Rethinking consumers' data sharing decisions with the emergence of multi-party computation

An experimental design for evaluation

Agahari, W.; de Reuver, G.A.

Publication date
2022

Document Version
Final published version

Published in
30th European Conference on Information Systems (ECIS 2022)

6-18-2022

Rethinking consumers' data sharing decisions with the emergence of multi-party computation: an experimental design for evaluation

Wirawan Agahari
Delft University of Technology, w.agahari@tudelft.nl

Mark de Reuver
Delft University of Technology, g.a.dereuver@tudelft.nl

RETHINKING CONSUMERS' DATA SHARING DECISIONS WITH THE EMERGENCE OF MULTI-PARTY COMPUTATION: AN EXPERIMENTAL DESIGN FOR EVALUATION

Research in Progress

Wirawan Agahari, Delft University of Technology, Delft, Netherlands, w.agahari@tudelft.nl

Mark de Reuver, Delft University of Technology, Delft, Netherlands, g.a.dereuver@tudelft.nl

Abstract

Consumers are increasingly reluctant to share their personal data with businesses due to mounting concerns over privacy and control. Emerging privacy-enhancing technologies like multi-party computation (MPC), which allows generating insights while consumers retain data control, are challenging the current understanding of why consumers share their data. In this research-in-progress paper, we develop and evaluate an instrument and experimental design to investigate the impact of MPC on consumers' willingness to share data and its antecedents. Preliminary analysis from a pre-study (N=300) indicates a good fit for our model. Also, MPC enhances consumers' control and trust while reducing privacy concerns and risk, ultimately increasing data sharing willingness. The findings suggest that privacy-enhancing technologies significantly affect both the willingness to share data itself and its typical antecedents. The next step will conduct a large-scale online experiment using the developed instruments to evaluate further the impact of MPC on consumers' willingness to share data.

Keywords: Multi-party computation, willingness-to-share, experimental research, information privacy, privacy-enhancing technologies

1 Introduction

More and more (personal) data are being generated in today's digital world, thanks to recent advances in technologies such as the Internet-of-Things (IoT). As a result, data is viewed as fuel for our (digital) economy and society (The Economist, 2017). Especially when data is shared, it can benefit consumers, businesses, and society, such as automation, personalization, business efficiency, and knowledge creation (Cichy et al., 2021; Schomakers et al., 2020). However, consumers are increasingly concerned about their information privacy (Malhotra et al., 2004). Specifically, consumers feel that they have lost control over their data and often have no idea how companies handle it (S. Spiekermann & Novotny, 2015). Moreover, they are afraid of the possible data misuse by third parties, ultimately leading to consumers losing trust in sharing data and refraining from data sharing (Pal et al., 2021). As a result, economic opportunities may be missed, and businesses largely remain dependent on large data aggregators and cloud providers for access to data.

Novel privacy-enhancing approaches are expected to stimulate data sharing while respecting consumers' privacy and control over data. One prominent example is Multi-Party Computation (MPC), which could overcome data sharing barriers (Zare-Garizy et al., 2018) and foster new economic opportunities (Zafrir, 2020). MPC is a cryptographic technique that involves sharing information while not disclosing submitted data between any involved parties (Yao, 1986). Nevertheless, MPC implementation has long remained limited, and only recently, computing resources are sufficient to execute demanding MPC algorithms.

In the Information Systems (IS) literature, information privacy theory is commonly used to study consumers' willingness to share data (Smith et al., 2011). Prior studies focus on various antecedents for willingness to share data, such as control (Krasnova et al., 2010), privacy (Malhotra et al., 2004), risk (Dinev & Hart, 2006), and trust (Pavlou, 2003). However, MPC is essentially a different approach to data sharing that poses a new context to understand consumers' willingness to share data. MPC only computes and shares the analysis results (data insights) while keeping the input data private (Bestavros et al., 2017; Elliot & Quest, 2020). In this way, consumers could be more willing to share data with MPC, as they retain privacy and control while facing less risks and lower need to trust other actors. Therefore, IS scholars need to understand the mechanisms through which MPC affects consumers' willingness to share data. As MPC is currently emerging in the market, IS scholars should start conceptualizing now what MPC means for data sharing decisions.

This research-in-progress paper aims to develop and evaluate an instrument and experimental design to investigate the impact of MPC on consumers' willingness to share data and its antecedents. Based on this objective, we formulate the following research question: *how to evaluate the impact of MPC on consumers' willingness to share data and its antecedents?* To do this, we draw upon information privacy theory (Malhotra et al., 2004; Smith et al., 2011) and look into key factors that scholars have been focusing on when researching consumers' willingness to share data, namely perceived control, privacy concerns, perceived risk, and trust. In this way, we can examine why consumers' willingness to share data could be greater with MPC in place. As an approach, we opt for the between-subjects experimental research design by comparing MPC and Trusted Third Party (TTP) as a conventional, non-MPC-based solution. We specify our context to sharing driving data in a personal data marketplace, meaning that consumers can voluntarily sell their data (Schomakers et al., 2020) instead of being collected by third parties as part of digital data devices usage (Cichy et al., 2021). Nevertheless, such data are highly sensitive and create mounting concerns for consumers (Docherty et al., 2018), making them refrain from sharing data despite possibilities to benefit from data-driven services (Athanasopoulou et al., 2019; Kaiser et al., 2021). Hence, we established a unique setting with a high potential for value creation and a low willingness to share, making it relevant to see the impact of MPC.

