Data: To Share or Not To Share?

An analysis of factors impacting the intention of consumers to share transportation data in the Swedish inner-city mobility context

JONAS ADAMS
LENNARD DIETRICH

Master Thesis
Stockholm School of Economics
May 2019

Title:

Data: To Share or Not To Share – An analysis of factors impacting the intention of consumers to share transportation data in the Swedish inner-city mobility context

Abstract:

In this research project, we conducted a mixed methods study to analyze the factors that have an impact on the consumer's intention to share data. Whereas the vast majority of academic research on the consumer focuses on data privacy concerns, there is little research on the intention to share data from the consumer's perspective. Based on the pre-study interviews and literature review we adapted the Unified Theory of Acceptance and Use of Technology (UTAUT) as our theoretical model and extended it by four additional theories. Our empirical data collection focused on non-sensitive transportation data in the Swedish mobility setting and consisted of two steps. Our first step was to collect and analyze 391 questionnaire results to generate insights about the factors for consumers to share their data in the mobility context. We conducted a Structural Equation Modelling analysis to test the hypotheses of our theoretical model. Our second step was to conduct three focus groups with selected participants based on their previous questionnaire answers to understand the reasoning behind these factors. The results were that high perceived benefits one receives in exchange for sharing data, i.e. Performance Expectancy, and low perceived barriers of sharing the data, i.e. Effort Expectancy, have a significant impact on the consumer's intention to share data. Furthermore, we found that transparency is a highly valued factor together with the overall feeling of sharing data that impact individuals' decision making in those situations. Contrary, social relations are not impacting individuals' intention to share their data. The results have implications for the government in terms of Open Data initiatives, city planning and legislation, and for businesses in terms of data collection and design of the transportation service.

Keywords:

Consumer Behavior, Data-Sharing, Mobility, Unified Theory of Acceptance and Use of Technology (UTAUT), Structural Equation Modelling (SEM)

Authors:

Jonas Adams (41178)

Lennard Dietrich (41179)

Supervisor:

Gianluca Chimenti, Department of Marketing and Strategy

Master Thesis

Master Program in Business and Management

Stockholm School of Economics

I. Acknowledgments

First of all, we would like to thank our family and friends, who have supported us during the course of the last years and were always there for us. A special thanks goes to the MBM Class of 2017-19 for the amazing two years and beyond! Thank you for all your support!

Additionally, we would like to thank our thesis supervisor Gianluca for his guidance and mentoring along this research project and the fruitful discussions. Thank you for sharing your insights with us!

Finally, we would like to express our gratitude to the pre-study interview partners, questionnaire respondents and focus group participants for their time and support for our research project. Thank you for sharing your data with us!

II. Glossary

Affect	We define Affect as the personal feeling associated with the act of sharing data.
Confirmatory Factor Analysis (CFA)	Analysis tool as part of the Structural Equation Modeling technique to examine whether the constructs of models in regard to the measured values are consistent with the researcher's theoretical understanding of those constructs or factors. (Jöreskog, 1969)
Construct	Theoretical variable, which is not directly measured but serves as a representative value "build" out of different items and is of an independent nature. (Muthén, 2002)
Dataset	We define a dataset as the information about an individual that can be voluntarily shared consisting of the following information of the individual for a specific trip: time, location and mean of transportation for one specific trip.
Economic Benefit	We define Economic Benefit as any reward in monetary or close to monetary (e.g. voucher, free usage) format that a consumer gets.
Effort Expectancy	"The degree to which a person believes that using a particular system would be free of effort." (Davis, 1989)
Enhanced Convenience	We define Enhanced Convenience as an improvement in the way a consumer is traveling in terms of less stress and time.
Facilitating Conditions	"The degree to which an individual believes that an organizational and technical infrastructure exists to support the use of the system." (Venkatesh, et al., 2003)
Government	We define government in a broader sense comprising all thinkable forms of political institutions from country and municipality level and the associated roles and responsibilities with it.
Guidance	We define Guidance as the support a consumer can get in the process of sharing the data.
Increased Safety	We define Increased Safety as the reduced risk of any physical damage that could occur during traveling.
Inner-city mobility	We define inner-city mobility as the conduction of trips of individuals inside an urban environment with the option to choose from different means of transportation regardless of the purpose of the trip.
Institutional Trust	We define Institutional Trust as the trust in a government or a related non-profit organization.
Intention to Share	We define Intention to Share as the willingness of a consumer to share his or her data with the service provider.
Item	We define an item to be the measurement scale questions used in the questionnaire to derive the constructs and latent variables. We used a Likert-scale measure to grasp the agreeableness towards statements.

Latent variable	Latin, *lateo*, "to lie hidden" – Similar to a construct, it is a theoretical variable, which is not directly measured but serves as a representative, which, in addition, is further inferred from other constructs in a following layer, thus can be seen as dependent. (Muthén, 2002)
Loading	Linear regression coefficients that show the correlation between the construct or latent variable and the item. High values indicate that the items explain the construct well and therefore displays the item reliability. Normally, there are differences between the items, which cannot be explained completely by the construct or latent variable. Thus, a residual error term is associated with each item. (Hox & Bechger, 1998)
Open Data	"Open Data is a philosophy- and increasingly a set of policies – that promotes transparency, accountability, and value creation by making government data available to all. [...] By encouraging the use, reuse and free distribution of datasets, governments promote business creation and innovative, citizen-centric services." (OECD, 2018)
Performance Expectancy	"The degree to which a person believes that using a particular system would enhance his or her job performance." (Davis, 1989)
Reciprocity	We define Reciprocity as the feeling of obligation to share data based on the fact that other members of society already do so for common benefit.
Regression weight	Standardized weights, that show how much the criterion variable increases when the predictor variable is increased by one. In this case, it displays specifically, how much of the variance of one variable is explained by the other. (Courville & Thompson, 2001)
Self-Efficacy	We define Self-Efficacy as the consumer's own assessment of the ability to be skilled and knowing the process of data-sharing.
Social Influence	"The degree to which an individual perceives that important others believe he or she should use the new system." (Venkatesh, et al., 2003)
Social Referral	We define Social Referral as the referral of family members and friends to share data.
Structural Equation Modeling (SEM)	"A very general statistical modeling technique, which is used in the behavioral sciences. It can be viewed as a combination of factor analysis and regression or path analysis. The interest in SEM is often on theoretical constructs, which are represented by regression or path coefficients between the factors." (Hox & Bechger, 1998)
Tailored Product	We define Tailored Product as a personalization of the service a consumer gets based on individual characteristics.
Transparency	We define Transparency as the availability of information regarding the collection and use of data observable for the consumer.
Trust in Business	We define Trust in Business as the perception of a consumer regarding the trustworthiness of a company.

III. Table of Contents

IV. List of Figures

1. Introduction

The introduction highlights the importance of the chosen topic and introduces the research question. It comprises a pre-study, which was conducted to narrow down the research scope and to test the intended research feasibility.

1.1 Background

"Ironically, although data is becoming ever more important, data about data is still hard to find."
- The World Bank, 2019

During the last years, data has increasingly gotten in the focus of companies, governments, and consumers (McKinsey, 2011). Especially from the company perspective, data is seen as the new resource that leads to competitive advantages in many industries (Economist, 2017; World Economic Forum, 2011). Data is valued equal to power in the case that by collecting data, a firm has more scope to improve its products, which attracts more users, who then generate even more data – a beneficial virtuous circle (MIT Technology Review & Oracle, 2016). An underlying principle for that virtuous circle is a sequential process in which consumers are generating data in the first place, which is then used by companies to optimize their operations.

In the last two decades, operation optimizations have especially been developed by companies in the transportation industry, in which large amounts of transportation data have brought unprecedented opportunities (Zheng, et al., 2016; Zhang, 2011). It is suitable for this research project to refer to transportation data due to three reasons: First, it is relevant because there is an increasing pressure on having a reliable and efficient urban transportation system as cities continue to grow in population whereas available land remains constant (Noland & Polak, 2002). Many scholars have already analyzed the potential of data analytics in urban passenger transportation, like Banister (2008) and Urry & Lyons (2005). Second, the amount of data supplied is increasing rapidly due to new data collection methods of auxiliary instruments, such as cameras, Global Positioning System (GPS)-based receivers, and microwave detectors (Zhang, 2011). Third, many new business models are emerging in the transportation sector, for example, numerous shared mobility services, like car- or bike-sharing services and e-scooters, which address the gap in supply and demand for sustainable mobility in cities (Firnkorn & Muller, 2011).

Many actors are interested in the access to transportation data and given the fact that there is a lot of demand for data, many scholars have taken various attempts to analyze the concept of *the value of data* (Grossklags & Acquisti, 2007; Lesk, 2012; Jentzsch, et al., 2012). One attempt is to identify the economic value of data (Pemberton Levy, 2015; Muschalle, et al., 2013; Schomm, et al., 2013), i.e. to assess the potential economic benefit one could gain from owning a certain transportation dataset. Another attempt is to interpret the value of data by analyzing the perceived value of it (Grossklags & Acquisti, 2007; Lesk, 2012; Jentzsch, et al., 2012), i.e. to assess how much a chosen actor would perceive a specific transportation dataset to be worth. In this second attempt, the valuation of the data is dependent on the chosen actor and under which circumstances these actors are willing to share their data. In order to find a perspective to interpret the value of data in the inner-city mobility context and to further understand the important actors in this environment, we have conducted a pre-study, which is described in the next part.

1.2 Pre-Study

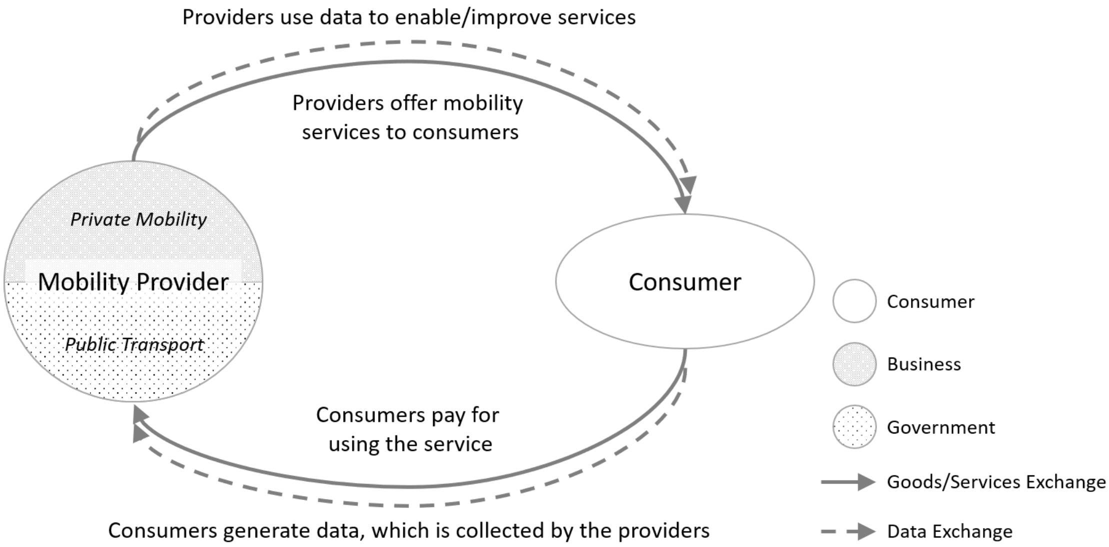
1.2.1 Purpose and Execution of the Pre-Study

As suggested by Holme & Solvang (1997), the purpose of the pre-study was to narrow down the research scope and simultaneously test its feasibility. We conducted interviews to get an overview of the transportation sector and to comprehend different perspectives from its various actors. Bryman and Bell (2015) suggest conducting semi-structured interviews by asking both pre-formulated questions and situationally related questions. Based on an initial read up on literature about the transportation sector (McKinsey, 2012), we chose to cover three main topics, which included (1) the most important trends in the inner-city mobility sector, (2) how the company/organization is currently collecting and using data, and (3) who the main business partners of the company/organizations are as well as their relation to them. We conducted five one-hour interviews with experts from relevant businesses and organizations linked to the transportation industry in Sweden (see Exhibit A). The interviews were analyzed by a simplistic content analysis technique, where we focused on the text content and actual messages of the interview and not aimed to interpret the abstract meaning behind each interviewer's input (Erlingsson & Brysiewicz, 2017). To ensure the ethics in qualitative research as suggested by Brinkmann and Kvale (2005), we informed each pre-study interview partner about the outline of our research and their participation as well as their right to withdraw at any time. In the end, all five interview partners gave permission to disclose the discovered material, their names, and their associated organization.

1.2.2 Relationship between the Mobility Provider and the Consumer

Before conducting the pre-study interviews, we did initial research on the transportation sector (McKinsey, 2012; Choudhury, 2018), which enabled us to sketch a simplified relationship of the typical actors in the mobility context (see Figure A). Simplistically, there are two parties involved, namely the mobility providers – divided into private businesses and governmental public transportation organizations – and the consumers. Both are interacting with each other in two dimensions. The first dimension is the trade of mobility services in exchange for money. The second and less visible dimension often underlies this trade – the exchange of data that is generated by using the mobility services and then collected and used by the mobility providers.

Figure A: Simplified Relationships between Mobility Providers and Consumer



After conducting the pre-study, we gained a better understanding of the relationships between mobility providers and consumers. We augmented Figure A to derive to an extended relationship of actors in the mobility context (see Figure B).

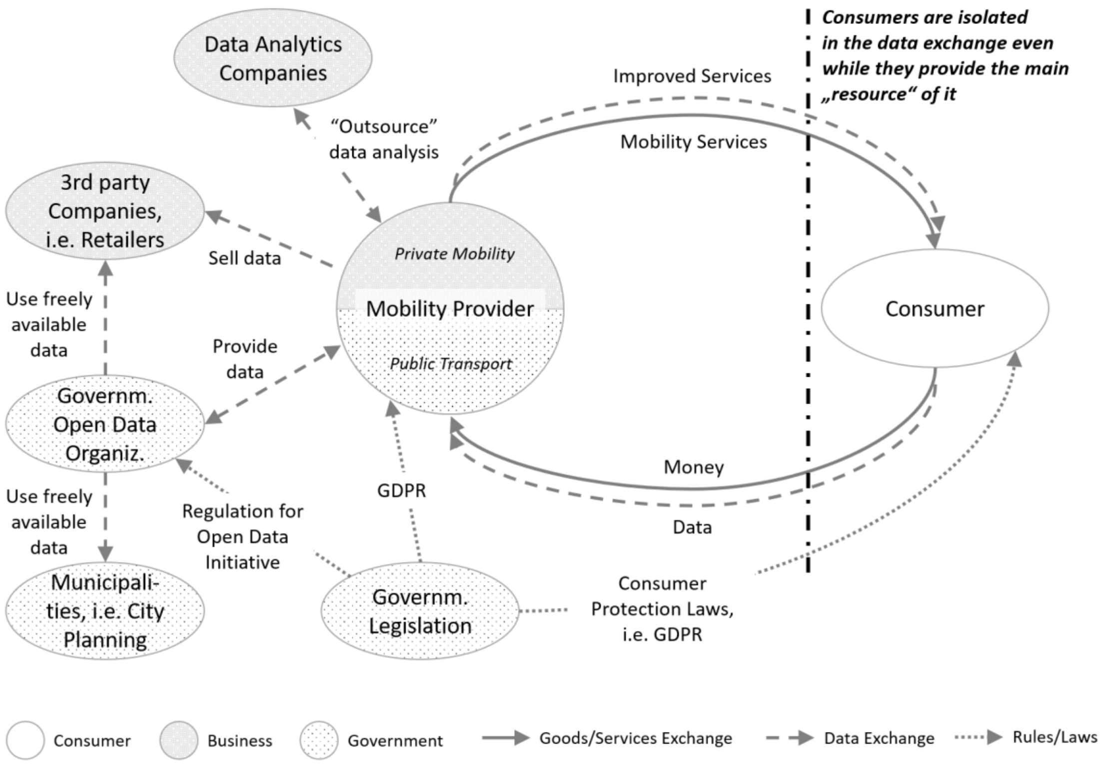
EXTENSION 1: THE GOVERNMENT

Apart from being active in the form of public transport providers, there are three additional interests of the government. First, governments are interested in city planning as highlighted by Elias Arnestrand from Samtrafiken. They aim to derive insights about current traffic and movement patterns to make profound decisions when planning infrastructure development. Second, the government has the responsibility to ensure a well-working legal system to protect businesses as well as consumers from abuses. The government seeks to regulate the markets by identifying what needs to be set as legal borders and thus creating the rules and practices of data exchanges. Third, the government is concerned about the general economic environment of a country and therefore it encourages business innovation through Open Data initiatives, e.g. allowing startups to use public data to develop products and services as mentioned by Peter Popovics from the Stockholm School of Economics. The government owns a lot of transportation datasets, which are made openly available to encourage innovation. For example, Arnestrand highlighted that the Swedish state has a specific legal regulation that makes data exchange between public transportation providers obligatory, which leads to the existence of the non-profit organization Samtrafiken – a joint-venture of those public transportation providers. The Swedish context is discussed further in Chapter 1.4.

EXTENSION 2: THIRD-PARTY COMPANIES

We furthermore identified two categories of third-party companies through our pre-study. First, there are companies like retailers, who are interested in buying data from the mobility providers to make use of the datasets for their own business. These companies also include specialized data aggregator firms, who act as intermediaries that combine and preprocess data to sell it further. Second, data analytics companies like Peltarion are specialized in generating insights from other companies' data. Both of these types of third-party companies have an interest to get as much data as possible, as it is their main resource of doing business.

Figure B: Extended Relationships between Mobility Providers and Consumer



1.2.3 Outcomes of the Pre-Study

Elaborating on the value of data in the mobility context, there are mainly two different approaches – the economic value of data and the perceived value of data – as mentioned before. Based on our pre-study, we decided to focus on the perceived value of data from the consumer perspective due to two reasons.

First, there are already many companies attempting to determine the value of data economically. As mentioned by interview partners from Samtrafiken, Peltarion, and the e-scooter sharing provider Voi, there are many businesses, including incumbents and startups, which share transportation datasets among each other. Facilitating platforms like the Open Data portal "Trafiklab", which has been established by Samtrafiken, are intermediating data among many players and hence, there is a lot of data exchange and valuation going on, which has already been investigated to a large extent.

Second, there is more uncertainty about the perceived value of data from the consumer perspective. As mentioned by Arnestrand, many organizations like Samtrafiken are mainly intermediating between businesses, the government, and third parties. We derived from our interviews that the consumer has gotten very little attention, even though it is their data that is exchanged among the businesses and third parties to create economic benefits. Furthermore, Arnestrand argued that Samtrafiken actively chooses not to focus on consumers, and therefore there is little understanding of the consumer's perception. Since consumers are not actively selling their data, we can only interpret the perceived value of their data by finding out under which circumstances they are intending to reveal and share it with third parties, as suggested by Popovics.

1.3 Research Question

In this project, we aimed to gain a holistic understanding of the various interrelated factors that impact a consumer's intention to share one's dataset in the Swedish mobility context. Hence, the following descriptive research question emerged.

What are the factors that influence the consumer's intention to share data in the Swedish mobility context?

In addition to the descriptive research question stated above, we decided to add a supporting explanatory question to explore the topic further.

Why do these factors influence the intention to share data?

We conducted a mixed methods study to answer these two research questions. Mixed methods research is defined as combining quantitative and qualitative research techniques into a single study (Johnson, et al., 2007). We argue that a mixed methods approach suited our topic best, because it is appropriate for analyzing broad patterns when there is a need for in-depth explanations, that a single quantitative study is unable to provide (Bryman & Bell, 2015). First, we conducted our quantitative main study to answer the first research question and then we conducted a qualitative follow-up study to answer the second question.

1.4 Research Scope

Defining the research scope, we focused on two themes, which we describe in the following part together with their implications on the research project.

FOCUS 1: THE UNIQUE INSTITUTIONAL STRUCTURE OF SWEDEN

As mentioned by Kye Andersson from Peltarion, there are differences among countries regarding the institutional structure and the development of a legal data framework. Whereas in the USA a liberal attitude towards the businesses and their usage of data has evolved, in China it is the government centralizing the usage of data (Wu, et al., 2011). Looking at Europe, Arnestrand pointed out that Sweden has a very unique institutional structure regarding the use of transportation data, since third-party collectors and mediators like Samtrafiken have already been established for more than 20 years and Sweden itself ranks at the third place in Europe regarding the development of the government digitalization (European Commission, 2017). In Sweden, there is a high interaction of multiple stakeholders, such as businesses, the government, and governmental institutions such as Samtrafiken, which stimulates further interest in the institutional structure of the Swedish transportation system.

FOCUS 2: DATA-SHARING NOT NECESSARY TO USE SERVICE AND LOW-SENSITIVE TRANSPORTATION DATA

When evaluating data-sharing in the transportation sector, the term *transportation data* had to be defined first. In the pre-study, Mehdi Rafinia from Samtrafiken pointed out that various levels of data have to be distinguished in the transportation context, and hence we defined the data that is shared in this study by two delimitations. First, consumers needed to be freely able to decide whether they share or not share their data, and thus, sharing the dataset should not be necessary in order to use the mobility service. The second limitation for the transportation data is that we did not focus on sensitive datasets, e.g. we did not focus on private credit card data connected with payments for using a transportation service, as it would shift the focus too much on privacy and security concerns. We thereby refer in the remainder of this project always to a transportation dataset that consists of the information of a consumer who is going from point A to point B with the associated timestamps and geolocations as well as the mean of transportation he or she uses.

1.5 Purpose and Expected Contribution

1.5.1 Theoretical Contribution

As mentioned in the outcome of the pre-study, there are many actors, especially businesses and the government that are interested in transportation data. However, most of them are ignoring the consumers, who actually generate transportation data. This is problematic because there is little theory on the consumer's perception on the value of data and therefore, in this project, we have the purpose to shed light upon the factors that impact the consumer's intention to share their dataset. Our study, combining existing theories in a previously unexpected relationship, to examine the holistic view on the interaction of various factors impacting a consumer to share a dataset contributes towards assessing the value of data in general and hence contributes to existing theory (Dinev & Hart, 2006; Berendt, et al., 2005; Grossklags & Acquisti, 2007). As indicated in Figure C, the purpose of this research project is to fill the research gap of the consumer's intention to share data as further identified in Chapter 2.1.3 and to demonstrate two implications for businesses and the government.

