahneman, D. & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*(2), 363-391

[11] Armstrong, K., Schwartz, J. S., Fitzgerald, G., Putt, M., & Ubel, P. A. (2002). Effect of framing as gain versus loss on understanding and hypothetical treatment choices: survival and mortality curves. *Medical Decision Making*, *22*(1), 76-83.

[12] Detweiler, J. B., Bedell, B. T., Salovey, P., Pronin, E., & Rothman, A. J. (1999). Message framing and sunscreen use: gain-framed messages motivate beach-goers. *Health Psychology*, *18*(2), 189

[13] Schneider, T. R., Salovey, P., Pallonen, U., Mundorf, N., Smith, N. F., & Steward, W. T. (2001). Visual and Auditory Message Framing Effects on Tobacco Smoking 1. *Journal of Applied Social Psychology*, *31*(4), 667-682.
[14] Kühberger, A. (1998). The influence of framing on risky decisions: A meta-analysis. *Organizational behavior and human decision processes*, *75*(1), 23-55.