2 Background

2.1 Multi-party computation

MPC is a powerful instrument because it provides a possible solution to Computation on Encrypted Data (CoED) (Archer et al., 2018). MPC comprises two or more input parties, each with a concealed dataset, whereby they jointly compute an objective functionality (e.g., an application-oriented task such as electronic voting) based on their inputs (Zhao et al., 2019). MPC used the secret sharing technique by splitting each parties' data into multiple parts (i.e., secret shares), computed based on the requested function, then recombined to generate the results (Pedersen et al., 2007). While the theoretical concept of MPC is not novel (Yao, 1982), recent advances in computational power and efficiency are bringing MPC increasingly closer to real-life applications.

A popular illustration of MPC is the millionaire's problem (Yao, 1986), a secure comparison function to determine which one of two millionaires is richest, without revealing the net worth to each other. MPC has been implemented in several use cases, such as auction-based pricing (Bogetoft et al., 2009), gender wage gap analysis (Lapets et al., 2018), and improving healthcare intervention (van Egmond, 2020). Nevertheless, its application within data marketplaces is lacking (Koch et al., 2021; Roman & Vu, 2019). Further, various barriers are hindering large scale implementations of MPC, such as usability issues (i.e., too complex to understand by non-experts), technical issues (i.e., performance limitations and scalability), and legal aspects (i.e., current regulations discourage cooperation) (Choi & Butler, 2019).

2.2 Data marketplaces

Data marketplaces are digital platforms managed by *data marketplace operators* that enable *data providers* (i.e., individuals or organizations) to share and sell data to *data buyers* (Koutroumpis et al., 2020; M. Spiekermann, 2019). Access to the data, manipulation, and the use of the data by other entities is commonly governed by the data marketplace using a range of standardized or negotiated licensing models (Stahl et al., 2016). Both static and dynamic data streams can be shared and traded in data marketplaces, in which it is accessible via individual file downloads, Application Programming Interfaces (APIs), or customized web interfaces (Fricker & Maksimov, 2017; M. Spiekermann, 2019). On top of that, data marketplaces also offer complementary applications and services such as data visualizations, data valuation, and data analytics (Mucha & Seppala, 2020; M. Spiekermann, 2019). Hence, such platforms would create value for its participants by lowering transaction costs, stimulating innovation by third-party developers, and generating network effects. This paper focuses on personal data marketplaces, where individuals can directly offer what kind of data they want to share in exchange for monetary compensation (Schomakers et al., 2020).

2.3 Information privacy theory

Information privacy refers to the ability of individuals to control when, how, and to what extent information about them is shared with others (Popovič et al., 2017; Westin, 1968). One theoretical stream of information privacy is the privacy calculus, where consumers weigh the benefits and costs of data sharing (Culnan & Armstrong, 1999). Nevertheless, we argue that the impact of MPC is more relevant to the costs of data sharing. Therefore, we control for perceived benefits by making it constant and assume that consumers will get benefits by sharing their data, such as money and better services (see Section 3.1).

Scholars have been using information privacy theory to explain consumers' information disclosure decisions in various contexts, such as e-commerce (Dinev & Hart, 2006), social media (Hajli & Lin, 2016), e-health (Juga et al., 2021), and IoT services (Bélanger et al., 2021). Information privacy theory stresses *privacy concerns* as an antecedent of data sharing, defined as the degree of an individual's concern on who has access to the data that is being shared and how other parties use it (Smith et al., 1996). This means that those with a deep concern over their information privacy are likely to refrain from data sharing and demand more privacy protection (Cichy et al., 2021; Schomakers et al., 2020).

Related to the definition of information privacy is the notion of *control over data,* making it one of the key factors in consumers' decision to share data (Dinev et al., 2013; S. Spiekermann, 2005). Perceived control refers to the extent to which an individual believes that he/she is able to manage the release and dissemination of personal information (Xu et al., 2011). This factor is essential as consumers are increasingly worried that they have lost control over their data and have no idea how other parties used their data (S. Spiekermann & Novotny, 2015), making them even more reluctant to share data.