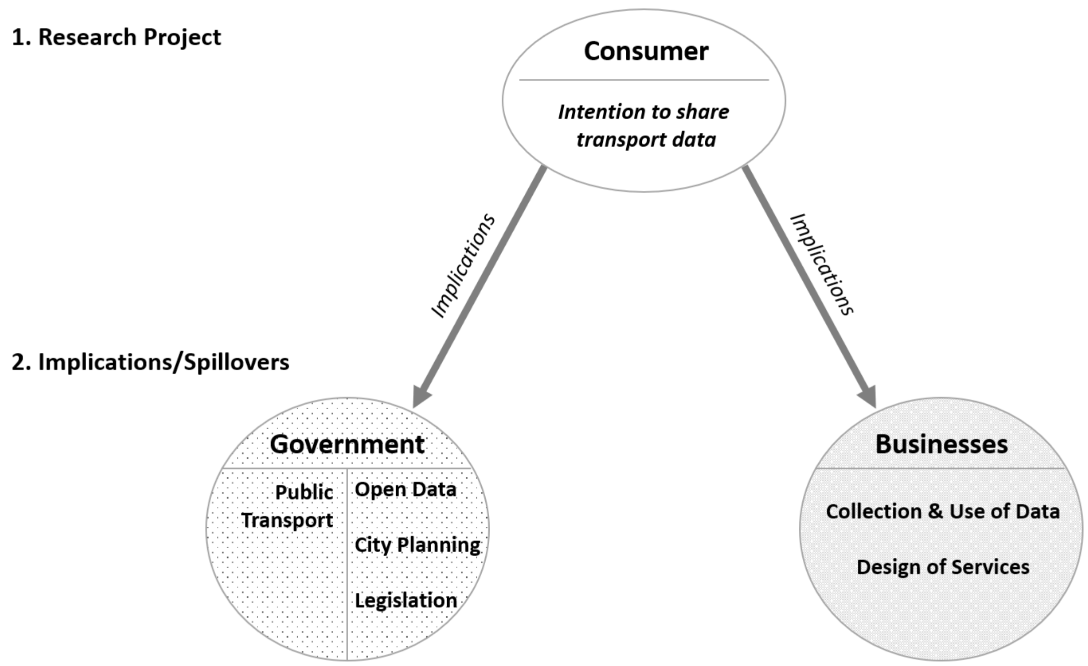
1.5.2 Implications for the Businesses

There are spillover effects to businesses because companies could benefit by targeting the consumer more specifically regarding data-sharing incentives. Whereas much literature in both academic research and managerial journals is about how businesses use transportation data to optimize their operations (Cohen & Kietzmann, 2014; Rehder, 2018; Adell, 2009), less research is conducted on how they actually retrieve data from consumers. Hence, determining why a consumer would intend to share their data can help businesses to improve their collection and use of data, and support them in designing their transportation services.

1.5.3 Implications for the Government

Furthermore, there are spillover effects to the government. The government has the mandate to create a legal environment to both support businesses as well as protect the individual (Cohen & Kietzmann, 2014). There have been many attempts to analyze the optimal relationship between businesses and the government to achieve the common objective of improved mobility (Magalhaes & Roseira, 2017; Chun, et al., 2010). If the government understands what factors are those to impact a consumer's intention to share data, the government can adjust the regulations on data ownership so that the transportation businesses and consumers can interact better. Additionally, governmental organizations concerned with Open Data initiatives and city planning are able to profit from our research if they can use our outcomes to optimize their handling towards consumers in regard to the data (Magalhaes & Roseira, 2017).

Figure C: Expected Contribution of the Research Project



1.6 Research Outline

This research project contains seven chapters. In the first chapter, the general topic was introduced, and the pre-study was explained that helped to narrow down the research scope and to find the two research questions. The remainder is structured as follows. In the second chapter, relevant academic literature is reviewed in order to derive to our theoretical model. In the third chapter, we explain the methodology of this research project concerning the main and the follow-up study. The results received from the main study, namely the questionnaire, are presented and analyzed in the fourth chapter. In the fifth chapter, we present the results from the follow-up study, which are based on the focus groups. In the sixth chapter, we discuss the ooutcomes from both research studies to compare those with theoretical expectations. Finally, the seventh chapter concludes our research by summarizing the main findings, discussing the implications and suggesting further research projects.

2. Theoretical Background

The theory section identifies and discusses relevant previous academic work. Referring to existing literature helps to build upon established knowledge and to further identify the research gap that this research project aims to fill. Afterwards, our own theoretical model is developed.

2.1 Literature Review

2.1.1 Role of Data for Businesses and Governments

There are many different causes contributing to the vast amount of data available. One source of generating data is the interaction between businesses and organizations because recorded transactions create data (MIT Technology Review & Oracle, 2016). Another source of providing data is Open Data, referring to the government making data publicly available, usually by supplying anonymized data (Vetrò, et al., 2016). According to the OECD, Open Data is “a philosophy- and increasingly a set of policies – that promotes transparency, accountability, and value creation by making government data available to all” (2018). An increasing amount of governments are becoming in favor of the Open Data mentality (Janssen, et al., 2012) and McKinsey analysts estimate that Open Data can potentially stimulate $3 trillion in benefits throughout the global economy through better decisions, new products and services, and greater transparency and accountability (Chui, et al., 2014).

According to a study of HM Treasury, firms adopting a data-driven decision-making process can have 5-6% higher output and productivity (HM Treasury, 2018). Companies are collecting trillions of bytes of information throughout their entire value chain, including customers, suppliers and their own operations (Manyika, et al., 2011). The vast amount of data, also referred to as Big Data, offers a lot of opportunities to create more value such as improving performance, segmenting populations to customize actions, replacing or supporting human decision processes and many more (Manyika, et al., 2011; Loebbecke & Picot, 2015). According to Mauro et al. (2016), Big Data describes “the Information asset characterized by such a high volume, velocity, and variety to require specific technology and analytical methods for its transformation into value”. However, this definition does not only emphasize the advantages of Big Data but also highlights its two major challenges (Spiekermann, et al., 2015), that are discussed in the following.

First, businesses are challenged to successfully extract a valuable and relevant analysis from a big amount of data. Ironically, the vast amount of available data could complicate the value creation in the commercial use for companies (Zuiderwijk, et al., 2014; Spiekermann, et al., 2015). It is the ability to gather and aggregate it effectively that proposes a challenge for the businesses. Merely a high volume of data is of little use if no insights are being generated, thus the collection of high-quality data is key.

The second challenge is that there is no universal consensus on data ownership and control of data (HM Treasury, 2018). In order to decrease confusion, Rock and Moran (2018) created a data ownership framework with four entities, data originator, primary data owner, co-owner of data, and enabled parties. Though in practical, it is still ambiguous how to differentiate each of these parties from another and how to identify the respective rights and responsibilities. This research project focuses on the consumer, who acts as data originator and who faces the decision to share or not to share their data with enabled parties.

2.1.2 Role of Data in the Swedish Mobility Sector

Data plays a distinct role in the mobility context because granular data analysis enables different opportunities, such as more efficient operations of the traveling service or improved customer service when moving from one point to another (Choudhury, 2018). There is a major underlying shift in the mobility context, starting in the 19th century with the industrial age in which railways and fixed public transportation were built, then moving towards the 20th century in which the car emerged empowering the mobility of the individual. In the 21st century, transportation is becoming even more tailored to the individual consumer by offering customized products (Goodall, et al., 2017). One example of offering tailored products in the mobility context is the concept of mobility-as-a-service (MaaS). MaaS combines "different transport modes to offer a tailored mobility package, similar to a monthly mobile phone contract and includes other complementary services, such as trip planning, reservation, and payments, through a single interface" (Hietanen, 2014). MaaS offers tailored transportation suggestions to consumers because they share their data, exemplary their GPS-location (Jittrapirom, et al., 2017). Hence, the success of mobility innovations such as MaaS relies both on technical conditions and social factors, which are shaped by the interaction of the business, the consumer, and the institutional structure (Teece, 2010). Again, a requirement for a business model innovation like this is a consumer who is willing to share their dataset, so that they can receive customized products in return.

Sweden has a very unique institutional structure that deserves further investigation. The supply of transportation data in Sweden is heavily fostered by governmental agencies such as Samtrafiken, Trafiklab or Trafik Analysis (Sandberg, 2014). The community Trafiklab is "a place that developers can share data and APIs for public transport in Sweden" (Trafiklab, 2019). In 2016, Trafiklab has founded the project Kraftsamling Öppna Trafikdata (KÖT) aiming to find a common national goal for public transportation (Lund, 2017). KÖT is a central pillar connecting six regional public transport agencies, the Swedish Transportation Administration, Samtrafiken and external third-party developers who are about to publicize Open Data from all actors in the transport industry to enable developers to create new smart digital services (Lund, 2017). In Generally, there are many other related initiatives to KÖT in Sweden like Drive Sweden, MaaS Alliance, and the Swedish Mobility program (Lund, 2017).

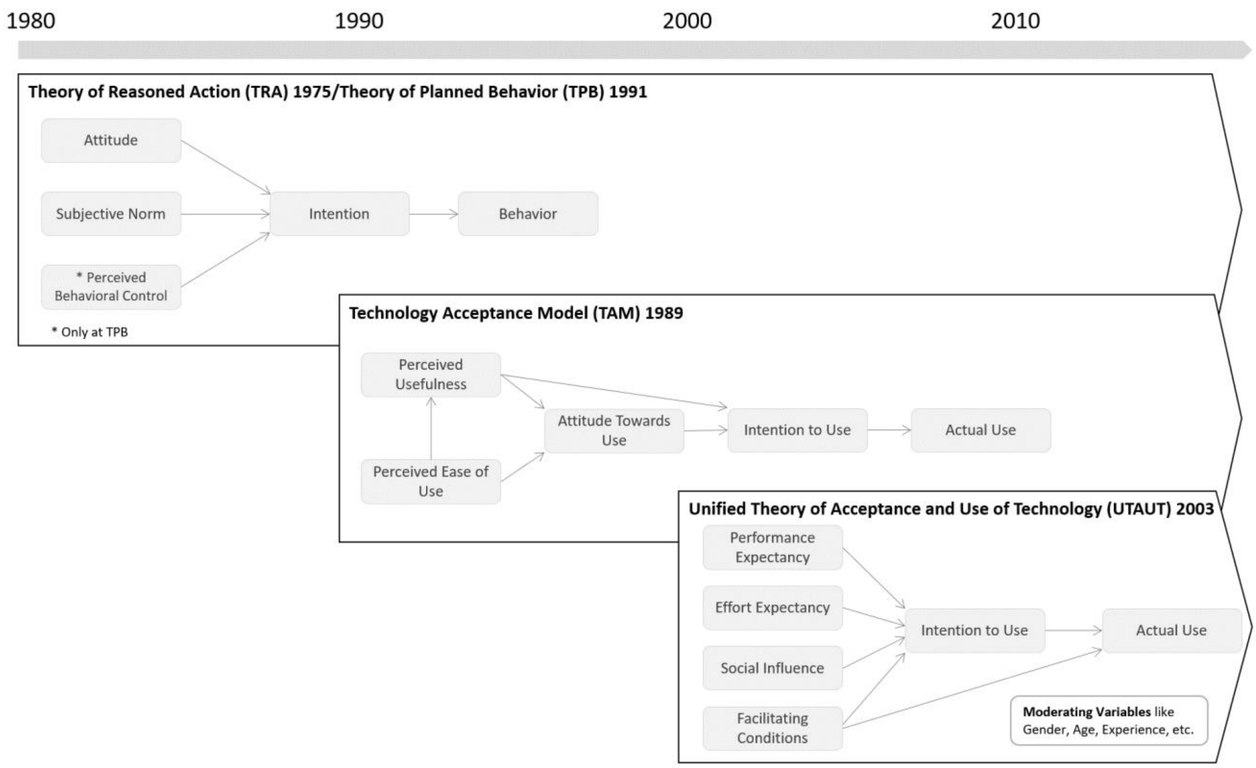
2.1.3 Research Gap: Consumer's Intention to Share Data

There is a lot of research on how transportation businesses use Open Data systems to optimize their business operations (Zheng, et al., 2016; Zhang, 2011). In order to support Open Data initiatives, like for the transportation system data, the feedback and insights of users are needed so that the Open Data initiative can continuously be improved (Janssen, et al., 2012). So far, the vast majority of academic research on the consumer focuses on data privacy concerns including scholars coming from various research areas like legal studies, philosophy, marketing, and consumer behavior (Solove, 2005; Iachello & Hong, 2007). However, there is little research on data ownership from the perspective of the individual consumer, e.g. how they think about data ownership, how they value data, and under which conditions they share their data. There is a theoretical gap about the consumer's intention to share data and this is in line with the outcomes of the pre-study mentioned in Chapter 1.2.3. In order to fill this research gap, we examine a holistic view of the various interrelated factors that influence a consumer's intention to share their data.

2.2 Theoretical Framework

In the following, we introduce existing theories as basis for our own model attempting to explain the factors that influence the intention to share, as illustrated in Figure D. We discuss these existing theories in Chapter 2.2 and introduce our own adopted model in Chapter 2.3.

Figure D: Evolution of Consumer Behavior Models towards UTAUT



2.2.1 Consumer Behavior Theories

Plenty of researchers from different academic areas have developed theories and models to set the basis for describing consumer behavior (Fishbein & Ajzen, 1975; Ajzen, 1991; Davis, et al., 1992). One of the first theories is the *Theory of Reasoned Action* (TRA) by Fishbein and Ajzen (1975), which is based on social psychology and aims to predict and explain human behavior in a given context. The theory states that a person's behavioral intention of conducting a specific task is influenced by the attitudes towards it and the subjective norms in the personal environment of the individual. According to Kim et al. (2009), "TRA is very general in nature and attempts to explain almost any human behavior." Ajzen (1991) further extended the TRA to the *Theory of Planned Behavior* (TPB). Ajzen added the construct of perceived behavioral control into the model to account for the person's belief in having control over the performance of this task. The TPB has become one of the most influential ones in this field of research and is still commonly used nowadays (Ajzen & Sheikh, 2013; Rise, et al., 2010).

Both TRA and TPB are consumer behavior models that are structured around the conceptual belief that various factors influence the intention to perform a specific task or behavior, which then leads to the actual performance or behavior. These models have been used as fundamental theories to study a variety of different consumer behaviors in diverse contexts, such as violations of speed limits (Wallén Warner, 2006) and the intention to try nano-foods (Chang, et al., 2017).

2.2.2 Technology Acceptance Model

Among the many adoptions of the TRA and TPB, the *Technology Acceptance Model* (TAM) has established itself as one of the most profound theories to explain consumer intentions regarding the acceptance of new technologies (Kim, et al., 2009). Davis (1989) developed the TAM based on the TRA and extended it with elements from other behavior theories, such as the Self-Efficacy Theory and the Innovation Diffusion Theory. Looking at its components, the TAM attempts to explain the dependent variable *Actual Use*, which is influenced directly by *Intention to Use*. Intention to Use is influenced by two factors: the *Perceived Usefulness*, dealing with the degree to which a person believes that using a particular technology enhances their job performance, and the *Perceived Ease of Use*, focusing on the degree to which a person believes that using a particular technology is free from effort (Davis, 1989).

The TAM has gained popularity in academia, especially when explaining adoption behavior in the information technology context, which newly emerged in the late 20th century (Hazen, et al., 2015). Additionally, Davis, Bagozzi, and Warshaw (1989) praised the TAM because it is as a "parsimonious and theoretically robust model which is applicable to the acceptance analysis and prediction of a broad range of computer-based technologies and in various contexts". The application of TAM instead of general consumer behavior models is thus especially suitable when focusing on user's adoption behavior towards technology-related systems (Taylor & Todd, 1995).

2.2.3 Unified Theory of Acceptance and Use of Technology

Building upon the TAM, Venkatesh et al. (2003) proposed a further developed model called the *Unified Theory of Acceptance and Use of Technology* (UTAUT) to better describe and predict technology acceptance. This theory integrates a total of eight prominent models from the field of user acceptance literature to develop a model that explains the individual's adoption behavior (Im, et al., 2011).

Developing the TAM further, the UTAUT integrated two major changes. The first change is that this theory uses four factors to explain the Intention to Use and Actual Use, whereas the TAM used two factors. The four UTAUT factors are *Performance Expectancy, Effort Expectancy,* and *Social Influence* and *Facilitating Conditions*, which are said to explain the Intention to Use and the Actual Use. Two out of these four UTAUT factors are derived from the TAM model, i.e. Performance Expectancy reflects the Perceived Usefulness whereas Effort Expectancy reflects the Perceived Ease of Use (Venkatesh, et al., 2003). The second change from the TAM to the UTAUT is that the researchers added a set of moderator variables that are assumed to influence the key relationship between the four UTAUT factors and the dependent variables Intention to Use and Actual Use.

The authors of the UTAUT recommend refining the model to deepen the understanding of dynamic influences on user acceptance (Venkatesh, et al., 2003). Therefore, this theory is constantly explored further, adapted and continuously improved in academia to better explain context-specific circumstances (Luarn & Lin, 2005) and it has been developed further along to refer to new emerging technologies. Hazen et al. (2015) states that "through each of these extensions, a variety of additional explanatory variables have been introduced", which emphasizes that the UTAUT can be adapted and applied to various contexts.

2.3 Synthesis: Modifying and Extending the UTAUT model

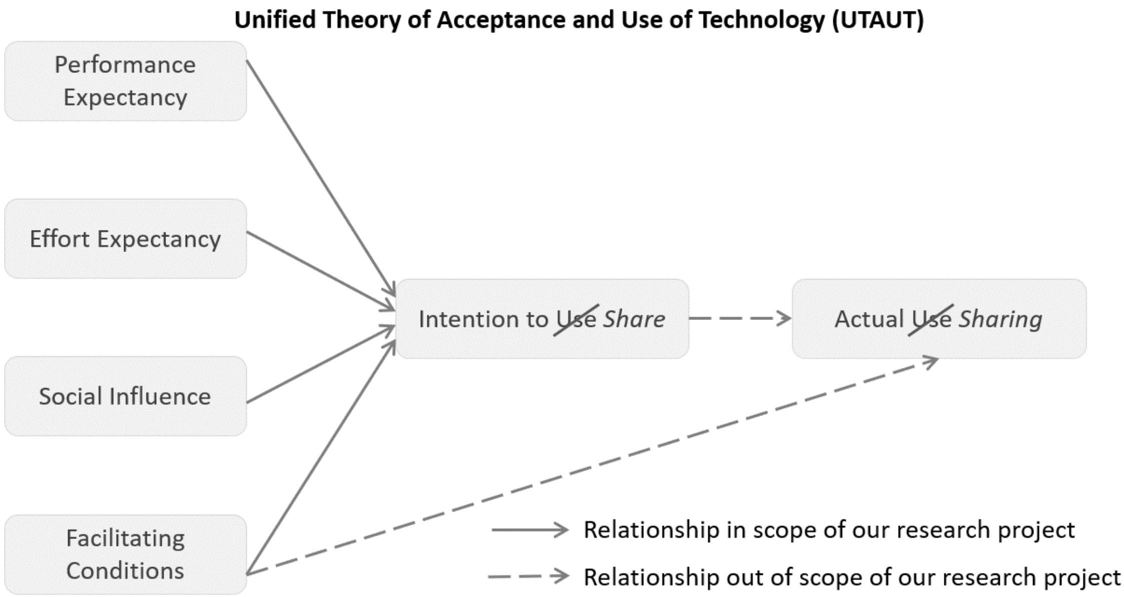
2.3.1 Introducing Our Adapted UTAUT Model

We took the UTAUT concept from Venkatesh et al. (2003) as the basis of our own theoretical model because this theory better suits complex scenarios of technology adoptions (Naranjo-Zolotov, et al., 2019). In the following part, we describe the two modifications compared to the original UTAUT (see Figure E).

The first modification was that instead of analyzing the acceptance of using technologies, we wanted to analyze the acceptance of sharing one's data. Hence, we replaced Intention to Use with *Intention to Share* and we replaced Actual Use with *Actual Sharing*, respectively. We can justify this modification because the UTAUT is a suitable explanation for describing adoptions of not only tangible technologies, e.g. the MP3 player adoption, but also has been validated for intangible technologies, e.g. online banking adoption or driver support systems adoption (Im, et al., 2011; Adell, 2009). We argue that the consumer's decision-making process to share or not to share their data is analogous to the decision-making process to use or not use new technologies.

The second modification was that we focused on Intention to Share as the final dependent variable and hence took Actual Sharing out of scope, due to three reasons. First, nowadays data-sharing in the mobility context is developed and implemented by only some services and is not the industry standard for transportation services. Prior research suggests that "Actual Use can be replaced by Intention to Use when the technology is still undergoing development, has a limited number of users, and when the objective of the research is to predict future use" (Tsai, 2014). Second, we collected data by introducing a fictive scenario and therefore a quantitative measurement of Actual Sharing is difficult to implement. Finally, the causal relationship between Intention to Use and Actual Use has been highly researched and empirically approved, thus replacing the variable usage by the variable intention is not a significant limitation to our model (Mathieson, 1991). Due to these three reasons, we argue that there is only a small discrepancy between Intention to Share and Actual Sharing, which justifies the second modification.

Figure E: Modifications towards the UTAUT



In the following, we introduce all 16 hypotheses of our adapted UTAUT model, and an overview can be found in Exhibit B. Based on the original UTAUT model described in Chapter 2.2.3, we expected the four factors Performance Expectancy, Effort Expectancy, Social Influence, and Facilitating Conditions to have a positive effect on the Intention to Share.

H1: Performance Expectancy has a significant positive effect on Intention to Share.

H2: Effort Expectancy has a significant positive effect on Intention to Share.

H3: Social Influence has a significant positive effect on Intention to Share.

H4: Facilitating Conditions has a significant positive effect on Intention to Share.