Information privacy theory also highlights the vital role of *trust* as a prerequisite of data sharing (Richter & Slowinski, 2019; M. Spiekermann, 2019). Following Ažderska (2012) and Kehr et al. (2015), we describe trust as an individual's belief that another party will act as expected and does not do harmful things such as misusing personal data. Since we focus on consumers' perspectives as data providers, trust is divided into trust in data buyers and data marketplaces operator. Further, trust is often associated with *perceived risk* (Hart & Saunders, 1997), in which a higher degree of trust reduces perceived risk in data sharing (Dinev & Hart, 2006; Pavlou, 2003). Building on Xu et al. (2011), perceived risk refers to the expectation of losses if someone decides to engage in data sharing. In this regard, if people think that sharing their data is a risky thing to do and could cause harm to them, they will refrain from data sharing (Wang et al., 2016).

2.4 Hypotheses development

We develop hypotheses for this study by applying information privacy theory to the context of MPC (see Table 1). We expect that MPC could empower consumers to exercise greater control over data than

the conventional solution (TTP). With MPC, data buyers will only receive insights from the computational analysis between multiple data consumers. In this way, consumers have more control over how data buyers utilize the data. Also, with MPC in place, consumers are expected to perceive a lower degree of risks and privacy concerns in data sharing than TTP. Since data buyers will not receive individual consumers' data, the risks for consumers to take part in data sharing might be lower. Moreover, MPC could also increase consumers' trust in data marketplace operators and data buyers. Both parties will only be able to access computation results and therefore cannot, in theory, misuse individual consumers' data. Furthermore, with all features of MPC, we expect that consumers are more willing to share data through MPC than TTP.

Hypotheses	
H_1	Consumers perceive more control over data sharing when data marketplaces use MPC rather than TTP
H_2	Consumers have less privacy concerns on data sharing when data marketplaces use MPC rather than TTP
H_3	Consumers perceive less risks over data sharing when data marketplaces use MPC rather than TTP
H_4	Consumers trust operators more when data marketplaces use MPC rather than TTP
H_5	Consumers trust data buyers more when data marketplaces use MPC rather than TTP
H_6	Consumers are more willing to share data when data marketplaces use MPC rather than TTP

Table 1. A summary of the hypotheses for this research

3 Research approach

3.1 Experimental design

We conduct a controlled, survey-based online experiment to investigate the impact of MPC on the willingness to share data in privacy-enhancing data marketplaces. We opt for a between-subject design with three experimental conditions: Trusted Third Party (TTP), MPC, and made-up privacy technology (referred to as Data-Computation-Protection/DCP). We include a made-up technology in our study because we use a description of MPC in our experiment rather than a working demonstrator or prototype. A critique could be that users attribute value to the term of MPC rather than to the underlying ideas in the technology. Therefore, we want to see if different privacy technologies would make any differences in perception or do not matter for users, even if the technology does not exist.

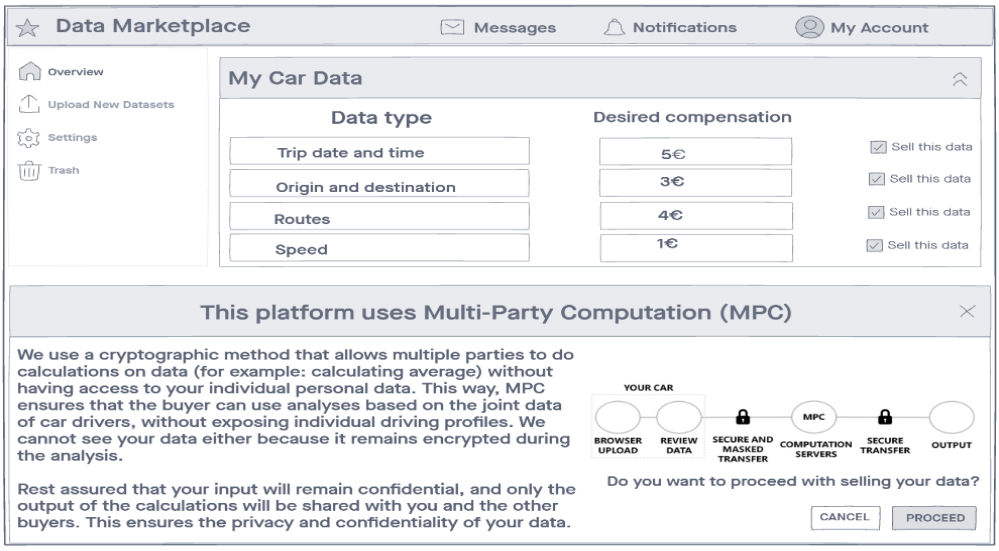

Figure 1. A screenshot preview of the mock-up for the MPC scenario

Each condition is different in terms of the description of the technology, how it works in the data marketplaces, and a screenshot preview of the mock-up (including a disclaimer on how the technology works, see Figure 1). The main difference is that, in TTP, users send data to a central system in which the data will be analyzed and stored. Meanwhile, in MPC and DCP scenarios, the data is encrypted in the car, and only the analysis results are revealed to the prospective buyer. Both MPC and DCP scenarios are identical, with the name of the technology being the only difference.

The experiment consisted of four parts. After introducing the purpose of the study and the consent form, we presented a persona that owns a connected car that generates driving data and could sell it via data marketplaces. We asked participants to imagine that the mobility service providers are interested in buying their driving data (e.g., trip date and time, destination and routes history, and driving speed) via data marketplaces. To control for benefits, we asked participants to assume that they would receive better driving advice and financial compensation by sharing their data in this scenario. Next, we randomly assigned participants to one of the three conditions (TTP, MPC, or DCP) and introduced them to their respective scenarios. Subsequently, participants filled out the post-test questionnaire to rate their perception of the data marketplace presented to them. We concluded the experiment with participants filling out the demographic questions, which are the same for all conditions.

3.2 Participants

We recruited participants using the online crowdsourcing platform Prolific (Palan & Schitter, 2018; Peer et al., 2017). Our population comprises consumers that have a driving license. For this pre-study, we restricted the sample to participants who are 18 years old and older as this is the minimum age to have a driving license in most countries. We also excluded participants from the United Kingdom (e.g., participants who have a UK nationality or currently live in the UK) as this will be our sample for the main study. Furthermore, we also excluded participants who had already taken part in other studies by the authors to ensure the reliability of the answers of our participants. We offered financial compensation to participants based on the recommendation provided by Prolific.

We conducted the data collection on 9 September 2021, and we managed to recruit a sample of 300 participants (165 male, 126 female, nine others/prefer not to say). The average age of participants was 30.1 years old (SD = 8.87), and about 70.7% of them are part of the younger generation (18-34 years old). Most of them reside in the United States (53.3%), France (20.7%), and South Africa (6.7%). The majority had already finished a graduate degree (35%), followed by an undergraduate degree (30.3%) and high school diploma/A-level education (18%). More than half of the participants currently work full-time (56%) or part-time (13.7%) and primarily work in the IT (19.7%) or finance industry (7.3%). About one-third of our participants hold a managerial position, either at a junior (5.7%), middle (18%), or upper management level (9.3%). In terms of access to and ownership of cars, only 10% of participants did not have access at all. The rest are either own a car (63.7%), have access via family members (22.3%), or have access via leasing or rental (4%). Further, 53.4% of participants claimed that they are familiar with data marketplaces, while only 23% of participants had prior knowledge about privacy-enhancing technologies before taking part in the survey. The underlying datasets are available at the following link: https://doi.org/10.4121/19403534

3.3 Scales and measurement variables

As quantitative studies on sharing driving data via privacy-enhancing data marketplaces are lacking, we developed a 5-scale Likert questionnaire by adapting measures from previous studies on information privacy (see Table 2). We modified survey items by Xu et al. (2011) to measure both perceived control and perceived risk. For privacy concerns, we adopted measures developed by Dinev and Hart (2006). Meanwhile, for trust in data buyers and data marketplace operator, we used measures by Kehr et al. (2015) and adjusted the items based on the actors in question. Finally, to measure willingness to share data, we used measures by Pavlou (2003).

4 Preliminary results

4.1 Confirmatory factor analysis

We conduct a Confirmatory Factor Analysis (CFA) using JASP to validate our constructs and measurement model (Brown & Moore, 2012). Through five rounds of analysis, we assess the model fit, construct validity, and identify areas of misfit (modification indices). To assess the model fit, we use measures like the Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), and Root Mean Square Error of Approximation (RMSEA). We follow a suggestion by Hu and Bentler (1999) for a good fit of those three measures: both CFI and TLI should be 0.95 or higher, and RMSEA should be 0.06 or lower. The results show a good level of the fit index of the model, with CFI = 0.988, TLI = 0.984, and RMSEA = 0.043.

Next, we looked into standardized factor loadings (λ) of each survey item using a threshold of 0.70 (Fornell & Larcker, 1981) and removed two items that did not meet the criteria. Then, we assessed the internal reliability of our model by looking at the Composite Reliability (CR) of each construct, which should have a value of 0.7 or higher. Subsequently, we assessed convergent validity through the Average Variance Extracted (AVE), which should be greater than 0.5 (Fornell & Larcker, 1981). We also examined the discriminant validity of the constructs by checking whether the correlation among constructs is lower than the square root of AVE (Fornell & Larcker, 1981). From our analyses, we establish internal reliability and convergent validity. We also establish the discriminant validity as all inter-construct correlation coefficients are well below the square root of AVE. Furthermore, we assessed modification indices to identify cross-loadings, which are items that load on the 'wrong' construct due to high correlations between the items from two different constructs. After five rounds of analysis, we removed two items for having very high modification indices (higher than 10) and one other item because it cross-loaded on all constructs. Our final model comprises six factors and 16 items (see Table 2). For source of the survey items, see section 3.3.