Furthermore, we extended the original UTAUT model, which is encouraged by its authors because they recommended other researchers that they "must pick and choose constructs across the models, [conduct] a review and synthesis [from other theories] in order to progress toward a unified view of user acceptance" (Venkatesh, et al., 2003). From the original UTAUT model, we adapted the determinant Intention to Share, which will hereafter be called the latent layer 2 factor, as well as the four UTAUT factors Performance Expectancy, Effort Expectancy, Social Influence, and Facilitating Conditions, which will hereafter be called the latent layer 1 factors. Additionally to the latent layer 1 and latent layer 2, we added a group of constructs for each of the four latent layer 1 factors, so that we can got a more granular understanding of the UTAUT model. Adding another level of constructs has already been done by other researchers applying the UTAUT such as Tsai (2014), Chang et al. (2017), and Naranjo-Zolotov et al. (2019). In the following Chapters 2.3.2 until 2.3.5 we introduce our four extensions to the original UTAUT model. Hereby, we list the theory that we added and then explain how we derived a relevant construct and hypotheses from that theory into our adapted model.

2.3.2 Extension 1: Social Capital Theory

According to *Social Capital Theory* (SCAT), individuals make decisions not only based on their own beliefs but because they are influenced by other people (Chang, et al., 2017). Therefore, SCAT attempts to explain the relationship between individuals and their social environment (Coleman, 1988). According to Valenzuela et al. (2009), social capital refers to the social value that is generated by the interactions among the diverse members within a certain social network, including social ties such as trusting relationships as well as reciprocal behavior (Valenzuela, et al., 2009; Chow & Chan, 2008; Granovetter, 1973). When analyzing behavioral intention, SCAT assumes that social resources, like trust, networks, and social relations facilitate collective action (Adger, 2003; Chang & Chuang, 2011).

The UTAUT authors assumed that the Intention to Use is affected by social capital and therefore based on SCAT, they identified Social Influence as one of the four UTAUT factors (Venkatesh, et al., 2003). To reflect the perspectives of the SCAT literature that are the most relevant for our data-sharing model, we extracted the following four constructs*: Trust in Business, Institutional Trust, Social Referral* and

Reciprocity. Next, we describe each of these four constructs and their suggested hypothesis describing the effect on Social Influence.

First, we expect that if consumers value to trust a business, they are more likely to be affected by Social Influence when deciding to share or not share their data. This is similar to a study on the customers' acceptance of banking information systems, in which the researchers used the TAM and explained how trust in businesses has a positive effect on the intention to use the system (Reid, 2008).

> *H5: Trust in Businesses has a significant positive effect on Social Influence.*

Similarly, we believe that Institutional Trust is relevant as we expect trust in governmental institutions to have a positive influence on Social Influence. A government is generally in place to regulate a market and protect consumers (Cohen & Kietzmann, 2014). Therefore, if a consumer values Institutional Trust, we expect him or her to have an enhanced Intention to Share in this setting. In a study on trust and TAM in online shopping, researchers have proven that institutional trust positively affects the intention to adapt to the system (Gefen, et al., 2003).

> *H6: Institutional Trust has a significant positive effect on Social Influence.*

The third construct is Social Referral. Hereby, we suggest that the relationship to peers and their referral is an important element for consumers when making decisions. This is based on Lazaric and Lorenz (1998), who claimed that social referral is especially relevant when a consumer evaluates whether to engage into a sharing activity or not, which is the case in our data-sharing context and thus, we consider Social Referral to have a positive effect on Social Influence.

> *H7: Social Referral has a significant positive effect on Social Influence.*

The last construct taken from the SCAT displays the concept of Reciprocity. Woolcock (1998) argued that social capital includes the information, trust, and norms of reciprocity inherent in social networks. Therefore, we expect that if consumers feel that they benefit from being in a certain social network, their intention to participate in a reciprocal manner is higher, hence they would tend to give back to the social network they are in. This is in line with a study by Portes and Sensenbrenner (1993), who argued that reciprocity transaction is one important example of social capital and we assume reciprocity to have a positive effect on Social Influence.

> *H8: Reciprocity has a significant positive effect on Social Influence.*

2.3.3 Extension 2: Social Cognitive Theory

Developed by Bandura (1989), the *Social Cognitive Theory* (SCOT) is seen as one of the most powerful theories assessing human motivation and thought, and it has been widely used by researchers to study computer utilization (Venkatesh, et al., 2003). SCOT assumes that behavior, cognition, personal factors, and environmental factors are influencing each other and are therefore interrelated (Tsai, 2014). Out of the SCOT literature, we used two constructs in our adapted model, and we assume that *Self-Efficacy* and *Affect* have a positive effect on Effort Expectancy.

First, Self-Efficacy indicates the consumer's own assessment of the "capability to organize and execute courses of action required to perform a task" (Tsai, 2014). Based on the level of Self-Efficacy, humans are more or less likely to choose, perform, and persist in conducting the actions (Hasan, 2007; Bandura, 1986; Gist, 1987). In the IT context, Self-Efficacy is considered to be a decisive factor determining whether a user will adopt a technology or not (Venkatesh & Davis, 1996). According to Tsai (2014), a stronger belief of having the capability to perform a task results in less perceived effort to actually perform the task, and therefore we assume that the construct Self-Efficacy positively influences Effort Expectancy.

> *H9: Self-Efficacy has a significant positive effect on Effort Expectancy.*

The second construct derived from SCOT is Affect and it was already integrated into the UTAUT theory by Venkatesh et al. (2003). Likewise, Dweck and Leggett (1988) argue based on studies with children that Affect is a vital part of the behavior and adoption process because affection has a direct impact on "real-adaptive behavior patterns". Furthermore, they concluded that "affective processes [...] promote adaptive performance". We assume that someone who has a positive opinion, hence higher Affect, towards data-sharing in general, finds less perceived hindrances to actually share their data, which is reflected in Effort Expectancy.

> *H10: Affect has a significant positive effect on Effort Expectancy.*

2.3.4 Extension 3: Internal System Characteristics

Analyzing the acceptance of new technologies in the TAM model, Davis (1989) suggested that the design characteristics of the technology significantly influence the Perceived Usefulness. In the UTAUT model, Perceived Usefulness is reflected in Performance Expectancy, both dealing with the expectation to get something in return by using the system (Venkatesh, et al., 2003). Similarly, according to Hennig-Thurau et al. (2007), the likeliness to participate in a sharing activity is the highest if benefits are maximized and costs are minimized. Likewise, it is said that utilitarian motives seem to play a major role for the use of interactive services such as data-sharing (Bardhi & Eckhardt, 2012; Moeller & Wittkowski, 2010; Lamberton & Rose, 2012). In line with this, the importance of economic value and convenience has been underlined in various empirical investigations studying factors that influence (non)participation in carsharing services of (non)users (de Luca & di Pace, 2015; Lindloff, et al., 2014).

Based on the findings from the pre-study interviews and the literature review on MaaS, we have identified four constructs that describe the benefits of sharing transportation data. These four constructs are *Tailored Product, Increased Safety, Enhanced Convenience,* and *Economic Benefit* and they are supposed to have a positive effect on the latent variable Performance Expectancy.

> *H11: Tailored Product has a significant positive effect on Performance Expectancy.*
>
> *H12: Increased Safety has a significant positive effect on Performance Expectancy.*
>
> *H13: Enhanced Convenience has a significant positive effect on Performance Expectancy.*
>
> *H14: Economic Benefit has a significant positive effect on Performance Expectancy.*

2.3.5 Extension 4: Environmental Factors

Similar to the system characteristics, there are specific contextual elements that could impact the consumer's intention to adopt a behavior. As found by Venkatesh et al. (2003), there is an "organizational and technical infrastructure to support the use of the system", which plays a role for individual's decision-making. Unlike the previously outlined internal system characteristics, that include the benefits of sharing the data, environmental factors describe the way the system is designed and do not necessarily have to convey perceived benefits. Based on the insights from our pre-study, we derived the two constructs *Transparency* and *Guidance* as part of the situational differentiators that are expected to positively influence the Facilitating Conditions.

> *H15: Transparency has a significant positive effect on Facilitating Conditions.*
>
> *H16: Guidance has a significant positive effect on Facilitating Conditions.*

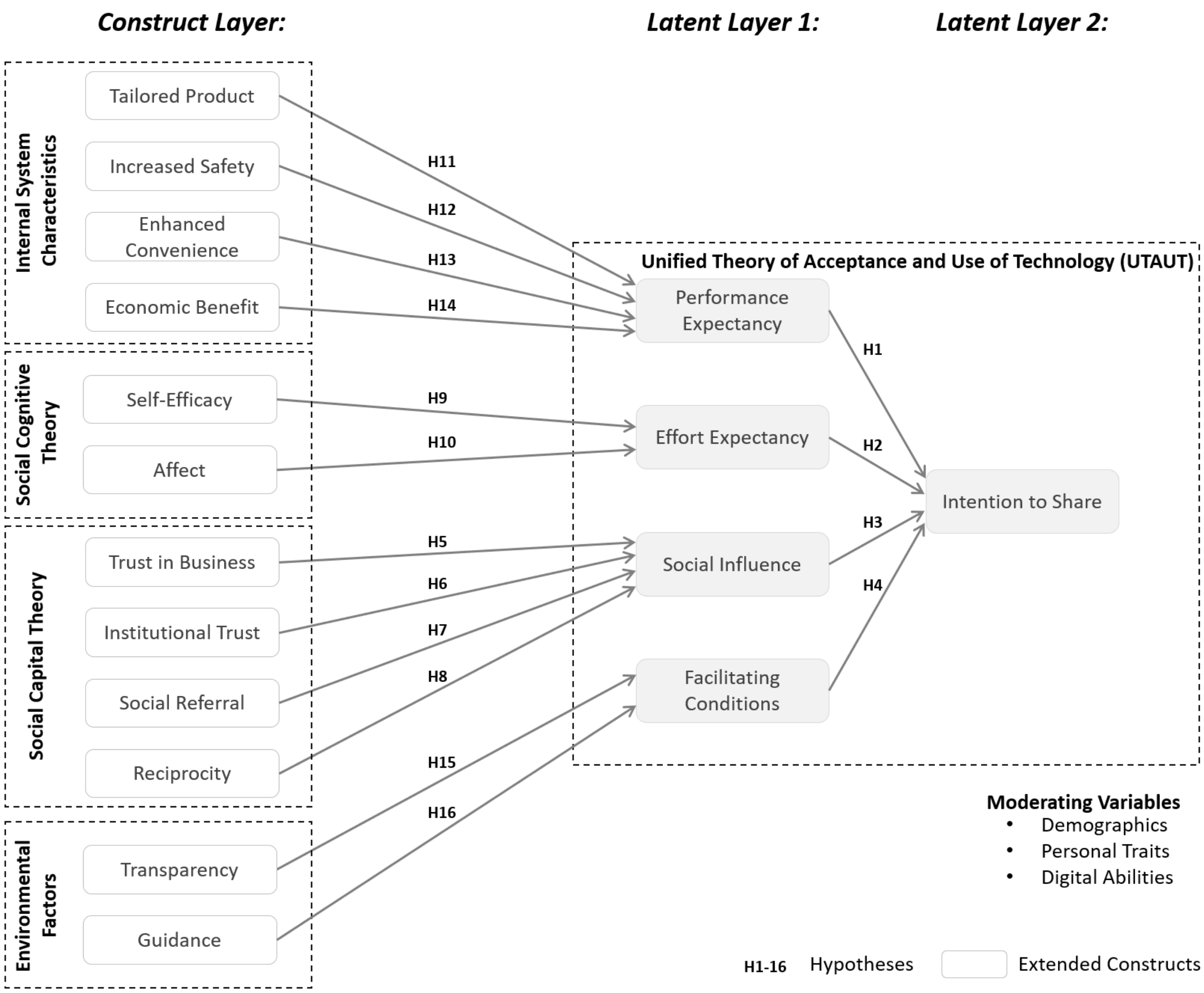
2.3.6 Moderating Variables

The developers of the UTAUT model introduced a set of moderating variables, namely Gender, Age, Experience, and Voluntariness of Use to inquire more information about the consumer (Venkatesh, et al., 2003) to advance the analysis. When Porter and Donthu (2006) applied the UTAUT, they advised other researchers to "conduct studies that explore the role of other belief or trait variables that could differentially impact members". Therefore, we also included moderating variables in our adapted UTAUT model to analyze different impacts on consumer segments. Based on the findings from our pre-study, we decided to include the following three sets of moderating variables into our model: Demographics, Personality Traits, and Digital Abilities (see Exhibit C).

2.3.7 Overview of Our Adapted UTAUT model

Based on the modified UTAUT and its contextual four extensions, we constructed the adapted UTAUT model, which can be seen in Figure F.

Figure F: Research Model based on the UTAUT



3. Methodology

The methodology section explains why we pursue a mixed methods approach. Furthermore, we describe the quantitative main study method and the qualitative follow-up study method.

3.1 Research Strategy

In this research project, we conducted a mixed methods study. For almost four decades, various scholars have conducted these studies by combining quantitative and qualitative research methods (Tashakkori & Creswell, 2007). Today's world is becoming increasingly interdisciplinary and complex, and therefore many researchers complement one research method with another (Johnson & Onwuegbuzie, 2004). We followed an explanatory sequential design, giving priority to the quantitative study, which was the main part of our project, and conducting a follow-up study to explore the rationales of the results from the main study (Creswell & Plano Clark, 2011). This is also called a quantitative dominant mixed methods research, because we relied on a quantitative view of the research process, and concurrently assumed "that the addition of qualitative data and approaches are likely to benefit the research project" (Johnson, et al., 2007).

In our research philosophy, we reject traditional dualism (e.g. rationalism vs. empiricism) and endorse pluralism by acknowledging the existence and importance of the physical world while simultaneously recognizing the emergent social world that includes subjective thoughts (Johnson & Onwuegbuzie, 2004). We use pragmatism as our research paradigm because it enables us to mix rationalism with empiricism, and we attempt to gain a holistic understanding of both studies to find a reasonable solution (Hoshmand, 2003). For our research progress, pragmatism helps us to decide which action to take next in order to better understand real-world phenomena and it offers a practical method of inquiry based on iterative action to further eliminate doubt (Johnson & Onwuegbuzie, 2004).

Our logic of inference was abductive because we engaged ourselves in the continuous cycle of reasoning (Feilzer, 2009). First, we created a theoretical model based on the pre-study and the literature review and then we chose our hypotheses of the model that would, if they are true, best explain the evidence we have gathered from the main study. Second, we conducted the follow-up study, to further investigate the topic and try to abductively uncover the best set of explanations for understanding the results from the questionnaire (de Waal, 2001). Coherent with our pragmatic philosophy, we started a loop, where we constantly attempt to improve upon past understandings in a way that fits in the world in which they operate (Johnson & Onwuegbuzie, 2004).

3.2 Research Design of the Main Study: Questionnaire

3.2.1 Purpose of the Questionnaire

We conducted the main study to attempt answering the first research question *"What are the factors that influence the consumer's intention to share data in the Swedish mobility context?"* We argue that a questionnaire is a suitable method to start our mixed methods study, as it allowed us to find correlations to show if the theoretically developed factors have a relationship towards the Intention to Share on a larger sample size.

3.2.2 Nature of the Questionnaire

In order to test our theoretical model, we developed an online self-completion questionnaire on the basis of best practices of Bryman and Bell (2015). The online questionnaire is assumed to be a suitable choice to reach out to a larger set of respondents for reasons of distribution and administration with low costs (Wright, 2005).

The questionnaire consisted of various parts. In the first part, we introduced the situation to the respondent (see Exhibit D) to set the mobility theme and describe the data-sharing context. Afterwards, we presented a set of statements for which respondents had to indicate their degree of agreeableness on a Likert-scale. We asked the participants to respond to 51 statements, which we define as items in the remainder, and these items were related to our constructs of the theoretical model (see Exhibit E). For each of the constructs and the latent layer variables, we chose to formulate multiple questions, which is favorable to account for possible outliers in the analysis due to misunderstood questions or personal cognitive associations towards specific wordings. This was a common approach in similar research studies that applied the UTAUT model (Attuquayefio & Addo, 2014; Chang, et al., 2017; Naranjo-Zolotov, et al., 2019). In addition to the statement questions about our adapted model's constructs, we asked about the moderating variables (see Chapter 2.3.6). All questions were randomized in its order of appearance to prevent biased responses.

Finally, we introduced the situational description about the mobility context at the beginning with a text and graphic. We placed the same graphic on the top of every page to ensure a coherent and cognitive reminder of the same situation and hence eliminate potential for biases as suggested by Dillmann (2007).

3.2.3 Coding of Answers

We used the online questionnaire software QualtricsXM from SAP to create the questionnaire and to collect and export the responses in a suitable manner for our statistical analysis. All multi-indicator questions were measured on an uneven ordinal Likert-scale from "0" to "10", which were labeled as strongly disagree ("0"), neutral ("5") and strongly agree ("10"). We expanded the typical 7-point Likert scale to a 10-point Likert scale to generate more distinct differences for the statistical analysis, which is suggested by Awang et al. (2016). This measurement allowed for comparison in statistical analysis, as we could compare the agreeableness towards specific constructs and their impact on the dependent variable. The values of the answer possibilities were consistently displayed to the questionnaire participant in the same order of ordinal nature (from 0 to 10), meaning that one can make claims about the relative comparison between different items of one survey (Norman, 2010).

3.2.4 Sampling

We distributed the questionnaire by sending an online link to network groups and to individual people via social media and personal messages. Inner-city mobility can be seen as a mundane topic, to which almost every person can relate to and therefore we targeted a broad sample and did not limit the choice of candidates a priori. It was more relevant that the respondents were able to understand the concept of sharing data and this is why we included the moderating variables of Digital Abilities. We included the other two sets of moderating variables to generate samples for our follow-up study based on the participant's responses.

The participation in our survey was voluntary and it was clearly stated that all data will be handled anonymously. Having distributed the self-completion questionnaire virtually, we could not physically observe the participant or make any judgments about their level of attention when filling out the questionnaire. Therefore, we included a control question to verify that the participant is reading each question carefully. In this question, we asked the participant to choose the answer "3" of the Likert-scale to prove that they are not just skipping through the questions without diligence.

3.2.5 Validity

In order to evaluate the academic strength of this research project, we discuss internal consistency reliability, face validity, and external validity (Bryman & Bell, 2015).

Internal consistency reliability describes whether the used questionnaire measures what it is designed to measure, hence the logic behind the theoretical model (Robson, 2011). Since we developed our own model and questionnaire, we first had to ensure its internal consistency. In the questionnaire, we asked for each of the constructs with the same amount of three questions and the same semantic structure, which ensures high comparability of the influencing factors to limit the potential for biases (Burns & Burns, 2008). Additionally, we designed the wording of the statements based on studies (Kim, et al., 2009; Venkatesh, et al., 2003; Attuquayefio & Addo, 2014), which already validated measures of the UTAUT. We are using similar constructs and adapt many of these formulations and adjust them to our data-sharing context. Moreover, we took our insights from the pre-study to formulate the statements. Finally, we analyzed the internal consistency reliability in a numerical manner by Cronbach's alpha to proof our questionnaire design (see Chapter 4.2.1).

In order to generate good face validity, we piloted the questionnaire with nine fellow students, who shared their thoughts with us when answering the questionnaire as it is recommended by Bryman & Bell (2015). We did this to improve the wording of the questions and to ensure a general understanding of the given scenario description and the 51 statements.

External validity describes whether the sample taken reflects the general population. As mentioned in Chapter 3.2.4, we did not focus on distributing to a broader audience reflecting the overall Swedish population but made a delimitation to target a sample that is acquainted with the concept of data sharing. We see this study as an exploratory approach to our research topic and focus on generating insights rather than disproving hypothesis as accurately as possible. Hence, this research project did not need to have a completely randomized sample, that accurately represents the entire population, but rather verifies whether the sample taken is able to validate that the questionnaire represents our theoretical model. This has been similarly done by other studies (Moreno, et al., 2014; Fortes & Rita, 2016).

3.2.6 Replicability and Generalizability

We decided to focus on the inner-city mobility context for our research. Nevertheless, we designed our quantitative study in a way to be generally applicable and replicable by other researchers, which is why our questions were not specifically directed towards a distinct company, i.e. a company case study, or focusing on a narrow target group, i.e. a demographic group. We expect other researchers, who use our model and replicate the questionnaire, to find coherent results to ours. This would prove high replicability of our research (Bryman & Bell, 2015).

Regarding the generalizability, it is important to us that our research can contribute to a high level of further research in this or related fields with suitable modifications, e.g. in other data sharing contexts not only linked to mobility. We assume that fellow researchers can use our theoretical model about the intention to engage in a specific interactive behavior. Hence, our research aims to have high generalizability, because it can appropriately be used in our specific mobility-context Sweden but is not limited to that specific context.

3.2.7 Analysis Technique

Structural Equation Modeling (SEM) is the widely used technique to analyze UTAUT models (Williams, et al., 2015). It is a general statistical technique for multivariate relations, which is commonly used in the behavioral sciences (Hox & Bechger, 1998; Byrne, 2009). To conduct SEM, a researcher initially specifies a structural model by proposing a path diagram between different elements a priori in order to define relationships between those elements according to theoretical concepts. By then integrating the observed data for all these elements, the SEM tool calculates all regressions of those paths, enlightening the patterns of the respondents (Moreno, et al., 2014). Anderson and Gerbing (1988) recommend using a two-step approach when conducting an SEM analysis. First, one should focus on a

Confirmatory Factor Analysis (CFA) to validate and potentially improve the proposed structural model. Through a CFA of the SEM, it is possible to assess how well the model fits the data and potentially to discover any previously undiscovered relationships that could be added by new paths (Kline, 2005). Thus, the proposed model can iteratively be re-specified to better fit the data. The second step is to evaluate the paths regressions of the model's elements to create insights about the relationships. The strength of SEM lies in its ability to giving a researcher the flexibility to shape and reshape the model to find intervening relationships between independent and dependent variables as well as latent variables, which are not directly observed (Hox & Bechger, 1998).