Factor	Item wording	λ	R^2	AVE	CR
Perceived control (CTRL)	I believe I have control over who can access the sensitive data I provided to this data marketplace.	0.84	0.71	0.65	0.85
	I think I have control over what kind of sensitive data is shared by this data marketplace to other companies.	0.82	0.68		
	I believe I have control over how other companies use the sensitive data I provided to this data marketplace.	0.75	0.57		
Privacy concerns (PRIV)	I am concerned that other parties could find sensitive information about me on this data marketplace.	0.82	0.67	0.72	0.84
	I am concerned about providing my sensitive data to this data marketplace because of what other parties might do with it.	0.88	0.78		
Perceived risk (RISK)	I find it risky to provide my sensitive data via this data marketplace.	0.85	0.73	0.76	0.86
	There would be too much uncertainty associated with providing my sensitive data to this data marketplace.	0.88	0.77		
Trust in data marketplace operator (TRSD)	I expect this data marketplace would be trustworthy regarding my sensitive data.	0.83	0.69	0.72	0.89
	This data marketplace would tell the truth and fulfill promises related to my sensitive data.	0.86	0.74		
	I expect this data marketplace would be honest with me regarding the sensitive data I would provide.	0.86	0.74		
Trust in data buyers (TRSB)	I expect that data buyers would be trustworthy in handling the data they got from this data marketplace.	0.92	0.85	0.87	0.95
	I expect that data buyers would tell the truth and fulfill promises in handling the data they got from this data marketplace.	0.95	0.89		

	I expect that data buyers would be honest when handling the data they got from this data marketplace.	0.93	0.86		
Willingness to share data (WTSD)	Given the chance, I would share my data via this data marketplace.	0.92	0.85	0.85	0.94
	Given the chance, I predict that I should share my data via this data marketplace in the future.	0.93	0.86		
	It is likely that I will share my data via this data marketplace in the near future.	0.91	0.83		

Table 2. Measurement model

In the last step, we conducted Multi-Group Confirmatory Factor Analysis (MGCFA) to see if we could compare the three experimental conditions (see section 3.1). We estimated the model using configural invariance testing and found a good level of the fit index, with CFI = 0.980, TLI = 0.975, and RMSEA = 0.053. All groups also showed convergent and discriminant validity, with all standardized estimates higher than 0.7, CR higher than 0.7, and AVE higher than 0.5. Moreover, the comparison between the square root of AVE and all inter-construct correlation coefficients in all groups suggests discriminant validity. In addition, modification indices in all groups were not an issue as we find no very high modification indices and no items that were cross-loaded in all other constructs.

4.2 Comparing the effect of three data sharing scenarios: one-way ANOVA

Before we proceed with further analysis, we compute composite scores for each construct by averaging the items that belong to the construct (see Table 2). For instance, since perceived control consists of three items, we computed a new variable in the dataset by calculating the average of these three items. For all factors, we will use these composite scores in the remainder of the analysis.

Factors	**TTP (N=100)**	**MPC (N=100)**	**DCP (N=100)**	**One-way ANOVA**				**Kruskal-Wallis**		
	Mean	**Mean**	**Mean**	**df**	**F**	**ω²**	**p**	**df**	**Statistic**	**p**
CTRL	2.78	3.16	3.15	2	4.00	0.02	0.019*			
PRIV	3.61	3.14	3.16					2	10.10	0.006**
RISK	3.35	2.95	3.01	2	3.89	0.02	0.021*			
TRSD	3.51	3.80	3.80	2	3.59	0.02	0.029*			
TRSB	3.16	3.61	3.15	2	5.51	0.03	0.004**			
WTSD	3.04	3.57	3.16					2	12.03	0.002**

Table 3. Comparing the effect of data sharing scenarios on all factors (p < .05, ** p < .01)*

In the next step of the analysis, we perform a one-way ANOVA to compare the effect of three data sharing scenarios (TTP, MPC, and DCP, see section 3.1) on all factors (see Table 3). We also want to provide evidence of the utility of the developed instruments (i.e., whether they can be used to compare different conditions in an experiment). Levene's test indicates that, except for privacy concerns ($p = 0.009$) and willingness to share data ($p = 0.018$), variances in all groups are equal for all other factors. Based on a one-way ANOVA, we find a significant effect of different data sharing scenarios on perceived control [$F(2,297) = 4.00$, $p = 0.019$, $\omega^2 = 0.02$], perceived risk [$F(2,297) = 3.89$, $p = 0.021$, $\omega^2 = 0.019$], trust in data marketplaces operator [$F(2,297) = 3.59$, $p = 0.029$, $\omega^2 = 0.017$], and trust in data buyers [$F(2,297) = 5.51$, $p = 0.004$, $\omega^2 = 0.029$] at the $p < .05$ level for the three scenarios. The post hoc tests using Tukey's correction find that perceived control in both MPC ($p = 0.037$) and DCP ($p = 0.042$) groups are greater than TTP. However, we find no significant differences between MPC and DCP ($p = 0.999$). Similarly, both MPC ($p = 0.008$) and DCP ($p = 0.018$) groups perceive higher trust in data buyers than TTP. However, we also find no significant differences between MPC and TTP ($p = 0.961$). Meanwhile, we find no significant differences in perceived risk in data sharing between DCP and TTP ($p = 0.073$) and between DCP and MPC ($p = 0.920$). However, the TTP group perceive a higher degree of risk in data sharing than MPC ($p = 0.027$). Finally, we find no significant differences between the three groups concerning trust in data marketplaces operator.