Our questionnaire was designed to fit the needs of conducting an SEM analysis, e.g. by asking several questions deriving to a single construct. We used AMOS 25 statistical software from IBM Statistical Package for the Social Sciences (SPSS) to evaluate the survey responses. This software is able to compute all common measurements and has a graphical interface of illustrating the paths for input as well as for output (Fathema, 2013).

3.3 Research Design of the Follow-Up Study: Focus Groups

3.3.1 Purpose of the Focus Groups

In the follow-up study, we aimed to answer the second research question, namely *"Why do these factors influence the intention to share data?"* We conducted focus groups to further explore the results that we have gotten from the main study. The focus groups supported us in eliciting a variety of views on data-sharing and to discover causal relationships and reasoning. We chose focus groups instead of individual interviews to allow people to comment on each other's inputs, probing different reasons and attitudes. Furthermore, we selected focus groups instead of group interviews, since the topic of a group interview span relatively widely, whereas in the focus group one specific theme is explored in depth (Onwuegbuzie, et al., 2009). Finally, in each focus group, we aimed to cover the same topics by following the same interview guide (see Exhibit F).

3.3.2 Nature of the Focus Groups

The group size has been between five and eight persons, which is in line with the recommended size according to Calder (1977) and Morgan (1998). We have conducted three focus groups because in the third focus group we have already encountered a lot of recurring answer patterns and hence we reached conceptual saturation (Calder, 1977). Furthermore, the instructor is important, who has two tasks to allow the discussion to flow freely and to intervene and bring out salient issues (Bryman & Bell, 2015). Overall, the instructor should guide the session in an appropriate manner that can be considered to be reactive and not proactively steering the discussion and we ensured to have the same instructor for all three focus groups (Onwuegbuzie, et al., 2009). We have chosen one of us researchers as the instructor and the other researcher to be present and take notes since we are already familiar with the theoretical model, which serves as an advantage when conducting and instructing the focus groups. This is in line with Knodel (1993), who advocates that the accuracy of the focus group analysis is enhanced if the analysts are involved with the actual data collection.

3.3.3 Sampling

In general, it is advised to choose interview participants based on their experience or involvement in a particular situation (Merton, et al., 1956). We selected the focus group participants not based on their experience with data-sharing, as we did not have any sufficient indication on this but based on their response in the questionnaire. Though the questionnaire was filled out anonymously, we gave the respondents the opportunity to leave their contact details if they are interested in participating in a follow-up study. This method enabled us in an ethical manner to pick out candidates that represent interesting patterns in their questionnaire responses, which we aimed to analyze further. We strived for identifying focus group participants, who are interested in sharing their opinions, and who scored high and low on the relevant moderating variables of Personality Traits and Demographics. This ensured a diverse set of people with different mindsets, which led to a more vivid focus group.

3.3.4 Analysis Technique

In order to analyze the focus groups, we adhered to the five stages framework by Ritchie and Spencer (1994). First, we familiarized ourselves with the content from the focus groups by complementing our observational notes from the focus groups, such as counting the votes in yes-no-questions or other impressions, with the transcript of the recorded audio information. Second, we linked the written transcript to our framework by including comments in the transcript. Third, we indexed the transcript by highlighting specific sentences and quotes. Hereby, we used color coding to mark each of the four UTAUT constructs with a different color. Fourth, we compared and contrasted the highlighted color coded quotes in a new table. As suggested by Knodel (1993), we used an Overview Grid, which is a large chart that provides a descriptive summary of the content of the focus group discussions, and we used a Microsoft Excel file for that. The various covered topics were listed on one axis and on the other axis we listed the three different focus groups. This technique enabled us to graphically get an overview of the extent of consensus regarding the same topic in various focus groups (Knodel, 1993). Finally, we mapped and interpreted the data from the Overview Grid. We did not only make sense of individual quotes but also interpreted the relationship between various quotes and linked them together.

4. Analysis Main Study: Questionnaire

In this chapter, we elaborate on the research findings of the quantitative main study and present the regressions between the different variables of our theoretical model.

4.1 Overview of the Main Study

We received 433 responses from March 13 until March 31, 2019, and removed 27 responses, because for those the control question was not answered correctly. Furthermore, we looked at the absolute differences of each construct's three items towards their mean and decided to erase another 15 responses from the set, as extreme outliers in some questions would have biased our analysis. Those 15 had given rather binary indications of 0 or 10 on the Likert-scale and even within the single construct groups, there were completely opposite answers due to this behavior. Thus, we had 391 responses to analyze after pre-cleaning (see Figure G).

Figure G: Overview of Demographics of Questionnaire Respondents

Category	Value	Amount	Percentage
Gender	Male	216	55.2%
	Female	173	44.2%
	Prefer not to say	2	0.5%
Age	Under 18	1	0.3%
	18-25	300	76.7%
	26-35	76	19.4%
	36-50	9	2.3%
	51-65	4	1.0%
	Over 65	1	0.3%
Education	Less than high school	2	0.5%
	High school graduate	14	3.6%
	Bachelor	133	34.0%
	Master/Diploma	239	61.1%
	Doctorate	3	0.8%
Location of living	City	325	83.1%
	Suburban area	51	13.0%
	Countryside	15	3.8%
Total		**391**	**100%**

Based on these demographics, our sample could be defined as younger adults with a higher level of education, who are mainly living in cities. The gender was fairly equally distributed. The moderating variables of Digital Abilities indicated that our sample was on average very digitally aware (Digital Understanding: mean 7.35, standard deviation 1.60; Digital Usage: mean 8.49, standard deviation 1.39; aggregated means of both items each). This was in line with the sampling approach, which we aimed for (see Chapter 3.2.4).

4.2 Structural Equation Modeling

4.2.1 Pre-Test for Internal Consistency Reliability via Cronbach's Alpha

Before we started the Confirmatory Factor Analysis (CFA) of the SEM, we ensured that our observed measurements were reliable by computing Cronbach's alpha for all constructs. As the survey was composed out of 51 items for the 17 constructs, each construct's three items were compared. According to common understanding, an alpha value of 0.70 or higher is regarded as a good fit, indicating that the items are having relatively high internal consistency (Hair, et al., 1998). Where necessary, one of the three items verifying each construct was erased to increase the internal reliability, however, due to restrictions in SPSS Amos, it is not possible to reduce the number of items below two. After the pre-test we proceeded with 43 questions (see Figure H). Most of the alpha values were in the range between 0.717 and 0.931, which indicates good reliability. Only the two constructs Social Influence and Facilitating Conditions were not above 0.70 and are thus seen as weaker consistent. We nevertheless decided to continue with both in our further analysis, as they were an integral part of our theoretical adapted UTAUT model. We refer to the weaker reliability later in Chapter 6.

Figure H: Cronbach's alpha for the Items of each Construct

Construct	**Initial Items***	**Cronbach's alpha**	**Erased Item**	**Retained Items**	**Cronbach's alpha afterwards**
IS *(Intention to Share)*	3	**0.865**		3	
PE *(Performance Expectancy)*	3	**0.732**	PE_1	2	**0.765**
EE *(Effort Expectancy)*	3	**0.625**	EE_3	2	**0.717**
SI *(Social Influence)*	3	**0.613**	SI_3	2	**0.638**
FC *(Facilitating Conditions)*	3	**0.464**	FC_2	2	**0.473**
TaPr *(Tailored Product)*	3	**0.817**		3	
InSa *(Increased Safety)*	3	**0.786**	InSa_3	2	**0.821**
EnCo *(Enhanced Convenience)*	3	**0.869**		3	

EcBe *(Economic Benefit)*	3	**0.833**		3	
SeEf *(Self-Efficacy)*	3	**0.624**	SeEf_3	2	**0.739**
Af *(Affect)*	3	**0.828**		3	
TrBu *(Trust in Businesses)*	3	**0.869**		3	
InTr *(Institutional Trust)*	3	**0.931**		3	
SoRe *(Social Referral)*	3	**0.874**		3	
Re *(Reciprocity)*	3	**0.824**	Re_1	2	**0.852**
Tr *(Transparency)*	3	**0.623**	Tr_2	2	**0.727**
Gu *(Guidance)*	3	**0.792**		3	
Total	**51**			**43**	

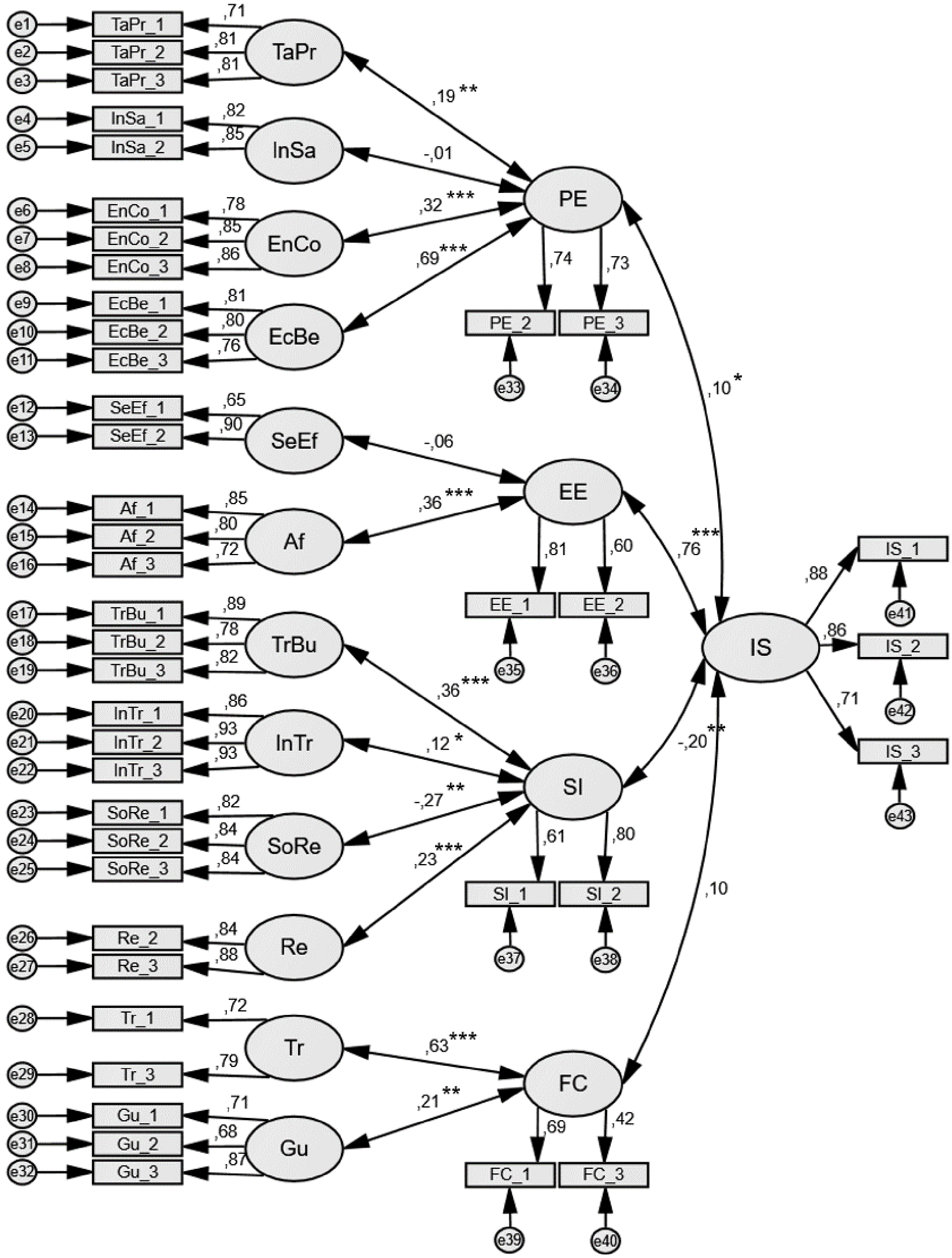
*see Exhibit E for an overview of all statements belonging to the items

4.2.2 Confirmatory Factor Analysis

To recall, it is recommended to test the model via a CFA to confirm a data-to-model fit. As a general rule of thumb, some researchers suggest having a sample size of *n* that is larger than five times the number of free parameters in order to effectively conduct a CFA and SEM study (Loehlin, 1998; Marsh, et al., 1988). As our model had 43 measurements questions, we needed more than 215 responses, which we successfully achieved and almost doubled. Before starting the CFA, we had to check if there is any missing data, as for example incomplete responses. Using QualtricsXM we were sure to collect complete questionnaires, as this software only recorded fully completed survey by our setting.

We started the CFA by drawing the structural model and attaching the recorded questionnaire data to the items. Our initial structural model (see Figure I) was displaying the drawn relationships from our theoretically derived adapted UTAUT model described in Chapter 2.3. The software SPSS Amos computed the regression weights and attached them to each path, and the program indicated the significance by using asterisks. The numbers linked to the one-directional arrows from any construct to the related items are called loadings. They show how well the items are able to describe the respective construct, such as Tailored Product or Increased Safety. According to Tabachnick and Fidell (2007), loadings over 0.71 are excellent, over 0.63 are very good, over 0.55 are good and over 0.45 are fair. As seen in Figure I, the majority of the loadings were above 0.71. This is in line with the internal consistency reliability assessment via Cronbach's alpha in the previous chapter.

Figure I: Initial Proposed Structural Model



*** p<0.001; ** p<0.01; * p<0.05

THREE CATEGORIES OF MODEL FIT

In the next step, we analyzed the actual fit of the model. There is a broad range of commonly accepted model-fit-measure indices which serve as indicators how well the model fits the input data. They are defined in the three categories – the absolute fit, the incremental fit, and the parsimony fit indices – and we have chosen four of them in total to indicate our model's fit based on recommendations from Hooper et al. (2008). Part of the first category, *Relative Chi-Square* (RCS) and *Root Mean Squared Error of Approximation* (RMSEA) are evaluating on a fundamental level how well the model fits the data compared to no model at all (Jöreskog & Sörbom, 1993). Second, the *Comparative Fit Inde*x (CFI) is a more specific measurement that compares the Chi-squares to a baseline model (McDonald & Ho, 2002) and takes sample size into account (Byrne, 1998). It requires all latent variables to be uncorrelated (Hooper, et al., 2008). The third category includes the *Akaike Information Criterion* (AIC) which considers the model complexity and is used to compare iterations with the same underlying input data. In contrast to the other measures there is no specific threshold value, but generally, a lower value is preferable when comparing two or more models (Akaike, 1974). A summary of the computed indices for this proposed model and of the further iterations together with the recommended threshold values can be found in Figure J.

The model-fit indices for this initial proposed model were below the commonly agreed thresholds with RCS 4.192, RMSEA 0.090, CFI 0.726, AIC 3,742. Thus, our initial structural model did not seem to fit the data very well according to these criteria. SPSS Amos is able to propose well-fitting relationships by calculating the impact of adding those new paths in regard to the model-fit indices. Therefore, the software suggested iterations of the model with a few new relationships that would increase the model fit. We evaluated all of these iteration suggestions based on the implication that this would bring to our model before changing the nature of it. Our decision rule was to only add a path if the link is logically supported by our theoretical framework and where the covariance was significant at minimum $p<0.05$ level.

Figure J: Model-Fit Measure Indices

	Recommended level of fit	Initial structural model	1st Iteration	2nd Iteration	Erasing Af & Tr
Absolute fit indices					
RCS (Relative Chi-Square)	2-5, <5 (Bentler, 1990)	4.192	3.166	3.162	2.579
RMSEA (Root Mean Square Error of Approximation)	<0.08 (Teo, 2012)	0.090	0.075	0.074	0.064
Incremental fit index					
CFI (Comparative Fit Index)	>0.90 (Browne & Cudeck, 1992)	0.726	0.817	0.817	0.904
Parsimonious fit index					
AIC (Akaike Information Criterion)	Smaller value better fit (model specific)	3,742	2,862	1,892	1,313

First iteration – Adding logical paths

At the first iterative step, we found that there are many relationships among the constructs of the construct layer, mainly inside each of the four blocks. Therefore, we decided to add those paths, as this seemed logical to us, because the constructs were derived from the same theoretical foundation and they were referring to the same group, which indicated that there should be a strong correlation. Based on these changes we derived our first iteration, which accounted for many covariances between the constructs of the construct layer. The model-fit indices for this new structural model were RCS 3.166, RMSEA 0.075, CFI 0.817, AIC 2,862. Compared to the initial model, RCS was much better and RMSEA was also fulfilled. In relative terms, the model improved a lot, as the AIC was by a fourth smaller. Only the CFI was still not at a satisfactory level, hence not above 0.9.

Second iteration – Erasing non-significant paths

As no more suggestions for adding paths in the first iteration were found to make logical sense, we did the second iteration, in which we evaluated the option to erase non-significant paths at the $p<0.05$ level. We found that the paths Facilitating Conditions towards Intention to Share, Reciprocity towards Social Influence and Self-Efficacy towards Effort Expectancy were not significant. After erasing those paths, we calculated the model-fit indices again, which were RCS 3.162, RMSEA 0.074, CFI 0.817, AIC 1,892. While the RCS, RMSEA, and CFI nearly did not change, the AIC improved by a third. This is logical since unnecessary complexity was erased, which is in line with the concept of AIC (Akaike, 1974). When comparing the regression weights to the first iterations, they only changed slightly, which was expected. Thus, the model did not need those three non-significant paths to explain the data. Our second iteration displayed a model that accounts for all relevant paths which would be logically supported by the theory. Nevertheless, the CFI index was not showing a good model fit.

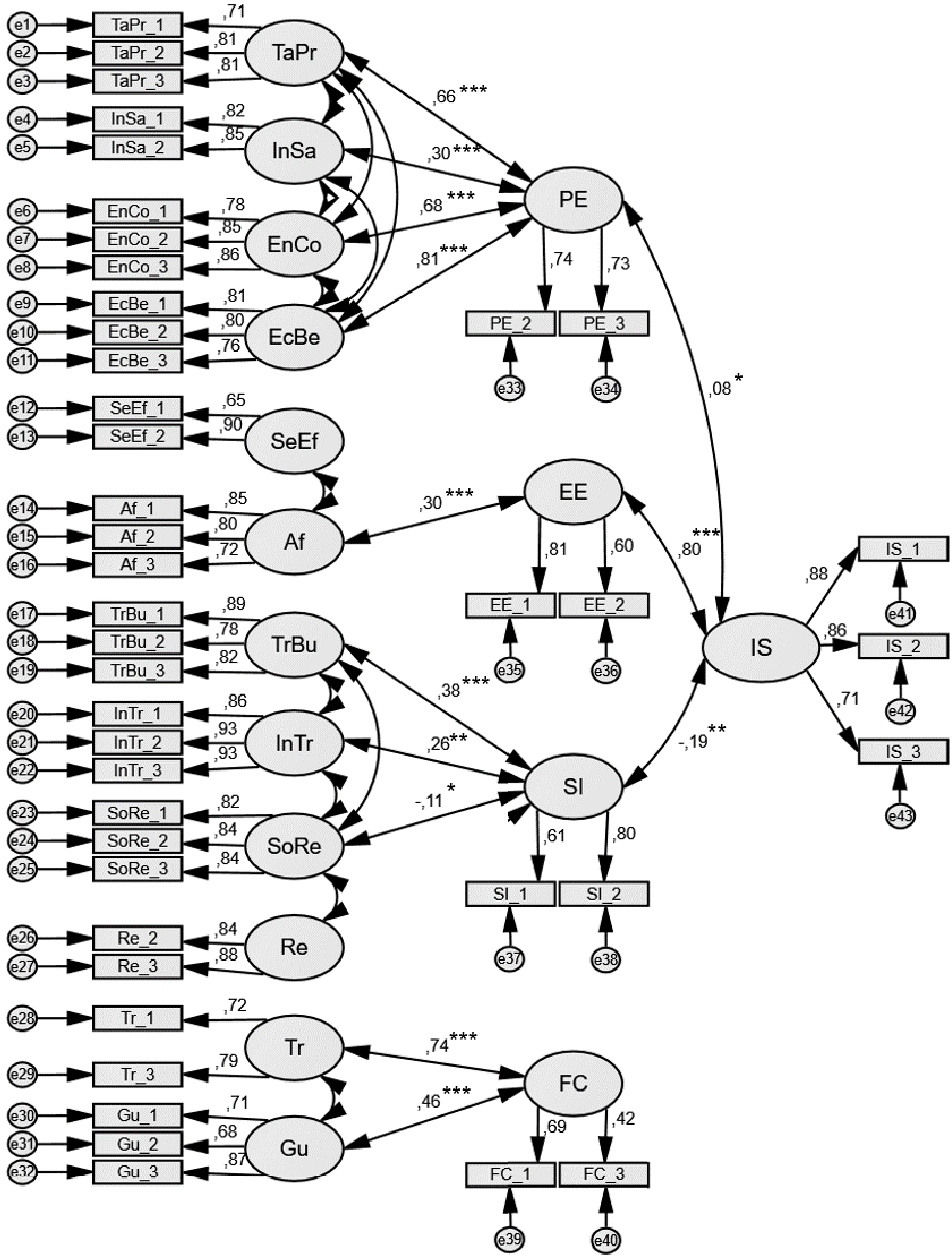
Option of erasing the constructs Affect and Transparency

After the first and second iteration, we evaluated the option of erasing the constructs Affect and Transparency, because both strongly correlated to Intention to Share, even if they are not directly linked via paths, which could explain the lower model fit. We tested an adoption of the structural model, where Affect and Transparency were erased. In this model, all indices were fulfilling the threshold with RCS 2.579, RMSEA 0.064, CFI 0.904, AIC 1,313. Especially the CFI and the ACI would be strongly affected, as it was a requirement of the CFI to not have strongly correlated latent variables and as the model would lose complexity for the AIC. Nevertheless, we decided to not iterate our model according to this option as we did not want to overfit it to get an extremely well-fitting model. Thereby, we would have to reject a major part of the initial theoretical model and would have the chance of a type I error, meaning to reject a hypothesis of an actually acceptable model. Our intention of this project was to find various factors influencing the Intention to Share, which is why we valued a slightly non-fitting model higher than consciously rejecting many hypotheses due to thresholds. In academia, there is also an ongoing debate about the usefulness of strict cutoff values, as it increases type I errors, and Barret (2007) for example claims that some researchers are in favor of even abandoning those fit indices. Summarized, this discussion suggests that we could accept our second iterated model, as the majority of fit indices is in a satisfactory range showing a good model fit.

4.2.3 Re-Specified Model

As argued in the previous part, we decided to take our second iteration as the best-fitting structural model for the needs of our research project to advance further with it. The detailed structural model can be seen in Figure K and all regression weights and significances are displayed accordingly.

Figure K: Re-Specified Structural Model after CFA



*** p<0.001; ** p<0.01; * p<0.05

4.3 Presentation of Findings

Based on the SEM we computed the regression weights for each of the paths and concluded the implications for our hypotheses (see Figure L). In the following Chapters 4.3.1 until 4.3.4, we present the findings according to the four UTAUT factors. We set our threshold of hypothesis rejection to p<0.05.

Figure L: Hypothesis Testing Results

Hypothesis		Path	Regression	Outcome
H1	*Performance Expectancy has a significant positive effect on Intention to Share.*	PE → IS	0.079*	Not rejected
H2	*Effort Expectancy has a significant positive effect on Intention to Share.*	EE → IS	0.799***	Not rejected
H3	*Social Influence has a significant positive effect on Intention to Share.*	SI → IS	-0.186**	Rejected
H4	*Facilitating Conditions has a significant positive effect on Intention to Share.*	FC → IS	Not significant	Rejected
H5	*Trust in Businesses has a significant positive effect on Social Influence.*	TrBu → SI	0.379***	Not rejected
H6	*Institutional Trust has a significant positive effect on Social Influence.*	InTr → SI	0.264***	Not rejected
H7	*Social Referral has a significant positive effect on Social Influence.*	SoRe → SI	-0.108*	Rejected
H8	*Reciprocity has a significant positive effect on Social Influence.*	Re → SI	Not significant	Rejected
H9	*Self-Efficacy has a significant positive effect on Effort Expectancy.*	SeEf → EE	Not significant	Rejected
H10	*Affect has a significant positive effect on Effort Expectancy.*	Af → EE	0.304***	Not rejected
H11	*Tailored Product has a significant positive effect on Performance Expectancy.*	TaPr → PE	0.660***	Not rejected
H12	*Increased Safety has a significant positive effect on Performance Expectancy.*	InSa → PE	0.299***	Not rejected
H13	*Enhanced Convenience has a significant positive effect on Performance Expectancy.*	EnCo → PE	0.683***	Not rejected
H14	*Economic Benefit has a significant positive effect on Performance Expectancy.*	EcBe → PE	0.809***	Not rejected
H15	*Transparency has a significant positive effect on Facilitating Conditions.*	Tr → FC	0.738***	Not rejected
H16	*Guidance has a significant positive effect on Facilitating Conditions.*	Gu → FC	0.457***	Not rejected

*** p<0.001; ** p<0.01; * p<0.05

4.3.1 Performance Expectancy

The latent variable Performance Expectancy was slightly positively related to Intention to Share with a regression weight of 0.079 and a low statistical significance at $p<0.05$ level. Thus, there was a positive effect measurable, but the strength of it was comparably low. The four constructs linked to Performance Expectancy all showed high significance. Tailored Product had a high positive effect on Performance Expectancy (0.660), which was significant at $p<0.001$ level. Similarly, Enhanced Convenience was related to Performance Expectancy with 0.683, and Economic Benefits was related to Performance Expectancy with 0.809. The path of Increased Safety towards Performance Expectancy was significant but lower with 0.299 compared to the previously mentioned three constructs. Thus, all five hypotheses (H1, H11, H12, H13, H14) were not rejected.

4.3.2 Effort Expectancy

We found that the path Effort Expectancy towards Intention to Share was the one that had the strongest effect within the latent layer 1 to latent layer 2 paths, having a regression weight of 0.799 and high significance. Hence, out of the four UTAUT factors, Effort Expectancy seemed to have the greatest effect on Intention to Share. The two paths of the constructs Self-Efficacy and Affect towards Effort Expectancy showed different results. Self-Efficacy did not have a significant effect on Effort Expectancy, whereas Affect had a significant effect on it with a regression weight of 0.304. Furthermore, we figured out that Affect was directly highly correlated with Intention to Share, which serves as an interesting point that we discuss later in Chapter 6.1.2. To summarize, we did not reject H2 and H10, but we rejected H9.

4.3.3 Social Influence

At the Social Influence block, we found very interesting results. The path between Social Influence and Intention to Share was rejected because it was significant at the $p<0.01$ level but negative and thus had no positive regression, which was unexpected based on our theory. Similarly, Social Referral had a negative regression weight at the path towards Social Influence and therefore, the hypothesis had to be rejected. Next, we could not find a significant effect of Reciprocity on Social Influence either, which is why this path was also rejected. Furthermore, the paths Trust in Business towards Social Influence and Institutional Trust towards Social Influence had both significant regression weights of 0.379 and 0.264, respectively. Concluding, we rejected the hypotheses H3, H7, and H8 and did not reject H5 and H6.

4.3.4 Facilitating Conditions

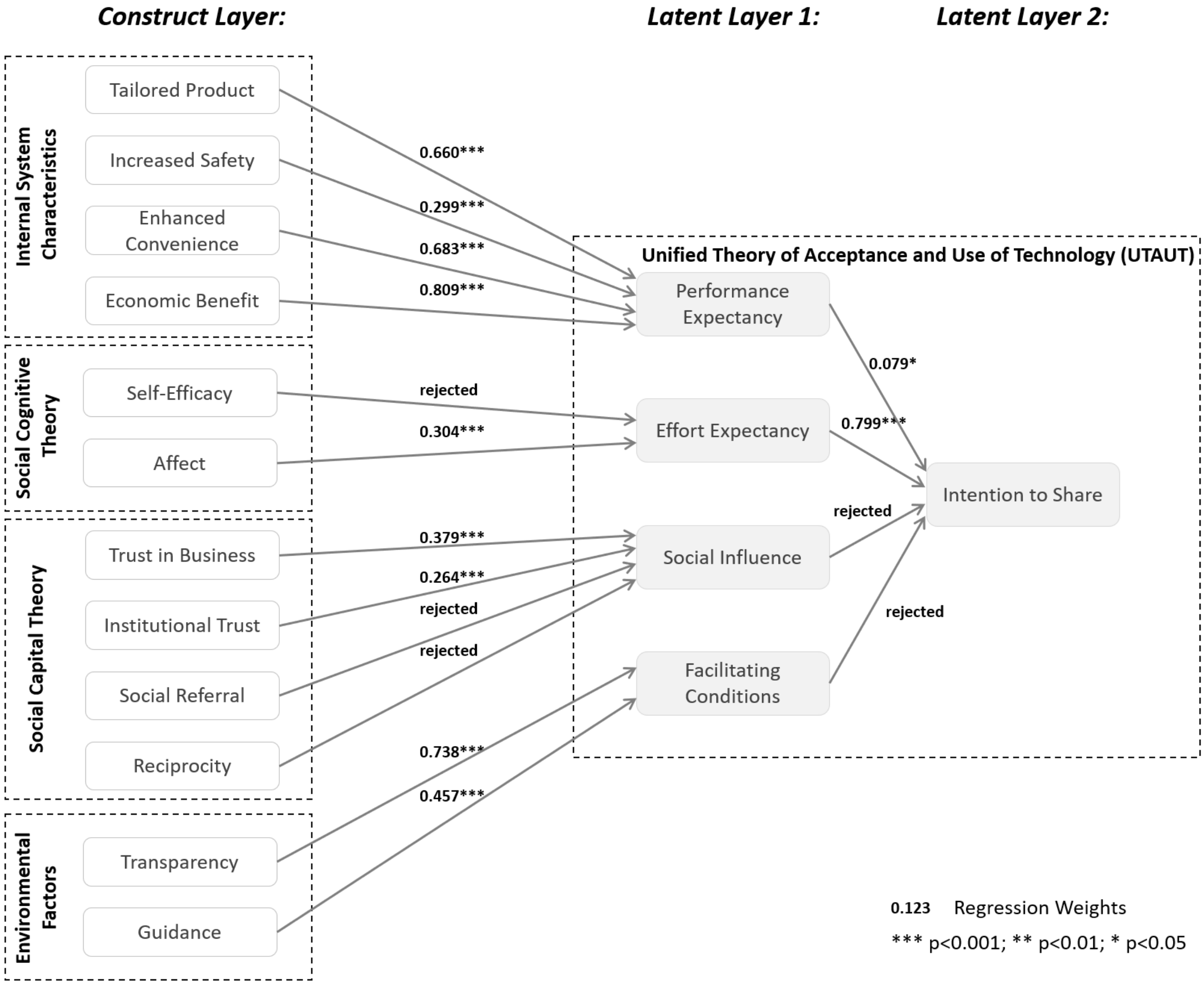
There was no significant regression between Facilitating Conditions and Intention to Share, which led us to reject the respective hypothesis H4. Apart from that, the two constructs Transparency and Guidance both had a significant positive effect on Facilitating Conditions with high regression weights of 0.738 and 0.457. Given its higher regression weight, Transparency was a more valued factor compared to Guidance. Additionally, we found out during the CFA that Transparency is strongly correlated directly to Intention to Share, which we refer to in the discussion. Summarizing, the two hypotheses H15 and H16 were not rejected and H4 was rejected.

4.4 Outcomes of the Main Study

In the process of SEM, we iteratively improved our structural modeling to finally derive a satisfactory depiction of the adapted UTAUT model. Based on the given input data from the questionnaire we computed the regression weights for all paths. We had to reject three hypotheses (H4, H8, H9) as the paths were not significant. Furthermore, as the regression weights were significant but negative, we additionally rejected two hypotheses (H3, H7). This left eleven hypotheses, which we were not able to reject. After the SEM we derived the final model (see Figure M). We originally included the interrelations between each construct on the construct layer (i.e. Tailored Product towards Enhanced Convenience) during the SEM analysis, though we did not include them in Figure M anymore, because it would overcomplicate the illustration and would not give any additional insights.

Taken the outcomes of the main study to a more interpretative level, this does not mean that we found any evidence for a causal relationship of the eleven paths, where we did not reject the hypotheses. SEM is merely a tool to discover relationships and estimate the correlations between variables. It is not there to prove anything true, but rather to not display clear falsified statements (Cliff, 1983). Since this quantitative method did not enlighten any causal explanation, we analyze these relationships further in Chapter 5.

Figure M: Regression Weights Applied to the Adapted UTAUT



5. Analysis Follow-Up Study: Focus Groups

In this section, we describe the outcomes of the qualitative follow-up study and relate them to our theoretical model.

5.1 Overview of the Follow-Up Study

We conducted three focus groups from April 15 to April 17, 2019, with 5 to 8 participants in each group. During that time, all participants lived in Stockholm, Sweden and we chose these individuals based on their questionnaire outcome. All participants signed a consent sheet, giving us the permission to use and publish their input in our research and to access their questionnaire responses when analyzing the discussions. Afterwards, we anonymized all names of the 21 participants (see Exhibit G).

5.2 Presentation of Findings

Based on the focus group Overview Grid mentioned in Chapter 3.3.4, we aggregated the mentioned topics that occurred during the focus groups and framed them into the categories based on the adapted UTAUT model. It is indicated how much overall support was given for those constructs.

5.2.1 Performance Expectancy

In general, when asked about what would influence the participants to share their data, the majority of them agreed that they would share if they get something in return as a main reason and thereby the focus groups supported the overall concept of Performance Expectancy. Next, we list further details of what the participants expected to get in return as well as their underlying motivations.

TAILORED PRODUCT – MUCH SUPPORT

Many participants mentioned that they would share their data if they receive a more personalized service in the form of an optimized route planning. One participant stated, “I could also imagine that I appreciate sort of suggestions of >Here, on the ways to school you could also go this way<” (P1b). Some participants explained that a more tailored product would increase their convenience, “I would expect that if I share my data, my routes are going to be more personalized in some way. So, my life will get easier. Maybe not immediately, but more in the long term“(P3c). Taking the idea of personalization to a more general level, two participants mentioned that sharing their data and track history with Spotify and Netflix yields benefits because these services would “give recommendations that you wouldn’t have gotten if you wouldn’t share your data” (P2d).

INCREASED SAFETY – NOT MENTIONED AT ALL

No participant in any of the three focus groups mentioned anything that would be related to the category of Increased Safety.

Enhanced Convenience – Much support

Related to the category of Tailored Product, many participants directly mentioned convenience as something they expect to get in return for data-sharing. Some participants expected increased convenience for themselves such as, "I would share it, if it improves the quality of the service that I receive" (P3e), whereas others emphasized the societal perspective, stating, "I place a lot of value in collecting data for beneficial purposes, particularly for collectively beneficial projects" (P1a). We saw that many arguments of participants throughout all focus groups associated the ideas of a more tailored product with convenience, thus those seem to be interrelated.

Economic Benefit – Much support

The majority of participants agreed to expect some more tangible benefits, as exemplified by the quote, "I would also share if there is some monetary reward for me since I am a student with little money" (P3d). Concerning the amount of the economic benefit, it seems that small amounts already matter as one participant concludes, "a little is just enough" (P2g). Thus, it seems that the simple act of getting a return is enough. One participant mentioned that a greater financial incentive would actually have the opposite effect by stating "And if I would get 100 SEK for my data, I would say >Wow, what would they do with that?<" (P2d). Hence, there seems to be a lower and an upper limit where economic benefits would enhance the consumer's intention to share. However, a few candidates rejected the idea of economic returns, like one participant stated, "Moneywise, I don't think a company can pay that many incentives. So, if you would get cash, it might be cents and then I don't really care about these cents" (P2f).

5.2.2 Effort Expectancy

Many of the focus group participants talked about the hindrances they perceive towards data-sharing, thus there was a lot of support for the concept of Effort Expectancy. Only a few participants felt little hindrances to share their data at all, and one of them stated, "For me, it is very hard to find something that hinders me because I feel like I am already sharing so much data" (P3c). Therefore, it seems that the personal attitude and individual behavior patterns are important in this category, as we expected according to our theoretical framework. In the following, we analyze the intentions based on the cognitive attributes Self-Efficacy and Affect.

Self-Efficacy – Little support

A small amount of the participants mentioned that their feeling of confidence affects their intention to share data, hence there was little support for the hypothesis of Self-Efficacy. Only a few participants felt confident when deciding to share their data and demonstrated a strong degree of self-efficacy, as mentioned by one candidate, "I think it is positive that I can control it, and that I feel I can make that decision" (P1b). But as many of the participants which had lower confidence would have shared their data anyways and saw no big effort for it, we could not identify any strong causal link between Self-Efficacy and Effort Expectancy.

AFFECT – MUCH SUPPORT

The construct Affect was supported a lot, because most of the participants stated that they have a certain feeling and opinion towards data-sharing, either positive or negative, which influences their data-sharing decision. The majority of participants had a strong negative feeling about data-sharing, for example, one person stated that "there is already a lot of information that is shared out there and I don't like that" (P2e). Nearly all of those who initially stated that they would not share their data, had negative feelings about it, while some who indicated they would also mention a bad gut feeling. On the opposite side, a few participants had a more indifferent opinion towards sharing their data and one person stated, "I probably have given away so much more information, so that feels just like some small detail" (P2h). Overall the participants' opinion towards data-sharing was aligned with their initial vote whether to share their data or not.

5.2.3 Social Influence

Building upon the theoretical framework, we expected individuals to make the data-sharing decision based on their social surrounding. The concept Social Influence has been discussed in the focus groups by asking questions about trust, social referral, and reciprocity.

TRUST IN BUSINESS – MUCH SUPPORT

The majority of participants agreed that it highly depends on the trustworthiness of a company if they would share their data with that company. We asked how a company could establish trust to understand the participants' reasoning. One important factor was the company reputation, as one participant stated, "I think it really depends on the company reputation. If I know the company, people who work there and if it is a reliable company" (P3a). Another factor of trust was the awareness about the company's scandals, as one person mentioned a specific example, "I have been very closely following the Cambridge Analytica case, [I am] kind of a little bit more skeptical over the past few years about data usage and whether I can trust an organization" (P1a). A third factor determining the trustworthiness was the mission of the company, as hinted by one participant, "And I know about the history of the company, and I know about the mission of the company. They are not there to exploit my data. [...] I would trust them more" (P3c). We found that trust was a type of pre-requisite factor for the majority of participants because everyone agreed that trust in a business is a necessity to allow the company to use their data. Furthermore, in two of the three focus groups, the discussions about trust were directly accompanied by discussions about transparency. Acting transparently seems to be connected to being perceived trustworthy, which we refer to in the discussion later.

Institutional Trust – Much support

When shifting from trust in businesses to trust towards a government, one pattern among participants was to verify which government we are talking about before determining whether they would trust the government as summarized by "I would have more confidence in the Swedish government and opposed to other countries, it might be different" (P3b). Many participants referred to the Swedish government, but a few referred to their home country government resulting in different answers, for instance one participant explained "if I am in Russia, I would never ever share my data, because there is no trust in the government" (P3c) and another one said "I would trust the Swedish government, not sure if I would trust the American government" (P1e). Similar to Trust in Businesses, reputation, scandals or the general mission of a government seem to influence the trustworthiness.

One person explained the motivation for her trust in the Swedish government, "because for the government there is nothing really like a direct profit and the benefits are sort of shared in the end" (P1b). However, some participants distrusted the government for various reasons, as stated by "I would never share my data with the government because I don't want them to see what I am doing during my whole life" (P3f) or "I wouldn't share data with the public authority, because they can use it also for other purposes and they have a lot of power" (P3h). There was no clear majority for trusting or distrusting the government, but in comparison towards trust in businesses, the tendency was slightly more negative. Based on their answers what affects their data-sharing opinion, it seems that especially skeptical individuals are distrusting the government in general.

Social Referral – Contradictory to original hypothesis

The majority of the participants stated that a friend's or family member's referral would not affect them in their data-sharing decision. One participant answered the question of whether they share their data after a referral with "No, I would ask them >why?<. Because I think I wouldn't just do it because they do it, though I know that I am very much influenced by my surroundings and especially by my family and my closest friends" (P1b). Another participant highlighted that they still need more knowledge about the specific data-sharing context in addition to the referral, "If you just say >Share your data<, I would need to know about the benefits they got" (P3c) and one individual even doubted the credibility of the referral, "I say no, because they do not have any knowledge of how this data is used" (P2e). Thus, other factors are seemingly higher valued for decision-making than the social surrounding.

A few individuals even stated they would become more cautious when friends refer them to share their data. One participant gave the following reasoning, "I like the idea of owning my own thoughts and my own decision" (P1c). This was in line with what another candidate said during a different focus group, "So if a friend would say >Share your data< without mentioning any benefits, I would may even be less likely to share my data because I now start to think about it before I use the service" (P3d). People seem to be more skeptical and question the motives of the referrer instead of simply accepting it and following their advice. This was contradicting to our theoretically derived hypothesis in Chapter 2.3.2.

RECIPROCITY – LITTLE SUPPORT

We asked the participants whether they feel more obliged to share their data knowing that the majority of the society is already doing so to improve the service in any beneficial way for everyone. There was little support that reciprocity would be important to the participants. One person explained that one's attitude on data-sharing does not depend on other people by stating "No, I wouldn't feel like I would share it because of the pressure from people around me, but I would feel like I want to share because I want to share." (P1d). Another participant added further reasoning why one does not feel obliged to share data saying "There is less pressure [...]. Not sharing data is not something that is bad for people" (P2f). Finally, a third participant reasoned that Reciprocity is not so relevant because "no one knows which button I pressed unless my app turns red and yours turns green" (P3c).

However, a few of the participants felt more willing to share their data based on their surroundings. One person felt "a bit of a moral obligation because I really don't want to be a freerider on something" (P1b) and another one concluded "If everyone shares their data, [...] this could be cool to improve society, and in my opinion, maybe change some minds in sharing their data, too" (P2a). Those two participants were identified as highly altruistic in our main study. Hence, we assume that reciprocity plays an important role for altruistic individuals but seems to be less important for individuals in general.

5.2.4 Facilitating Conditions

When discussing the concept of Facilitating Conditions, we directly asked the participants how the system should be set up when they are asked to share their data.

TRANSPARENCY – MUCH SUPPORT

Almost every participant highlighted the importance of transparency in the data-sharing context. Two participants immediately mentioned transparency when answering how the data-sharing system should be designed by requesting, "just full transparency, what is going to happen with the data in a very simple manner" (P3e) and by claiming, "I think it is transparency, by making it simple. Make it as stupid to understand as possible" (P3a). The participants highlighted the importance of transparency with a different emphasis. First, many participants wanted transparent information about the purpose of data-sharing. They mentioned that "there should be a section that explains what the data is being used for and what company is using the data" (P1d) and, "what would be the consequences of me choosing no or yes" (P1d). Second, according to the focus group participants, the terms and conditions should be transparent.

There was some advocating for a short and concise overview of terms and conditions, whereas others favored a more elaborated manner. One participant challenged the focus group, “But would you rather have a long terms and conditions list where you just click on the green button on the right side rather than short and concise terms and conditions that can be read in less than a minute?” (P1a). In general, most participants favored a more concise, yet transparent, overview of the terms and conditions. Third, many participants required transparency of what happens with the shared data afterwards, especially the access to collected data about themselves. One candidate requested a “very easy access to your own data, that would be something I would really much appreciate” (P1a). Moreover, no candidate disagreed with having transparency as a positive influence when deciding to share his or her data, which strengthens the importance of Transparency.

GUIDANCE – LITTLE SUPPORT

In the focus groups, we asked the participants whether they would like to have support when deciding to share their data. Almost every participant denied that offer, thus the importance of guidance seemed to have little support from the participants. One of the reasons that came up several times was that it would make the data-sharing decision overly complicated and one participant explained, “No, I mean there are millions of apps that work without guidance, so why does this one need special guidance? [...] Yeah it signals the customer that we are either weird or incompetent” (P2e). Even though the majority rejected the offer of guidance, a small group of participants advocated in favor of it as exemplified by the following statement, ”So if anyone has doubts about something then they can easily like write questions and get answers and that could be appreciated by people, so I feel like yes, there should be a functions where you can ask questions” (P1d). Finally, a few participants were indifferent regarding a guidance option by explaining “I would probably be too lazy to write and wait until someone answers” (P3f). As the majority was not expressing the need of being guided, it seems as if this aspect is only a subordinate environmental factor to influence the intention to share data.