Furthermore, a Kruskal-Wallis test reveals that privacy concerns [H(2) = 10.102, p = 0.006] and willingness to share data [H(2) = 12.030, p = 0.002] are significantly affected by different scenarios. Pairwise comparisons show that both MPC (p_{holm} = 0.006) and DCP (p_{holm} = 0.008) groups perceive lower privacy concerns than TTP, but we find no significant differences between MPC and DCP (p_{holm} = 0.411). Both MPC and DCP also significantly increase participants' willingness to share data (p_{holm} = 0.001 and p_{holm} = 0.022 respectively) than TTP. However, there are no significant differences between MPC and DCP (p_{holm} = 0.133).

5 Conclusions and outlook

This research-in-progress paper has developed and evaluated instrument and experimental design to investigate the impact of MPC on consumers' willingness to share data and its antecedents. We focused on a specific context of personal data marketplaces in the automotive sector. Using confirmatory factor analysis, we find a good fit for our model based on the adapted survey items, including measurement invariance across the three conditions (MPC, TTP, and made-up tech called DCP). In this regard, we make a methodological contribution by developing an instrument and experimental design for evaluating the impact of MPC on decisions about data sharing. Ultimately, we can generate insights on the impact of MPC on consumers' data sharing decisions and its antecedents. Furthermore, we will be able to investigate the causal mechanisms through which MPC affects data sharing decisions.

We also find that the MPC and DCP groups scored significantly higher than the TTP groups for all measured constructs except one, implying that car drivers would be more willing to share their driving data via privacy-enhancing approaches than a conventional solution. One reason might be that, as we found, car drivers feel MPC (and even made-up technology like DCP) provides control over data, reduced risk, and lower privacy concerns. As a result, car drivers could perceive a higher degree of trust towards data buyers, ultimately increasing their willingness to share. Nevertheless, we find no differences between MPC and DCP for all constructs, suggesting that car drivers might not care how the privacy-enhancing technology is named.

A limitation is that we excluded perceived benefits in our model despite being a dominant factor explaining individual data sharing decisions (Culnan & Armstrong, 1999) since we expect MPC will not affect this factor. Thus, MPC only changes the costs in the privacy calculus model and not the benefits. Nevertheless, if data buyers understand that MPC reduces the risks and concerns of consumers, data buyers can pay less for consumers' data, implying that the benefits could also be affected by MPC. In this way, the privacy calculus would remain the same. A second limitation is that we only used hypothetical scenarios and mock-ups that are not working prototypes due to the context of privacy-enhancing data marketplaces based on MPC, which is still limited. To counter this, we extensively informed participants about the setting at the beginning of the survey. As the next step, we will conduct the main survey with larger participants (N=1500) using Prolific and restrict the sample to the United Kingdom's population. We will analyze the data by (1) comparing means between TTP, MPC, and DCP groups (one-way ANOVA); and (2) using Structural Equation Modelling (SEM) to model relations between perceived control, perceived risk, privacy concerns, trust, and willingness to share data and examine if MPC moderates these established theoretical relationships.

Once we conduct and analyze the final survey results, we expect to contribute toward theory and practice. Specifically, we will contribute to understanding how MPC could challenge the core assumption behind antecedents for sharing personal data, as derived from information privacy theory. Additionally, our study will provide a fundamental basis for further user evaluation research on the impact of MPC on consumers' willingness to share data. Future studies could build on our research by (1) employing design science research (DSR) to develop privacy-enhancing approaches to share data; (2) further conceptualizing and developing an instrument to evaluate the impact of MPC on data sharing decisions (cf. Hoehle & Venkatesh, 2015); and (3) utilize qualitative approaches to understand how MPC could challenge the core assumption of data sharing. Finally, our study will also provide practical insights for businesses in general and data marketplaces operators in particular to consider privacy-enhancing approaches like MPC to offer assurances and stimulate data sharing by individuals.

Acknowledgments

The work leading to this paper has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 825225.

References

Archer, D. W., Bogdanov, D., Lindell, Y., Kamm, L., Nielsen, K., Pagter, J. I., Smart, N. P., & Wright, R. N. (2018). From Keys to Databases—Real-World Applications of Secure Multi-Party Computation. *The Computer Journal*, *61*(12), 1749–1771. https://doi.org/10.1093/comjnl/bxy090

Athanasopoulou, A., de Reuver, M., Nikou, S., & Bouwman, H. (2019). What technology enabled services impact business models in the automotive industry? An exploratory study. *Futures*, *109*, 73–83.

Ažderska, T. (2012). Co-evolving trust mechanisms for catering user behavior. *IFIP International Conference on Trust Management*, 1–16.