5.3 Additional Input

During the interviews, some participants mentioned aspects, which we did not consider yet and therefore serves as additional input regarding data-sharing in the mobility contexts. First, we found that we could extend the Performance Expectancy by one more category, called *Fairness*. Especially in the last focus groups, individuals stated that they expect some type of benefit in return for sharing their data because the third party collecting their data benefits from the shared data as well. Hence due to fairness reasons, some of the third-party benefits should be forwarded to the individual according to the participants. Second, we could add the idea of *Selective Ignorance* as another cognitive characteristic to the Effort Expectancy. As many individuals said, they would share their data without any hesitation even though they know they should probably think more about that decision. Exemplary, many participants stated that they ignore the urge of reading the terms and conditions carefully and rather choose to be deliberately ignorant by just clicking the accept-button. Finally, many participants stated new personal characteristics that influence their opinion towards data-sharing. We detected four new moderating variables, and these are the profession, the interest in innovation, the interest in technology, and the cultural background.

5.4 Outcomes of the Follow-Up Study

In general, many of the focus group participants explained under which circumstances they would be more or less willing to share their data and shared their thoughts and reasoning with us. Summarized, we made the following observations: First, benefits of sharing data, i.e. the Performance Expectancy, were mentioned a lot, and especially personalization, convenience, and economic rewards were frequently mentioned. Second, potential hindrances and the effort of sharing were important, and the majority of participants expressed a clear attitude towards data-sharing in general, hence Affect seemed to play an important role. Third, Social Influences and its four mentioned factors showed different results, and there were many discussions on trust in businesses and the institutional trust, which both impact the data-sharing attitude. However, a personal referral did not seem to have a positive influence on the Intention to Share, as many participants explained that they would become even more cautious and reluctant to share their data if their friends only refer them to do so without mentioning specific benefits of doing so. Finally, many of the participants had a strong preference for how the system should be set up, thus the inherent Facilitating Conditions of the data-sharing process were confirmed to have a great impact on the intention to share data. Whereas almost all of the participants requested full transparency, most of the participants rejected the option to have some form of guidance.

6. Discussion

In the discussion, we reflect on the research findings and analyze them regarding the expectations based on our theoretical framework in order to present an updated research model.

6.1 Comparison Between Main Study and Follow-Up Study

We compared the outcomes of the main study with the outcomes of the follow-up study in Figure N. In this figure, we listed our 16 hypotheses and the outcomes from the main and the follow-up study. Finally, we evaluated how much overall evidence we found for each respective hypothesis, as summarized in the last column. In the next parts, we discuss Performance Expectancy, Effort Expectancy, Social Influence, and Facilitating Conditions in detail.

Figure N: Comparison of Findings from Main and Follow-Up Study

Hypothesis		**Outcome Main Study**	**Outcome Follow-up Study**	**Evidence strength for hypothesis**
H1	*Performance Expectancy has a significant positive effect on Intention to Share.*	Not rejected (0.079*)	Much support for hypothesis	**strong**
H2	*Effort Expectancy has a significant positive effect on Intention to Share.*	Not rejected (0.799***)	Much support for hypothesis	**strong**
H3	*Social Influence has a significant positive effect on Intention to Share.*	Rejected (negative effect)	Some support for hypothesis	**weak**
H4	*Facilitating Conditions has a significant positive effect on Intention to Share.*	Rejected (not significant)	Much support for hypothesis	**medium**
H5	*Trust in Businesses has a significant positive effect on Social Influence.*	Not rejected (0.379***)	Much support for hypothesis	**strong**
H6	*Institutional Trust has a significant positive effect on Social Influence.*	Not rejected (0.264***)	Much support for hypothesis	**strong**
H7	*Social Referral has a significant positive effect on Social Influence.*	Rejected (negative effect)	Contradictory to hypothesis	**weak**
H8	*Reciprocity has a significant positive effect on Social Influence.*	Rejected (not significant)	Little support for hypothesis	**weak**
H9	*Self-Efficacy has a significant positive effect on Effort Expectancy.*	Rejected (not significant)	Little support for hypothesis	**weak**
H10	*Affect has a significant positive effect on Effort Expectancy.*	Not rejected (0.304***)	Much support for hypothesis	**strong**

H11	*Tailored Product has a significant positive effect on Performance Expectancy.*	Not rejected (0.660***)	Much support for hypothesis	**strong**
H12	*Increased Safety has a significant positive effect on Performance Expectancy.*	Not rejected (0.299***)	Not mentioned at all as a reason	**medium**
H13	*Enhanced Convenience has a significant positive effect on Performance Expectancy.*	Not rejected (0.683***)	Much support for hypothesis	**strong**
H14	*Economic Benefit has a significant positive effect on Performance Expectancy.*	Not rejected (0.809***)	Much support for hypothesis	**strong**
H15	*Transparency has a significant positive effect on Facilitating Conditions.*	Not rejected (0.738***)	Much support for hypothesis	**strong**
H16	*Guidance has a significant positive effect on Facilitating Conditions.*	Not rejected (0.457***)	Little support for hypothesis	**medium**

*** p<0.001; ** p<0.01; * p<0.05

6.1.1 Performance Expectancy

The UTAUT model has its foundation in the TAM by Davis (1989), who argues that a system should have specific useful characteristics in order to become adapted. Davis emphasizes the importance of the Perceived Usefulness, i.e. "the degree to which a person believes that using a particular system would enhance his or her job performance" (1989). We reflected this thought of Performance Expectancy in the hypothesis H1 and in our research we found strong evidence for the Performance Expectancy to have a positive effect on the Intention to Share. This is because the questionnaire outcome for testing this hypothesis was statistically significant and we found much support from the participants in the focus groups. In the following, we discuss each hypothesis belonging to Performance Expectancy.

TAILORED PRODUCT – STRONG EVIDENCE FOR H11

We assumed that a Tailored Product has a positive effect on Performance Expectancy. We found strong evidence for this hypothesis since this regression was of statistical significance in the questionnaire and many focus group participants wished to receive a more personalized product when sharing their data, as exemplified in the desire to get suggested routes.

INCREASED SAFETY – MEDIUM EVIDENCE FOR H12

There was medium evidence for Increased Safety to be a relevant factor of the Performance Expectancy. Even though this hypothesis was not statistically rejected, not a single participant has proactively mentioned safety when talking about data-sharing in the mobility context. Even though during the pre-study Increased Safety was mentioned by company representatives as one of the advantages of data analytics in the transportation sector, it seems that the majority of consumers are not aware or interested in that benefit of sharing data.

Enhanced Convenience – Strong Evidence for H13

We hypothesized that Enhanced Convenience has a positive effect on Performance Expectancy. Statistically, this hypothesis has been significant and in the focus groups we found much support, leading to strong evidence overall. As seen in the focus groups, some participants perceived a tailored product to cause enhanced convenience, so there seems to be a linkage between these two constructs. This specific example of construct interrelation reinforced our decision to include paths between various constructs in our model, as we already observed at our first iteration during the CFA.

Economic Benefit – Strong Evidence for H14

Finally, we argued that Economic Benefit impacts Performance Expectancy and this hypothesis had strong support, both from the main study and the follow-up study. It is notable that economic incentives had the strongest regression among the four internal system characteristics in the quantitative study, thus highlighting its particular importance. Furthermore, during the focus groups we observed, that the consumers do not emphasize the amount of economic incentives, however they valued the existence of a reward and the visibility of such. These findings are especially interesting because businesses can integrate economic compensations towards the customer when designing their service, as elaborated on in the managerial implications.

6.1.2 Effort Expectancy

Attempting to understand an individual's reasoning to accept new technologies, we used the variable Perceived Ease of Use in the TAM model, which is reflected in the concept of Effort Expectancy of the UTAUT (Venkatesh, et al., 2003). The original definition is "the degree to which a person believes that using a particular system would be free of effort" (Davis, 1989) and we tested Effort Expectancy via H2. Given the responses from the questionnaire, we conclude that the path is statistically significant and has the highest effect towards Intention to Share among all four UTAUT factors. In the focus groups, we also found much support. Hence, we conclude that the concept of Effort Expectancy has strong evidence to impact the individual's Intention to Share, and we argue that the process of data-sharing should be perceived to be free of effort and hindrances for the consumer.

As seen in the literature review, Effort Expectancy is mainly shaped by motivation or feelings of the individual, and during the focus groups we found out that these feelings are impacted by other factors, not mainly limited to our developed constructs. Additional associations with the perceived effort have been for example Selective Ignorance as seen in Chapter 5.3, which led us to believe that our initial factors are not able to explain all perspectives of Effort Expectancy.

Self-Efficacy – Weak Evidence for H9

We considered Self-Efficacy to have an impact on Effort Expectancy. Based on the main study outcome, we rejected the hypothesis and there was only little support in the focus groups. Thus, we conclude that Self-Efficacy, i.e. one's belief to make the right decision when it comes to data-sharing, has relatively weak support in our research. This finding is contradicting to our theoretical expectations, as mentioned in Chapter 2.3.3. As our questionnaire and focus group samples had high Digital Abilities, and did not emphasized Self-Efficacy, we interpret that these individuals do not reflect much on their confidence to share or not share data and perhaps even take data-sharing for granted. This was highlighted by P3c who said "For me it is very hard to find something that hinders me, because I feel like I am already sharing so much data." It would give additional insights to analyze a different sample with low digital abilities. We hypothesize that these individuals reflect on data-sharing more thoroughly, by also integrating their own capabilities in their considerations, leading to an increased importance of Self-Efficacy in our theoretical model.

Affect – Strong Evidence for H10

Based on the literature, we assumed that Affect dealing with the feeling when sharing data has a positive impact on Effort Expectancy. This path was of statistical significance in our quantitative analysis, and we found a lot of reasoning for it in our qualitative analysis. Therefore, we assume that an individual's Affect has a strong impact on the perceived amount of effort regarding data-sharing. As seen in the focus groups, the majority of participants proactively mentioned additional factors when answering the question of how they feel about data-sharing. Participants associated their Affect with other issues, such as trust or transparency, which led us to the conclusion that Affect needs to be considered in relation to other reasons. Therefore, we assume that the feelings and motivations, hence Affect, not only shape the Intention to Share through the UTAUT factor Effort Expectancy but also that they are influenced by and influence other variables.

6.1.3 Social Influence

Based on the SCAT, we assumed that individuals decide whether to share or not to share their data based on their social environment. This thought served as the underlying principle for the factor Social Influence that is said to impact an individual's behavior. In H3 we expected Social Influence to have a significant positive impact on the Intention to Share and throughout our research, we found weak evidence for this hypothesis. On the one hand, we rejected H3 in our main study outcome because there was a statistical negative regression, hence the opposite of what we expected. On the other hand, we found some support from the focus group participants regarding the concept of Social Influence and therefore this hypothesis deserves further attention.

Trust in Business – Strong Evidence for H5

We assumed that Trust in Business has a significant positive effect on Social Influence. We found strong evidence for this hypothesis because the quantitative analysis revealed a significant regression and the qualitative analysis revealed that many participants consider Trust in a Business to be important when evaluating to share their data. Trust is conveyed in various ways, as indicated in the focus groups. Moreover, the focus group participants stated that they have a higher trust in businesses, if businesses act in a transparent manner. This is another case of two constructs seemingly correlating with each other.

Institutional Trust – Strong Evidence for H6

Furthermore, we analyzed Institutional Trust and found that there is strong evidence for this hypothesis. This concept is statistically significant and was backed up by a lot of support during the focus group because many participants considered trust in the government to be relevant regarding data-sharing. It is important to not generalize Institutional Trust but to acknowledge differences among various nations, as highlighted during the focus groups.

In the questionnaire, we found out that more people have a higher trust in businesses than in a government. In all focus groups, we encountered that the participants slightly favored trusting the business instead of the trusting the government similar to the quantitative result. During the pre-study, the company representative from Samtrafiken mentioned that the Swedish government evaluates whether it should increase its Open Data initiatives and we see a potential risk in the future. It seems that the majority of people asked in this research project would favor a business, hence a non-governmental institution, to collect and supply data. If the government aims to implement more Open Data initiatives, it should keep in mind this tendency for distrust, which we will refer to in Chapter 7.2.

Social Referral – Weak Evidence for H7

Whether the referral of friends or family has an impact on Social Influence was analyzed next. The quantitative analysis showed that there is statistical significance, though there is a negative regression. Hence, we rejected our original hypothesis, which assumes a positive regression of Social Referral with Social Influence. During our follow-up study, many participants explained their reasoning and said that the referral of a trusted individual from their social surrounding without further information would make them actually less likely to also engage into data-sharing. One participant shared her reasoning “If I can say >No<, I will say no to sharing the data. It doesn’t really matter if someone else would recommend me to do so” (P3a) and another one emphasized that he would actually consider doing the opposite of the friend’s recommendation: “And it may be a reaction in the other direction if my family says >This is the new thing that’s what we are gonna do< and I would be like >No and I am not gonna do it<” (P1c). Reflecting on these reasons, we agreed that Social Influence might have an opposite impact on the

Intention to Share because a referral without background information might lead to more mistrust and hesitation to engage into data-sharing.

RECIPROCITY – WEAK EVIDENCE FOR H8

Finally, we analyzed the concept of Reciprocity, questioning whether knowing that other people share their data have an impact on the individual's intention. Based on the main study, we rejected this hypothesis and during the focus groups, we found little support for this hypothesis. Hence, there was weak evidence in our research project for the hypothesis to be true. In the focus group, we found various reasonings why individuals are not affected by Reciprocity, as mentioned in Chapter 5.2.3. Interestingly, the two participants who would be affected by Reciprocity, were identified as showing a higher than average altruism trait. Therefore, we do not want to rule out the possibility that Reciprocity has no impact at all, and suspect that the importance of Reciprocity depends on the personality traits.

6.1.4 Facilitating Conditions

Based on the theoretical framework, we expected that the Facilitating Conditions of the system have a positive impact on the consumers' intention to share their data, which we stated in H4. There was medium evidence for this hypothesis to be true, since we rejected the hypothesis in our main study but we found much support for this hypothesis in the follow-up study. While we argued in Chapter 4 that one reason could be the low internal consistency reliability of Facilitating Conditions, due to the abstract wording of the questionnaire items, another reason could be that it simply does not affect the Intention to Share. This conflict of ambiguous evidence is similarly mentioned in other studies (Karahanna & Straub, 1999; Thompson, et al., 1991). Given the fact that many focus group participants had clarification questions when talking about the Facilitating Conditions, we assumed that the abstract wording has been one of the causes why we found only medium evidence for the hypothesis, but we do not want to limit the abstract wording to be the only root cause for our medium evidence. Therefore, we recommend other researchers to explore further explanations of why there was ambiguous evidence.

TRANSPARENCY – STRONG EVIDENCE FOR H15

We expected Transparency to have a significant positive effect on Facilitating Conditions and based on our questionnaire results, this was considered statistically significant. Additionally, almost all of the 21 focus group participants mentioned Transparency as one of the main factors when analyzing how the system should be set up in the data-sharing context. Taken the results from both studies, we can clearly state that Transparency has found strong evidence in this research project. Furthermore, many participants of the focus groups explained that a company which demonstrates a transparent behavior is more likely to be trusted. Further interrelations among constructs are discussed in Chapter 6.3.

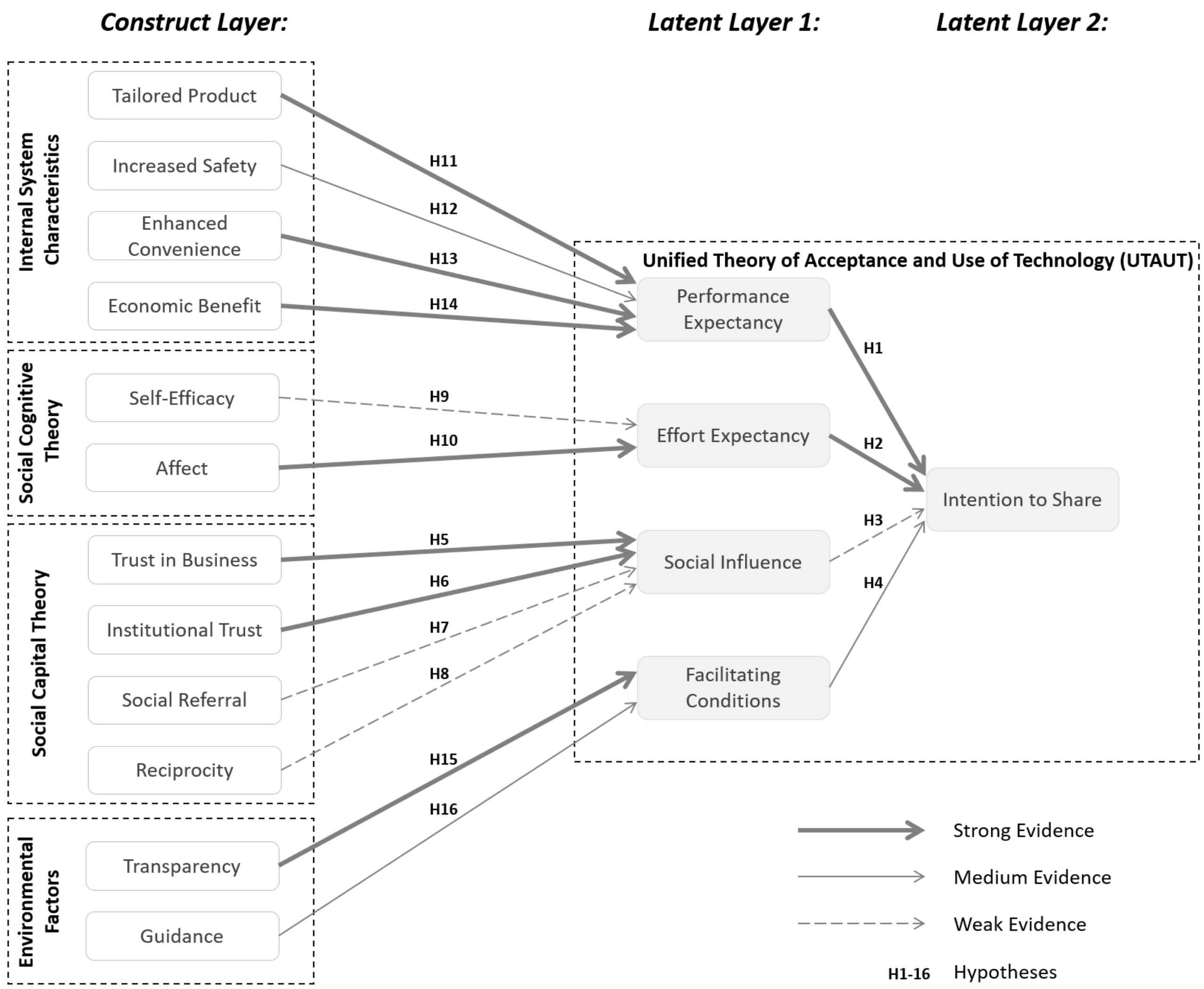
GUIDANCE – MEDIUM EVIDENCE FOR H16

Additionally, we hypothesized that Guidance has a positive impact on Facilitating Conditions, and we only found medium evidence for this hypothesis. In the qualitative study, we found little support because mostly reasons rejecting Guidance were mentioned. In the quantitative study, we observed that the path was statistically significant, even though it had a lower regression weight than Transparency, hence suggesting that Guidance is of subordinate importance than Transparency within Facilitating Conditions.

6.2 Updated Research Model

In Chapter 2.3.7, we introduced the adapted UTAUT model based on the literature review. Having collected data via the main study and the follow-up study, we discussed our findings in Chapter 6.1 and summarized the results for each hypothesis in Figure N. In Figure O, we illustrated the updated research model with its 16 hypotheses and whether the evidence for each is strong, medium or weak.

Figure O: Updated Research Model Based on the UTAUT



6.3 Interrelations Among Concepts

Furthermore, we were able to depict two new interrelations when reflecting on the results from the questionnaire and the focus group.

The first interrelation is between the constructs Transparency and Affect. In the outcome of the questionnaire, both showed a high correlation directly to the dependent variable Intention to Share, indicating that both have a strong impact on an individual's Intention to Share. Furthermore, as we explored during the focus groups, both Affect and Transparency seem to directly relate to each other. As one participant explained her thoughts "I feel better towards data-sharing when I have more information" (P1e), hence an individual's Affect might change depending on the availability of information, which in turn is shaped by the Transparency. We assume that some of the constructs that correlate with each other and possibly also impact each other, even if they are from different UTAUT factors.

The second interrelation deals with Performance Expectancy and Effort Expectancy. Especially in the TAM, the Perceived Usefulness and the Perceived Ease of Use are the only factors determining whether an individual accepts or rejects a new technology. After the insights from the focus groups, there was some evidence that individuals weight performance and effort against each other when evaluating data-sharing. The term "cost-benefit analysis" was mentioned frequently suggesting that an individual makes a holistic evaluation on both the costs and benefits when it comes to data-sharing. One participant summarized "If I feel there is no benefit, I assume my costs to be higher" (P2d). This interrelation between Performance Expectancy and Effort Expectancy is in line with our expectations based on the theory from Hennig-Thurau et al. (2007), who suggested that participation in a sharing is most likely when benefits are maximized and costs minimized, as discussed in Chapter 2.3.4.

6.4 Generalizability of the Results

In this part, we elaborate on how much these results from Chapter 6.1 depend on the mobility context and thus can or cannot be generalized on a higher level and applied to different contexts. At the end of each focus group, we asked the participants how their attitude towards data-sharing would change in another context and next, we discuss similar and different contexts.

SIMILAR CONTEXTS FOR DATA-SHARING

First, many participants favored data-sharing when they got something in return, either for themselves or for society at large. One participant directly mentioned consumer goods and retail as a positive example. She referred to the department store Åhlens and their loyalty program for cosmetics by saying "Åhlens does a really good job because based on my purchase history, I get monthly offers based on what I bought historically" (P1e). This example showed that especially the construct Tailored

Products in combination with Economic Benefits could be applied to other data-sharing situations besides just the mobility context.

Additionally, the context of entertainment was also mentioned a lot by participants. Hereby, the participants actively wanted to share their data, which can be recorded, analyzed and used for further recommendations based on past behavior. Specific examples that were mentioned in two focus groups were the music streaming service Spotify and the media provider Netflix. One participant summarized, "And there it works super well, and I am very happy that they use the information I am revealing, because I had Netflix for 2 years, and by now, the recommendations are extremely accurate" (P2d).

Finally, data-sharing was supported when the results are beneficial for society at large. One candidate said, "I think that sleep in the same way as mobility is interesting and that it is interesting on a societal level" (P1c). Similar to the inner-city mobility sector, where transportation systems are improved for the whole society, there seems to be similar contexts in which individuals are willing to share their data so that general systems can be improved to distribute gained insights and benefits.

DIFFERENT CONTEXTS FOR DATA-SHARING

Conversely, many participants highlighted that their intention to share their data decreases when shifting away from transportation data to another context, as exemplified in two examples from the focus groups. First, many participants mentioned they are reluctant to share any personal data that is very sensitive, such as political preferences or personal consumption habits like nutrition. Regarding the data-sharing of consumption habits we need to differentiate between non-personal consumption habits, like cosmetics as mentioned above, and personal consumption habits, such as nutrition. Additionally, participants are more reluctant to share their data if it could potentially harm them. One participant said, "I am less likely to share my data if there is something that can backfire, like medical information or how much money I have on the bank" (P3e) and there was a lot of agreement from the other participants.

7. Conclusion

The conclusion summarizes the main findings and introduces implications for managers and policymakers. Afterwards, we list the limitations of the conducted research and suggest further research issues.

7.1 Main Findings

We conducted a mixed methods study and based on the pre-study and the literature review, we adapted the UTAUT model to determine factors that influence an individual's intention to share data and the reasoning behind these factors. In Chapter 1.3 we listed our two research questions and started by asking what the factors are that influence the intention to share data in the Swedish mobility context. We found that Effort Expectancy and Performance Expectancy are strong factors that influence data-sharing decisions. Furthermore, we found that Affect, hence the feeling about the topic of data-sharing, and a high need for Transparency of the data-sharing process strongly correlate with the Intention to Share.

Henceforth, we elaborated on the second research question why these factors influence the Intention to Share. Many focus group participants explained that they conduct a cost-benefit analysis when evaluating whether to share their transportation data. Effort Expectancy and Performance Expectancy seem to be balanced against each other, which is in line with the results from the quantitative study. Additionally, many focus group participants had a very strong positive or negative opinion towards data-sharing, both supporting the importance of Affect. Furthermore, the focus groups gave us insights in the participant's reasoning why Social Referral would have a negative impact on the Intention to Share. This is due to the fact that they still want to know about the benefits of data-sharing, and some participants would actually become more cautious and reluctant to share their data after a referral. Finally, we were able to understand what kind of Transparency the participants request and based on the focus groups we found that Transparency is most valued regarding the purpose of the data collection, the terms and conditions, and what happens with the shared data afterwards.

7.2 Contributions

7.2.1 Theoretical Contribution

We successfully have proven that the UTAUT model can be extended by the SCAT, the SCOT, Internal System Characteristics, and Environmental Factors. It is possible to use the UTAUT for data-sharing adoption, hence beyond the context of technology and IT. Therefore, we contribute to academia by connecting previously unrelated theories in order to explain the consumer's intention to share data.

Furthermore, our research has two practical implication (see Chapter 1.5), which are discussed in the following.

7.2.2 Implications for the Businesses

The managerial implications are divided into the two areas of collection and use of data and design of services. First, it became evident in our study that consumers favor effortless processes. When companies are designing the collection and use-of-data process, they should make the system clear and simple and not overwhelm their consumers with information via very detailed terms and conditions. Since consumers value transparency and trust, we advocate that the terms and conditions should be short and intuitively understood to lower perceived hindrances.

Next, when the companies design their service, they should make the benefits of data-sharing observable for the consumer. Thereby, already incremental optimizations of the system are sufficient, as long as the consumer perceives some benefit either immediately or later on. Benefits are most valued in the form of enhanced services, such as saving time, or tailored products, such as suggesting optimized routes, or monetary incentives, such as discounts. To exemplify, we recommend a business, which lowers its costs because it analyzes shared data to optimize their operations, to forward a share of these savings directly to the consumer.

7.2.3 Implications for the Government

We identify three areas of interest for the government, which are Open Data, city planning, and legislation. First, we found that Open Data initiatives are seen as beneficial for society, even though many people lack an understanding of the rationale behind them. Therefore, consumers could be afraid when they are informed that their data is publicly shared via an Open Data initiative. Even though this happens in an anonymized way, there might remain some mistrust in the government's handling of data. We suggest governments to further educate about the reasons for Open Data initiatives and clearly state which data is accessible to prevent negative backlashes in the form of mistrust and decreasing reputation.

Second, mobility data can be important for city planning and we found evidence that people generally want to contribute to improving society when there are no negative consequences involved. We advise to educate about the benefits of sharing data in the mobility sector for urban development and actively promote services that require data-sharing by highlighting that it enhances the future city development to improve in terms of convenience, which would be highly perceived as beneficial by the participants.

Finally, regarding the legislative mandate, governments are advised to secure trust in the data economy in general and to ensure a fair competition for data as a good. As seen in our study, data-sharing is connected to some degree of distrust for many people, even if the data is not sensitive. If there are more clear mechanisms that regulate and shape the market and make sharing a more trustworthy process, those current externalities that limit the effectiveness could be reduced and thus, the market could be more beneficial for all stakeholders. Therefore, we suggest policymakers to design a suitable legal frame to protect consumer rights and in the same manner guarantee that data-sharing benefits are shared among all stakeholders.

7.3 Limitations

7.3.1 Conceptual Limitations

We developed the list of factors based on the pre-study and literature and thereby we see a risk for a conceptual limitation. Even though these factors were diverse and other researchers have proven that they were relevant, we cannot exclude the possibility that there are other relevant factors that influence the intention to share data. Hence, there is the risk that the factors that we have found were true, yet that they were not collectively exhaustive, meaning there were some factors that we have missed. For example, we have found two potential new constructs in Chapter 5.3 that could be added – Fairness and Selective Ignorance.

7.3.2 Methodological Limitations

MAIN STUDY

Considering our methodology for the quantitative study, we accumulated suggestions for the general improvement of the model fit in the process of SEM concerning further studies. First, the number of items per construct turned out to be low with three, as we also had to erase a few items for some constructs. We assume that four to five items for each construct would be more reliable and would have led to a better model fit. However, due to practical reasons we deliberately set the number of questions quite low to minimize the risk that participants will not finish the questionnaire due to its length.

Second, a few constructs seemed to lack internal consistency reliability, as seen with Facilitating Conditions during CFA. The most likely explanation is that the respondents of the questionnaire did not understand the questions correctly. We suspect that some questions were of abstract nature, which left much space for interpretation, leading to a higher variance. Nevertheless, our conducted piloting of the questionnaire did not give us any reason to believe in large misunderstandings of these questions.

FOLLOW-UP STUDY

We chose the focus group participants based on their answers in the questionnaire. This yielded the advantage of exploring specific answer patterns, though also served as a two-folded limitation. First, the focus group attendees were self-selected since they left their contact details, which possibly yielded to more extreme opinions since participants actively wanted to explain their reasoning. Furthermore, the focus group participants were already exposed to the questionnaire, hence biased with the statements that we had already presented to them.

7.4 Further Research Suggestions

We encountered the following areas for further research. First, researchers could use and modify our adapted UTAUT model to find more interrelated factors, elaborate on further reasons, or to apply it to other contexts. Even though we had to remove some variables in our adapted UTAUT model, we could add others into our model. Discussing the generalizability of the results, we listed different contexts where individuals have a higher or lower intention to share their data and we advise further researchers to build upon our theoretic foundation when exploring data-sharing in contexts such as the retail sector.

Furthermore, we started researching on moderating variables that have an influence on the various constructs and latent variables. We were able to derive initial insights on personality traits and digital abilities; a larger research study of the influence of those moderators could lead to more insights. Including additional moderating variables enables more granular consumer segmentation, which could have a great impact on the service design for mobility providers.

Finally, our research led us to reject the hypothesis regarding Social Referral and Reciprocity, indicating that in the data-sharing context individuals are not influenced by their social surrounding. We even found a negative regression for Social Referral and this seems contradictory to our literature review. On the opposite, the social surrounding in the form of trust, however, does have a positive effect on data-sharing. We advise further researchers to build upon our results and try to find the underlying reasons for this phenomenon.

V. References

Adell, E., 2009. *Driver Experience and Acceptance of Driver Support Systems - A Case of Speed Adaptation,* Lund: Lund University.

Adger, N., 2003. Social Capital, Collective Action, and Adaptation to Climate Change. *Economic Geography,* 79(4), pp. 287-404.

Ajzen, I., 1991. The Theory of Planned Behavior. *Organizational Behavior and Human Decision Processes,* 50(1), pp. 179-211.

Ajzen, I. & Sheikh, S., 2013. Action versus inaction: anticipated affect in the theory of planned behavior. *Journal of Applied Social Psychology,* 43(1), pp. 155-162.

Akaike, H., 1974. A New Look at the Statistical Model Identification. *IEE Transactions on Automatic Control,* 19(6), pp. 716-723.

Anderson, J. C. & Gerbing, D. W., 1988. Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin,* 103(3), pp. 411-423.

Attuquayefio, S. N. & Addo, H., 2014. Using the UTAUT model to analyze students' ICT adoption. *International Journal of Education and Development using Information and Communication Technology,* 10(3), pp. 75-86.

Awang, Z., Afthanorhan, A. & Mamat, M., 2016. The Likert scale analysis using parametric based Structural Equation Modeling (SEM).

Bandura, A., 1986. *Social Foundation of Thought and Action: A Social Cognitive Theory.* Englewood Cliffs: Prentice-Hall.

Bandura, A., 1989. Social Cognitive Theory. In: R. Vasta, ed. *Annals of child development.* Greenwich: JAI Press, pp. 1-60.

Banister, D., 2008. The sustainability mobility paradigm. *Transport Policy,* 15(1), pp. 73-80.

Bardhi, F. & Eckhardt, G., 2012. Access-Based Consumption: The Case of Car Sharing. *Journal of Consumer Research,* 39(4), pp. 881-898.

Barret, P., 2007. Structural Equation Modelling: Adjudging Model Fit. *Personality and Individual Differences,* 42(5), pp. 815-824.

Bentler, P. M., 1990. Comparative fit indexes in structural models. *Psychological Bulletin,* 107(2), pp. 238-246.

Berendt, B., Guenther, O. & Spiekermann, S., 2005. Privacy in e-commerce: stated preference vs. actual behavior. *Communication of the ACM,* 48(4), pp. 101-106.

Brinkmann, S. & Kvale, S., 2005. Confronting the ethics of qualitative research. *Journal of Constructivist Psychology,* 18(2), pp. 157-181.

Browne, M. W. & Cudeck, R., 1992. Alternative ways of assessing model fit. *Sociological Methods & Research,* 21(2), pp. 230-258.

Bryman, A. & Bell, E., 2015. *Business Research Methods.* 4th ed. Oxford: Oxford University Press.

Burns, R. P. & Burns, R., 2008. *Business Research Methods and Statistics Using SPSS.* 1 ed. Thousand Oaks, CA: SAGE Publications Ltd.

Byrne, B., 2009. *Structural Equation Modeling with AMOS: Basic Concepts, Applications, and Programming.* 2nd ed. NY, London: Routledge.

Byrne, B. M., 1998. *Structural Equation Modeling with LISREL, PRELIS and SIMPLIS: Basic Concepts, Applications and Programming.* 1 ed. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Calder, B. J., 1977. Focus Groups and the Nature of Qualitative Marketing Research. *Journal of Marketing Research,* 14(1), pp. 353-364.

Chang, H. H. & Chuang, S.-S., 2011. Social capital and individual motivations on knowledge sharing: Participant involvement as a moderator. *Information & Management,* 48(1), pp. 9-18.

Chang, H. H., Huang, C. Y., Fu, C. S. & Hsu, M. T., 2017. The effects of innovative, consumer and social characteristics on willingness to try nano-foods: Product uncertainty as a moderator. *Information Technology & People,* 30(3), pp. 653-690.

Choudhury, A., 2018. *Top Benefits of Big Data in the Transportation Industry.* [Online]
Available at: https://www.businesswire.com/news/home/20180914005108/en/Top-Benefits-Big-Data-Transportation-Industry-–
[Accessed 23 March 2019].

Chow, W. S. & Chan, L. S., 2008. Social network, social trust and shared goals in organizational knowledge sharing. *Information & Management,* 45(7), pp. 458-465.

Chui, M., Farrell, D. & Jackson, K., 2014. *How government can promote open data and help unleash over $3 trillion in economic value,* s.l.: McKinsey Global Institute.

Chun, S. A., Shulman, S., Sandoval, R. & Hovyd, E., 2010. Government 2.0: Making connections between citizens, data and government. *Information Polity,* 15(1), pp. 1-9.

Cliff, N., 1983. Some cautions concerning the application of causal modeling methods. *Multivariate Behavioral Research,* 18(1), pp. 115-126.

Cohen, B. & Kietzmann, J., 2014. Ride On! Mobility Business Models for the Sharing Economy.

Coleman, J., 1988. Social Capital in the Creation of Human Capital. *The American Journal of Sociology,* 94(1), pp. 95-120.

Courville, T. & Thompson, B., 2001. Use of structure coefficients in published multiple regression articles: β is not enough. *Educational and Psychological Measurement,* 61(1), pp. 229-248.

Creswell, J. W. & Plano Clark, V. L., 2011. *Designing and Conducting Mixed Methods Research.* 2nd ed. Los Angeles: Sage Publications.

Davis, F., 1989. Perceived Usefulness, Perceived Ease of use, and User Acceptance of Information Technology. *MIS Quarterly,* 13(3), pp. 319-340.

Davis, F., Bagozzi, R. & Warshaw, P., 1989. User Acceptance of Computer Technology: A Comparison of Two Theoretical Models. *Management Science,* 35(8), pp. 982-1003.

Davis, F., Bagozzi, R. & Warshaw, P., 1992. Extrinsic and Intrinsic Motivation to Use Computers in the Workplace. *Journal of Applied Social Psychology,* 22(14), pp. 1111-1132.

de Luca, S. & di Pace, R., 2015. Modelling users' behaviour in inter-urban carsharing program: a stated preference approach. *Transportation Research,* 71(1), pp. 59-76.

de Waal, C., 2001. *On Peirce.* Belmont, CA: Wadsworth.

Dillman, D., 2007. *Mail and Internet Surveys: The Tailored Design Method.* 2nd ed. Hoboken, New Jersey: John Wiley & Sons.

Dinev, T. & Hart, P., 2006. An extended privacy calculus model for e-commerce transactions. *Information Systems Research,* 17(1), pp. 61-80.

Dweck, C. & Leggett, E., 1988. A Social-Cognitive Approach to Motivation and Personality. *Psychology Review,* 95(2), pp. 256-273.

Economist, 2017. *The world's most valuable resource is no longer oil, but data.* [Online]
Available at: https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data
[Accessed 23 March 2019].

Erlingsson, C. & Brysiewicz, P., 2017. A hands-on guide to doing content analysis. *African Journal of Emergency Medicine,* 7(3), pp. 93-99.

European Commission, 2017. *The Digital Economy and Society Index (DESI).* Brussels: European Commission.

Fathema, N., 2013. *Structural Equation Modeling (SEM) of an Extended Technology Acceptance Model (TAM) to Report Web Technology Adoption Behavior in Higher Education Institutions,* Auburn, Alabama: Auburn University.

Feilzer, M. Y., 2009. Doing Mixed Methods Research Pragmatically: Implications for the Rediscovery of Pragmatism as a Research Paradigm. *Journal of Mixed Methods Research,* 4(1), pp. 6-16.

Firnkorn, J. & Muller, M., 2011. What will be the environmental effects of new free-floating car-sharing systems? The case of Car2Go in Ulm. *Ecological Economics,* 70(1), pp. 1519-1528.

Fishbein, M. & Ajzen, I., 1975. *Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research.* 1st ed. Massachusetts: Addison-Wesley.

Fortes, N. & Rita, P., 2016. Privacy concerns and online purchasing behaviour. *European Research on Business and Management Economics,* 22(1), pp. 167-176.

Gefen, D., Karahanna, E. & Straub, D., 2003. Trust and TAM in online shopping: an integrated model. *MIS Quarterly,* 27(1), pp. 51-90.

Gist, M. E., 1987. Self-efficacy: Implications for organizational behavior and human resource management. *Academic Management Review,* 12(1), pp. 472-485.

Goodall, W., Fishman, T. D., Bornstein, J. & Bonthron, B., 2017. The rise of mobility as a service. *Deloitte Review,* Issue 20, pp. 112-129.

Granovetter, M., 1973. The Strength of Weak Ties. *The American Journal of Sociology,* 78(6), pp. 1360-1380.

Grossklags, J. & Acquisti, A., 2007. When 25 cents is too much: An experiment on willingness-to-sell and willingness-to-protect personal information. *WEIS.*

Hair, J. F., Anderson, R. E., Tatham, R. L. & Black, W. C., 1998. *Multivariate Data Analysis.* 5th ed. New Jersey: Prentice-Hall International, Inc..

Hasan, B., 2007. Examining the effects of computer self-efficacy and system complexity on technology acceptance. *Information Resources Management,* 20(1), pp. 76-88.

Hazen, B., Overstreet, R. & Wang, Y., 2015. Predicting Public Bicycle Adoption Using the Technology Acceptance Model. *Sustainability,* 7(1), pp. 12558-14573.

Hennig-Thurau, T., Henning, V. & Sattler, H., 2007. Consumer file sharing of motion pictures. *Journal of Marketing,* 71(1), pp. 1-18.

Hietanen, S., 2014. "Mobility as a Service" - The new transport model?. *Eurotransport,* 12(2), pp. 2-4.

HM Treasury, 2018. *The Economic Value of Data: Discussion Paper,* London: HM Treasury.

Holme, I. M. & Solvang, B. K., 1997. *Forskningsmetodik: Om kvalitativa och kvantitativa metoder.* 2nd ed. Lund: Studentlitteratur.

Hooper, D., Coughlan, J. & Mullen, M. R., 2008. Structural Equation Modelling: Guidelines for Determining Model Fit. *Electronic Journal of Business Research Methods,* 6(1), pp. 53-60.

Hoshmand, L. T., 2003. Can lessons of history and logical analysis ensure progress in psychological science?. *Theory and Psychology,* 13(1), pp. 39-44.

Hox, J. & Bechger, T., 1998. An Introduction to Structural Equation Modeling. *Family Science Review,* Volume 11, pp. 354-373.

Iachello, G. & Hong, J., 2007. End-user privacy in human-computer interaction. *Foundations and Trends in Human-Computer Interactions,* 1(1), pp. 1-137.

Im, I., Hong, S. & Kang, M. S., 2011. An international comparison of technology adoption: Testing the UTAUT model. *Information & Management,* 48(1), pp. 1-8.

Janssen, M., Charalabidis, Y. & Zuiderwijk, A., 2012. Benefits, Adoption Barriers and Myths of Open Data and Open Government. *Information Systems Management,* 29(4), pp. 258-268.

Jentzsch, N., Preibusch, S. & Harasser, A., 2012. *Study on montezing privacy,* Heraklion: European Network and Information Security Agency (ENISA).

Jittrapirom, P. et al., 2017. Mobility as a Service: A Critical Review of Definitions, Assessments of Schemes, and Key Challenges. *Urban Planning,* 2(2), pp. 13-25.

Johnson, R. B. & Onwuegbuzie, A. J., 2004. Mixed Methods Research: A Research Paradigm Whose Time Has Come. *Educational Researcher,* 33(7), pp. 14-26.

Johnson, R. B., Onwuegbuzie, A. J. & Turner, L. A., 2007. Toward a Definition of Mixed Methods Research. *Journal of Mixed Methods Research,* 1(2), pp. 112-133.

Jöreskog, K. G., 1969. A General Approach to Confirmatory Factor Analysis. *Psychometrika,* 34(2), pp. 183-202.

Jöreskog, K. & Sörbom, D., 1993. *LISREL 8: Structural Equation Modeling with the SIMPLIS Command Language.* 1st ed. Chicago, IL: Scientific Software International Inc..

Karahanna, E. & Straub, D. W., 1999. The psychological origins of perceived usefulness and ease-of-use. *Information & Management,* 35(4), pp. 237-250.

Kim, C., Mirusmonov, M. & Lee, I., 2009. An empirical examination of factors influencing the intention to use mobile payment. *Computers in Human behavior.*

Kline, R. B., 2005. *Principles and practice of structural equation modeling.* 1st ed. New York: Guilford Press.

Knodel, J., 1993. Chapter 3 | The Design and Analysis of Focus Group Studies: A Practical Approach . In: D. L. Morgan, ed. *Successful Focus Groups: Advancing the State of the Art* . Thousand Oaks, CA: SAGE Publications, pp. 35-50.

Lamberton, C. & Rose, R., 2012. When is Ours Better than Mine? A Framework for Understanding and Altering Participation in Commercial Sharing Systems. *Journal of Marketing,* 76(4), pp. 109-125.

Lazaric, N. & Lorenz, E., 1998. Trust and Economic Learning. *Journal des Economistes et des Etudes Humaines,* 8(2-3), p. 10.

Lesk, M., 2012. The price of privacy. *IEEE Security and Privacy Magazine,* 10(5), pp. 79-81.

Lindloff, K., Pieper, N., Bandelow, N. C. & Woisetschläger, D. M., 2014. Drivers of carsharing diffusion in Germany: an actor-centred approach. *Automotive Technology Management,* 7(1), pp. 217-245.