Bélanger, F., Crossler, R. E., & Correia, J. (2021). Privacy Maintenance in Self-Digitization: The Effect of Information Disclosure on Continuance Intentions. *ACM SIGMIS Database: The DATABASE for Advances in Information Systems*, *52*(2), 7–24.

Bestavros, A., Lapets, A., & Varia, M. (2017). User-centric distributed solutions for privacy-preserving analytics. *Communications of the ACM*, *60*(2), 37–39.

Bogetoft, P., Christensen, D. L., Damgård, I., Geisler, M., Jakobsen, T., Krøigaard, M., Nielsen, J. D., Nielsen, J. B., Nielsen, K., Pagter, J., Schwartzbach, M., & Toft, T. (2009). Secure Multiparty Computation Goes Live. In R. Dingledine & P. Golle (Eds.), *Financial Cryptography and Data Security* (pp. 325–343). Springer. https://doi.org/10.1007/978-3-642-03549-4_20

Brown, T. A., & Moore, M. T. (2012). Confirmatory factor analysis. *Handbook of Structural Equation Modeling*, 361–379.

Choi, J. I., & Butler, K. R. B. (2019). Secure Multiparty Computation and Trusted Hardware: Examining Adoption Challenges and Opportunities. *Security and Communication Networks*, *2019*, 1368905. https://doi.org/10.1155/2019/1368905

Cichy, P., Salge, T.-O., & Kohli, R. (2021). Privacy Concerns and Data Sharing in the Internet of Things: Mixed Methods Evidence from Connected Cars. *MIS Quarterly*.

Culnan, M. J., & Armstrong, P. K. (1999). Information privacy concerns, procedural fairness, and impersonal trust: An empirical investigation. *Organization Science*, *10*(1), 104–115.

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

Dinev, T., Xu, H., Smith, J. H., & Hart, P. (2013). Information privacy and correlates: An empirical attempt to bridge and distinguish privacy-related concepts. *European Journal of Information Systems*, *22*(3), 295–316.

Docherty, I., Marsden, G., & Anable, J. (2018). The governance of smart mobility. *Transportation Research Part A: Policy and Practice*, *115*, 114–125.

Elliot, D., & Quest, L. (2020, January 14). *It's time to redefine how data is governed, controlled and shared. Here's how.* World Economic Forum. https://www.weforum.org/agenda/2020/01/future-of-data-protect-and-regulation/

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, *18*(1), 39–50.

Fricker, S. A., & Maksimov, Y. V. (2017). Pricing of data products in data marketplaces. *International Conference of Software Business*, 49–66.

Hajli, N., & Lin, X. (2016). Exploring the security of information sharing on social networking sites: The role of perceived control of information. *Journal of Business Ethics*, *133*(1), 111–123.

Hart, P., & Saunders, C. (1997). Power and trust: Critical factors in the adoption and use of electronic data interchange. *Organization Science*, *8*(1), 23–42.

Hoehle, H., & Venkatesh, V. (2015). Mobile Application Usability: Conceptualization and Instrument Development. *MIS Quarterly*, *39*(2), 435–472.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55.

Juga, J., Juntunen, J., & Koivumäki, T. (2021). Willingness to share personal health information: Impact of attitudes, trust and control. *Records Management Journal*.

Kaiser, C., Stocker, A., Viscusi, G., Fellmann, M., & Richter, A. (2021). Conceptualising value creation in data-driven services: The case of vehicle data. *International Journal of Information Management*, *59*, 102335.

Kehr, F., Kowatsch, T., Wentzel, D., & Fleisch, E. (2015). Blissfully ignorant: The effects of general privacy concerns, general institutional trust, and affect in the privacy calculus. *Information Systems Journal*, *25*(6), 607–635.

Koch, K., Krenn, S., Pellegrino, D., & Ramacher, S. (2021). Privacy-preserving Analytics for Data Markets using MPC. *ArXiv Preprint ArXiv:2103.03739*.

Koutroumpis, P., Leiponen, A., & Thomas, L. D. (2020). Markets for data. *Industrial and Corporate Change*, *29*(3), 645–660.

Krasnova, H., Spiekermann, S., Koroleva, K., & Hildebrand, T. (2010). Online social networks: Why we disclose. *Journal of Information Technology*, *25*(2), 109–125.

Lapets, A., Jansen, F., Albab, K. D., Issa, R., Qin, L., Varia, M., & Bestavros, A. (2018). Accessible Privacy-Preserving Web-Based Data Analysis for Assessing and Addressing Economic Inequalities. *Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies*, 1–5. https://doi.org/10.1145/3209811.3212701

Malhotra, N. K., Kim, S. S., & Agarwal, J. (2004). Internet Users' Information Privacy Concerns (IUIPC): The Construct, the Scale, and a Causal Model. *Information Systems Research*, *15*(4), 336–355. https://doi.org/10.1287/isre.1040.0032

Mucha, T., & Seppala, T. (2020). *Artificial Intelligence Platforms–A New Research Agenda for Digital Platform Economy*.

Pal, D., Funilkul, S., & Zhang, X. (2021). Should I Disclose My Personal Data? Perspectives From Internet of Things Services. *IEEE Access*, *9*, 4141–4157. https://doi.org/10.1109/ACCESS.2020.3048163

Palan, S., & Schitter, C. (2018). Prolific. Ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, *17*, 22–27.

Pavlou, P. A. (2003). Consumer acceptance of electronic commerce: Integrating trust and risk with the technology acceptance model. *International Journal of Electronic Commerce*, *7*(3), 101–134.

Pedersen, T. B., Saygın, Y., & Savas, E. (2007). *Secret Sharing vs. Encryption-based Techniques For Privacy Preserving Data Mining*.

Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017). Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology*, *70*, 153–163.

Popovič, A., Smith, H. J., Thong, J. Y. L., & Wattal, S. (2017). Information Privacy. In Ashley Bush & Arun Rai (Eds.), *MIS Quarterly Research Curations*. http://misq.org/research-curations

Richter, H., & Slowinski, P. R. (2019). The data sharing economy: On the emergence of new intermediaries. *IIC-International Review of Intellectual Property and Competition Law*, *50*(1), 4–29.

Roman, D., & Vu, K. (2019). Enabling Data Markets Using Smart Contracts and Multi-party Computation. In W. Abramowicz & A. Paschke (Eds.), *Business Information Systems Workshops* (pp. 258–263). Springer International Publishing. https://doi.org/10.1007/978-3-030-04849-5_23

Schomakers, E.-M., Lidynia, C., & Ziefle, M. (2020). All of me? Users' preferences for privacy-preserving data markets and the importance of anonymity. *Electronic Markets*, *30*(3), 649–665.

Smith, H. J., Dinev, T., & Xu, H. (2011). Information privacy research: An interdisciplinary review. *MIS Quarterly*, 989–1015.

Smith, H. J., Milberg, S. J., & Burke, S. J. (1996). Information privacy: Measuring individuals' concerns about organizational practices. *MIS Quarterly*, 167–196.

Spiekermann, M. (2019). Data marketplaces: Trends and monetisation of data goods. *Intereconomics*, *54*(4), 208–216.

Spiekermann, S. (2005). *Perceived Control: Scales for Privacy in Ubiquitous Computing* (SSRN Scholarly Paper ID 761109). Social Science Research Network. https://doi.org/10.2139/ssrn.761109

Spiekermann, S., & Novotny, A. (2015). A vision for global privacy bridges: Technical and legal measures for international data markets. *Computer Law & Security Review*, *31*(2), 181–200. https://doi.org/10.1016/j.clsr.2015.01.009

Stahl, F., Schomm, F., Vossen, G., & Vomfell, L. (2016). A classification framework for data marketplaces. *Vietnam Journal of Computer Science*, *3*(3), 137–143.

The Economist. (2017). *The world's most valuable resource is no longer oil, but data*. https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data

van Egmond, M. B. (2020). *Identifying heart failure patients at high risk using MPC*. The Sugar Beet: Applied MPC. https://medium.com/applied-mpc/identifying-heart-failure-patients-at-high-risk-using-mpc-ab8900e75295

Wang, T., Duong, T. D., & Chen, C. C. (2016). Intention to disclose personal information via mobile applications: A privacy calculus perspective. *International Journal of Information Management*, *36*(4), 531–542.

Westin, A. F. (1968). Privacy and freedom. *Washington and Lee Law Review*, *25*(1), 166.

Xu, H., Dinev, T., Smith, J., & Hart, P. (2011). Information privacy concerns: Linking individual perceptions with institutional privacy assurances. *Journal of the Association for Information Systems*, *12*(12), 1.

Yao, A. C.-C. (1982). Protocols for secure computations. *23rd Annual Symposium on Foundations of Computer Science (Sfcs 1982)*, 160–164. https://doi.org/10.1109/SFCS.1982.38

Yao, A. C.-C. (1986). How to generate and exchange secrets. *27th Annual Symposium on Foundations of Computer Science (Sfcs 1986)*, 162–167.

Zafrir, N. (2020). *Beyond trust: Why we need a paradigm shift in data-sharing*. World Economic Forum. https://www.weforum.org/agenda/2020/01/new-paradigm-data-sharing/

Zare-Garizy, T., Fridgen, G., & Wederhake, L. (2018). A Privacy Preserving Approach to Collaborative Systemic Risk Identification: The Use-Case of Supply Chain Networks. *Security and Communication Networks*, *2018*, e3858592. https://doi.org/10.1155/2018/3858592

Zhao, C., Zhao, S., Zhao, M., Chen, Z., Gao, C.-Z., Li, H., & Tan, Y. (2019). Secure Multi-Party Computation: Theory, practice and applications. *Information Sciences*, *476*, 357–372. https://doi.org/10.1016/j.ins.2018.10.024