Loebbecke, C. & Picot, A., 2015. Reflections on societal and business model transformation arising from digitization and big data analytics: A research agenda. *The Journal of Strategic Information Systems,* 24(3), pp. 149-157.

Loehlin, J. C., 1998. *Latent variable models: An introduction to factor, path, and structural analysis.* 3rd ed. Mahwah, NJ: Lawrence Erlbaum Associates, Inc..

Luarn, P. & Lin, H. H., 2005. Toward an Understanding of the Behavioral Intention to Use Mobile Banking. *Computers in Human Behavior,* 21(1), pp. 873-891.

Lund, A., 2017. *Kraftsamling Öppna Trafikdata,* Stockholm: Forum för transportinnovation.

Magalhaes, G. & Roseira, C., 2017. Open government data and the private sector: An empirical view on business models and value creation. *Government Information Quarterly.*

Manyika, J. et al., 2011. *Big data: The next frontier for innovation, competition, and productivity,* s.l.: McKinsey Global Institute.

Marsh, H. W., Balla, J. R. & McDonald, R. P., 1988. Goodness-of-fit indexes in confirmatory factor analysis: The effect of sample size. *Psychological Bulletin,* 103(3), pp. 391-410.

Mathieson, K., 1991. Predicting user intention: comparing the technology acceptance model with theory of planned behavior. *Information System Research,* 2(1), pp. 173-191.

Mauro, A., Greco, M. & Grimaldi, M., 2016. A formal definition of Big Data based on its essential features. *Library Review,* 65(3), pp. 122-135.

McDonald, R. P. & Ho, M.-H. R., 2002. Principles and Practice in Reporting Statistical Equation Analyses. *Psychological Methods,* 7(1), pp. 64-82.

McKinsey, 2011. *Big data: The next frontier for innovation, competition, and productivity,* Washington: McKinsey Global Institute.

McKinsey, 2012. *Mobility of the Future,* Munich: McKinsey & Company, Inc.

Merton, R. K., Fiske, M. & Kendall, P. L., 1956. *The Focused Interview.* Glencoe, Illinois: The Free Press.

MIT Technology Review & Oracle, 2016. *The Rise of Data Capital,* s.l.: MIT Technology Review.

Moeller, S. & Wittkowski, K., 2010. The burdens of ownership: Reasons for preferring renting. *Journal of Service Theory and Practice,* 20(2), pp. 176-191.

Moreno, E. M. O., Luna, E. B. d., Gómez, M. d. C. O. & López, J. E., 2014. Structural Equations Model (SEM) of a questionnaire on the evaluation of intercultural secondary education classrooms. *Suma Psicologica,* 21(2), pp. 107-115.

Morgan, D. L., 1998. *Planning Focus Groups.* Thousand Oaks, CA: Sage.

Muschalle, A., Stahl, F., Löser, A. & Vossen, G., 2013. Pricing approaches for data markets. *Enabling Real-Time Business Intelligence,* pp. 129-144.

Muthén, B. O., 2002. Beyond SEM: General Latent Variable Modeling. *Behaviormetrika,* 29(1), pp. 81-117.

Naranjo-Zolotov, M. et al., 2019. Examining social capital and individual motivators to explain the adoption of online citizen participation. *Future Generation Computer Systems,* 92(1), pp. 302-311.

Noland, R. & Polak, J., 2002. Travel time variability: A review of theoretical and empirical issues. *Transport Reviews,* 22(1), pp. 39-54.

Norman, G., 2010. Likert scales, levels of measurement and the "laws" of statistics. *Advances in Health Science Education,* 15(1), pp. 625-632.

OECD, 2018. *Open Government Data.* [Online]
Available at: http://www.oecd.org/gov/digital-government/open-government-data.htm
[Accessed 23 March 2019].

Onwuegbuzie, A. J., Dickinson, W. B., Leech, N. L. & Zoran, A. G., 2009. A Qualitative Framework for Collecting and Analyzing Data in Focus Group Research. *International Journal of Qualitative Methods,* 8(3), pp. 1-21.

Pemberton Levy, H., 2015. *Why and How to Value Your Information as an Asset.* [Online]
Available at: https://www.gartner.com/smarterwithgartner/why-and-how-to-value-your-information-as-an-asset/
[Accessed 23 March 2019].

Porter, C. & Donthu, N., 2006. Using the technology acceptance model to explain how attitudes determine Internet usage: The role of perceived access barriers and demographics. *Journal of Business Research,* 59(1), pp. 999-1007.

Portes, A. & Sensenbrenner, J., 1993. Embeddedness and Immigration: Notes on the Social Determinants of Economic Action. *The American Journal of Sociology,* 98(6), pp. 1320-1350.

Rehder, C., 2018. *Make Shipping Green Again: How big data practices affect environmental sustainability in shipping,* Stockholm: Stockholm School of Economics.

Reid, M., 2008. Integrating trust and computer self-efficacy with TAM: An empirical assessment of customers' acceptance of banking information systems (BIS) in Jamaica. *Journal of International Banking Commerce,* 12(1), pp. 1-18.

Rise, J., Sheeran, P. & Hukkelberg, S., 2010. The Role of Self-identity in the Theory of Planned Behavior: A Meta-Analysis. *Journal of Applied Social Psychology,* 40(5), pp. 1085-1105.

Ritchie, J. & Spencer, L., 1994. Qualitative data analysis for applied policy research. In: *Analysing Qualitative Data.* London: Routledge, pp. 173-194.

Robson, C., 2011. *Real World Research: A Resource for Users of Social Research Methods in Applied Settings.* 2nd ed. Sussex: A. John Wiley and Sons Ltd. .

Rock, R. & Moran, M., 2018. *A 21st Century Framework for Data Ownership,* Philadelphia: Microshare Inc.

Sandberg, K., 2014. *Transport Trends in Sweden,* Stockholm: Trafik Analys.

Schomm, F., Stahl, F. & Vossen, G., 2013. Marketplaces for data: an initial survey. *ACM SIGMOD Record,* 41(1), pp. 15-26.

Solove, D. J., 2005. A taxonomy of privacy. *University of Pennsylvania Law Review,* 154(3), pp. 477-560.

Spiekermann, S., Acquisti, A., Böhme, R. & Hui, K.-L., 2015. The challenges of personal data markets and privacy. *Electron Markets,* pp. 161-167.

Tabachnick, B. G. & Fidell, L. S., 2007. *Using Multivariate Statistics.* 5th ed. Boston, MA: Pearson Education.

Tashakkori, A. & Creswell, J. W., 2007. The New Era of Mixed Methods. *Journal of Mixed Methods Research,* 1(3), pp. 3-7.

Taylor, S. & Todd, P. A., 1995. Understanding information technology usage: A test of competing models. *Information Systems Research,* 6(1), pp. 144-176.

Teece, D. J., 2010. Business Models, Business Strategy and Innovation. *Long Range Planning,* 43(2-3), pp. 172-194.

Teo, T., 2012. Examining the intention to use technology among pre-service teachers: an integration of the technology acceptance model and theory of planned behavior. *Interactive Learning Environments,* 20(1).

Thompson, R. L., Higgins, C. A. & M., H. J., 1991. Personal Computing: Toward a Conceptual Model of Utilization. *MIS Quarterly,* 15(1), pp. 125-143.

Trafiklab, 2019. *Trafiklab.* [Online]
Available at: https://www.trafiklab.se/
[Accessed 23 March 2019].

Tsai, C.-H., 2014. Integrating Social Capital Theory, Social Cognitive Theory, and the Technology Acceptance Model to Explore a Behavioral Model of Telehealth Systems. *International Journal of Environmental Research and Public Health,* 11(1), pp. 4905-4925.

Urry, J. & Lyons, G., 2005. Travel time use in the information age. *Transportation Research Part A,* 39(1), pp. 257-276.

Valenzuela, S., Park, N. & Kee, K., 2009. Is There Social Capital in a Social Network Site?: Facebook Use and College Students' Life Satisfaction, Trust, and Participation. *Journal of Computer-Mediated Communication,* 14(4), pp. 875-901.

Venkatesh, V. & Davis, F. D., 1996. A model of the antecedents of perceived ease of use: Development and test. *Decision Science,* 27(1), pp. 451-481.

Venkatesh, V., Morris, M., Davis, G. & Davis, F., 2003. User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly,* 27(3), pp. 425-478.

Vetrò, A. et al., 2016. Open data quality measurement framework: Definition and application to Open Government Data. *Government Information Quarterly,* 33(1), pp. 325-337.

Wallén Warner, H., 2006. *Factors Influencing Drivers' Speeding Behaviour,* Uppsala: Uppsala Universitet.

Williams, M. D., Rana, N. & Dwivedi, Y. K., 2015. The unified theory of acceptance and use of technology (UTAUT): A literature review. *Journal of Enterprise Information Management,* 28(3), pp. 443-488.

Woolcock, M., 1998. Social capital and economic development: Toward a theoretical synthesis and policy framework. *Theoretical Sociology,* 27(1), pp. 151-208.

World Bank, 2019. *Information and Communication for Development: Data-Driven Development,* Washington: World Bank Group.

World Economic Forum, 2011. *Personal Data: The emergence of a new asset class,* Davos: World Economic Forum.

Wright, K. B., 2005. Researching Internet-Based Populations: Advantages and Disadvantages of Online Survey Research, Online Questionnaire Authoring Software Packages, and Web Survey Services. *Journal of Computer-Mediated Communication,* 10(3).

Wu, Y., Lau, T., Atkin, D. J. & Lin, C. A., 2011. A comparative study of online privacy regulations in the U.S. and China. *Telecommunications Policy,* 35(7), pp. 603-616.

Zhang, J., 2011. Data-Driven Intelligent Transportation Systems: A Survey. *IEEE Transactions on Intelligent Transportation Systems,* 12(4), pp. 1624-1639.

Zheng, X. et al., 2016. Big Data for Social Transportation. *IEEE Transactions on Intelligent Transportation Systems,* 17(3), pp. 620-630.

Zuiderwijk, A., Helbig, N., Gil-García, R. & Janssen, M., 2014. Special Issue on Innovation through Open Data - A Review of the State-of-the-Art and an Emerging Research Agenda. *Journal of theoretical and applied electronic commerce research,* 9(2), pp. 1-2.

VI. Appendix

Exhibit A: Overview of the Pre-Study Interview Partners

#	Date	Name	Organization	Function
1	2/08/2019	Peter Popovics	Stockholm School of Economics	Innovation Researcher
2	2/12/2019	Elias Arnestrand	Samtrafiken	Project Manager Strategy
3	2/20/2019	Kye Andersson	Peltarion	Head of Brand and Communications
4	3/01/2019	Douglas Stark	Voi Scooter	COO
5	3/25/2019	Mehdi Rafinia *(follow-up talk to interview #2 with E. Arnestrand)*	Samtrafiken	Project Manager

Exhibit B: Overview of Hypotheses and Theoretical Origin

Category	Hypothesis		Origin	Author
Performance Expectancy	H1	*Performance Expectancy has a significant positive effect on Intention to Share.*	UTAUT	Venkatesh et al. (2003)
Effort Expectancy	H2	*Effort Expectancy has a significant positive effect on Intention to Share.*	UTAUT	Venkatesh et al. (2003)
Social Influence	H3	*Social Influence has a significant positive effect on Intention to Share.*	UTAUT	Venkatesh et al. (2003)
Facilitating Conditions	H4	*Facilitating Conditions has a significant positive effect on Intention to Share.*	UTAUT	Venkatesh et al. (2003)
Social Influence	H5	*Trust in Businesses has a significant positive effect on Social Influence.*	SCAT	Reid (2008)
Social Influence	H6	*Institutional Trust has a significant positive effect on Social Influence.*	SCAT	Gefen et al. (2003)
Social Influence	H7	*Social Referral has a significant positive effect on Social Influence.*	SCAT	Lazaric & Lorenz (1998)
Social Influence	H8	*Reciprocity has a significant positive effect on Social Influence.*	SCAT	Portes & Sensenbrenner (1993)
Effort Expectancy	H9	*Self-Efficacy has a significant positive effect on Effort Expectancy.*	SCOT	Venkatesh & David (1996)
Effort Expectancy	H10	*Affect has a significant positive effect on Effort Expectancy.*	SCOT	Dweck & Leggett (1988)
Performance Expectancy	H11	*Tailored Product has a significant positive effect on Performance Expectancy.*	Internal system charac-teristics	Insights from the pre-study
Performance Expectancy	H12	*Increased Safety has a significant positive effect on Performance Expectancy.*	Internal system charac-teristics	Insights from the pre-study
Performance Expectancy	H13	*Enhanced Convenience has a significant positive effect on Performance Expectancy.*	Internal system charac-teristics	Insights from the pre-study
Performance Expectancy	H14	*Economic Benefit has a significant positive effect on Performance Expectancy.*	Internal system charac-teristics	Insights from the pre-study
Facilitating Conditions	H15	*Transparency has a significant positive effect on Facilitating Conditions.*	Environ-mental Factor	Insights from the pre-study
Facilitating Conditions	H16	*Guidance has a significant positive effect on Facilitating Conditions.*	Environ-mental Factor	Insights from the pre-study

Exhibit C: List of Moderating Variables

Demographics	Personality Traits	Digital Abilities
Gender	Skepticism	Digital Understanding
Age	Altruism	Digital Usage
Education	Curiosity	
Location of Living	Risk Attitude	

Exhibit D: Questionnaire Scenario Description

Imagine the following situation: It is a normal day in your life. You are about to move from point A to point B, which is from now on referred to as the *journey*. You can conduct this journey either by foot, bike, car, bus, or metro, or any combination of these means of transportation. Hereby, it does not matter whether you own the bike/car or whether it belongs to a bike- or car-sharing service.

Before you start the journey, you have the opportunity to share your data and reveal the following three pieces of information: your route including starting and final location, the according timestamps, and how you choose to travel. This is from now on called the *dataset*. This dataset of you will be anonymized and then stored with a uniquely recognizable address by a third-party company.

The company can make use of your data in any possible legal way, e.g. for product development, analytics etc.

Graphic taken from https://specials.nrc.nl/vodafoneslimmestad/

Exhibit E: Questionnaire Items

Model Measurement Scale – Question Items		*(All questions of this block appeared in random order)*
Intention to Share	IS_1	In general, I am willing to share my dataset.
	IS_2	In general, I am intending to share my dataset with the company.
	IS_3	I think there are many reasons why I should share my dataset.
Performance Expectancy	PE_1	I expect to get a benefit in return for sharing my dataset.
	PE_2	If I get something for it, I am more likely to share my dataset.
	PE_3	If it gives me a personal advantage, I will share my dataset.
Effort Expectancy	EE_1	If there is little or no work for me involved to share my dataset, I will do it.
	EE_2	If it takes little effort for me, I am more likely to share my dataset.
	EE_3	If it requires a lot of effort for me, I am less likely to share my dataset.
Social Influence	SI_1	I am less likely to share my dataset unless I know about the intentions of others.
	SI_2	Trust is very important to me when sharing my dataset.
	SI_3	The relationship to others is very important to me when sharing my dataset.
Facilitating Conditions	FC_1	If I know more about the specific context, I am more likely to share my dataset.
	FC_2	Unless I am able to evaluate the specific circumstances, I will not share my dataset.
	FC_3	It depends on the context whether I share my dataset.
Tailored Product	TaPr_1	If I share my dataset, I expect to get a service that is more customized to my preferences.
	TaPr_2	If I share my dataset, I expect to get a more personalized travel suggestion in return.
	TaPr_3	I share my dataset to get a suggested travel journey that is more relevant to me.
Increased Safety	InSa_1	If I share my dataset, I expect to conduct my travel journey in a safer way.
	InSa_2	I expect to conduct the journey with fewer risks of damage once I have shared my dataset.
	InSa_3	I share my dataset because it will reduce the amount of accidents on the road.
Enhanced Convenience	EnCo_1	If the required time of commuting will be decreased, I will share my dataset.
	EnCo_2	If I get a journey that is less stressful for me, I will share my dataset.
	EnCo_3	If I get a journey that is more pleasant for me, I will share my dataset.
Economic Benefit	EcBe_1	If I can conduct the journey for free, I will share my dataset.
	EcBe_2	I will share my dataset to receive a discount on the price I have to pay when conducting the journey.
	EcBe_3	I will share my dataset if I get money for doing so.
Self-Efficacy	SeEf_1	I am aware of my ability to handle my personal data.
	SeEf_2	I feel confident to make my decision whether to share my data or not.
	SeEf_3	Sharing my personal data is not stressful for me.
Affect	Af_1	I feel good about sharing my dataset.
	Af_2	I do not feel afraid about sharing my dataset.
	Af_3	It makes sense to share one's dataset.

Trust in Businesses	TrBu_1	I am more likely to share my dataset with a business if I trust the company to treat my data confidentially.
	TrBu_2	I am more likely to share my dataset with a business if I trust the company to not use it against me.
	TrBu_3	I am more likely to share my dataset with a business if I trust the company to protect my data against data theft.
Institutional Trust	InTr_1	I am more likely to share my dataset with a governmental organization if I trust the organization to treat my data confidentially.
	InTr_2	I am more likely to share my dataset with the government if I trust the government to not use it against me.
	InTr_3	I am more likely to share my dataset with a governmental organization if I trust the organization to protect my data against data theft.
Social Referral	SoRe_1	I will share my dataset if people that are important to me refer me to do so.
	SoRe_2	If close friends of mine tell me to share my dataset I will be convinced to do so.
	SoRe_3	If people that are close to me expect me to share my dataset, I will do so.
Reciprocity	Re_1	Because other people share data to improve the transportation ecosystem, I also feel obliged to share my dataset.
	Re_2	I feel peer pressured to share my dataset because others do so as well.
	Re_3	I feel the need to share my dataset because others have already done so.
Transparency	Tr_1	If I know what happens with my data, I rather tend to share my dataset.
	Tr_2	It is important to me to know at all times who can assess my data.
	Tr_3	Having transparency about the data that is collected about me makes it more likely for me to share my data.
Guidance	Gu_1	I expect to be guided by someone about what happens with my personal data.
	Gu_2	Having the option to ask someone for support is appealing to me when sharing my dataset.
	Gu_3	It is important to me to have someone who informs me about the process of sharing data.

Moderating Variable Items

(All questions of this block appeared in random order)

Skepticism	Sc_1	I generally believe that people have good intentions.
	Sc_2	If a stranger offers me a free cup of coffee or tea, I take it without hesitating.
Altruism	Al_1	I consider myself as being a helpful and generous part of the society.
	Al_2	In my freetime I like to volunteer for clubs, charity projects, or non-profit initiatives.
Curiosity	Cu_1	I consider myself as being an open-minded person.
	Cu_2	When I am travelling to a new place, I am curious to learn about the culture and history of it.
Risk Attitude	RiAt_1	In unexpected situations I feel confident and fearless.
	RiAt_2	I consider myself as an adventurous or thrill-seeking person.
Digital Understan-ding	DiUn_1	I am interested in articles or news that explain the latest technological trends.
	DiUn_2	I am aware of the wide range of functions that my smartphone has.
Digital Usage	DiUs_1	I use my phone to conduct various tasks throughout the day.
	DiUs_2	During the day I use my laptop or tablet quite a lot.

Exhibit F: Focus Group Interview Guide

Part 1: General questions about data sharing

Would you share your data in general according to this setting?

Why would you do so/do not so?

Part 2: Effort Expectancy

What are your hindrances to prevent you from sharing your data?

Why do you perceive these reasons as hindrances?

How do you feel about sharing your data and why do you feel so?

Part 3: Facilitating Conditions

Under which conditions should the setup be designed so that you agree to share the data?

Why are these circumstances relevant for you?

Part 4: Performance Expectancy

What incentives would make you more likely to share your data?

Why are these incentives important to you?

How should a mobility provider design its service for you to have an incentive to share your data?

What would be an appropriate amount to pay/save/earn to get you to share your data?

Part 5: Social Influence

Would you make your decision to share or not share your data depending on your social surrounding?

How important is trust in the mobility provider to you when sharing your data?

Does that differ whether a company, or governmental collects your data?

Why would it do so/do not so?

Part 6: Miscellaneous

Would your answers change if we are not asking you that for the transportation context, but in another context?

Why would they change?

What characteristics / experience mostly shaped your opinion regarding data-sharing?

Exhibit G: Overview Focus Groups Interview Partners

Focus Group 1

Time: *Monday, April 15th, 15:00 – 16:30*

Location: *Room C645, Stockholm School of Economics, Saltmätargatan 18-20*

Participant	Gender	Education	Share data	Attributes
P1a	male	Master	No	Altruism
P1b	female	Master	No	Altruism Curiosity
P1c	male	Bachelor	Yes	Curiosity Skepticism
P1d	female	Master	Yes	Altruism Skepticism
P1e	female	Master	No	Skepticism Curiosity Innovation Interest*

*additional input discovered during focus group

Focus Group 2

Time: *Tuesday, April 16th, 19:00 – 20:30*

Location: *Room 328, Stockholm School of Economics, Sveavägen 65*

Participant	Gender	Education	Share data	Attributes
P2a	male	Master	Yes	Altruism Education
P2b	female	Bachelor	Yes	Profession*
P2c	female	Master	Yes	Skepticism
P2d	male	Master	Yes	Technology Interest*
P2e	female	Master	No	Education Culture*
P2f	male	Master	No	Curiosity Culture*
P2g	female	Bachelor	Yes	Skepticism Culture*
P2h	female	Bachelor	Yes	Skepticism Culture*

*additional input discovered during focus group

Focus Group 3

Time: *Wednesday, April 17th, 15:00 – 16:30*

Location: *Room C645, Stockholm School of Economics, Saltmätargatan 18-20*

Name	Gender	Education	Share data	Attributes
P3a	female	Master	Yes	Culture*
P3b	male	Master	Yes	Skepticism
P3c	female	Master	Yes	Skepticism Culture*
P3d	female	Master	No	Altruism
P3e	male	Master	Yes	Digital Abilities
P3f	male	Master	Yes	Culture* Innovation Interest*
P3g	male	Master	Yes	Culture
P3h	male	Master	No	Altruism

*additional input discovered during focus group