Received: 19 April 2021 | Revised: 27 September 2021 | Accepted: 7 November 2021

DOI: 10.1002/cb.2012

ACADEMIC PAPER

WILEY

Willingness to share data: Contextual determinants of consumers' decisions to share private data with companies

Kurt Alexander Ackermann[1] | Linda Burkhalter[2] | Thoralf Mildenberger[3] | Martin Frey[3] | Angela Bearth[1]

[1]Institute of Marketing Management, School of Management and Law, Zurich University of Applied Sciences, Winterthur, Switzerland

[2]Institute of Natural Resource Sciences, Zurich University of Applied Sciences, Wädenswil, Switzerland

[3]Institute of Data Analysis and Process Design, Zurich University of Applied Sciences, Winterthur, Switzerland

Correspondence
Kurt Alexander Ackermann, Institute of Marketing Management, School of Management and Law, Zurich University of Applied Sciences, Winterthur CH-8400, Switzerland.
Email: acka@zhaw.ch

Present address
Angela Bearth, Consumer Behavior, Institute for Environmental Decisions, Federal Institute of Technology (ETH) Zurich, Zurich, Switzerland

Funding information
Innosuisse - Schweizerische Agentur für Innovationsförderung, Grant/Award Number: 19006.1PFES-ES

Abstract

In the age of digitalization, customer and consumer data have become a valuable source of information for companies. However, to obtain these data, companies depend on peoples' willingness to share (WTS) their private data with them. By means of a large-scale online experiment with more than 20,000 participants, we investigated the extent to which peoples' WTS private data is affected by contextual factors. We complement and extend previous research by (i) simultaneously addressing several contextual factors that companies can largely control themselves, (ii) comparing their relative impacts on WTS, and (iii) explicitly examining interactions between these contextual factors in addition to their specific univariate effects. Concretely, we investigate contextual factors, such as the type of data requested, the purpose for which the data are used, the industry sector a corresponding company belongs to, the type of compensation offered for the shared data, and the degree to which the data allows for personal identification. Our data suggest that all these factors do affect peoples' WTS significantly, while there are also multiple significant interaction effects between these contextual factors. For instance, we found that a better intuitive match between the core business a company is engaged in and the type of data that is requested, results in higher proportions of people who are willing to share the corresponding data with the corresponding company. Hence, companies may benefit from tuning their requests for consumer or customer data according to the specific context in which they operate.

1 | INTRODUCTION

It is a long-existing marketing practice to target consumers based on their needs and desires regarding products, services, and promotions. Due to innovations in technology, such as big data (Lim et al., 2018), internet of things (Xia et al., 2012; Ziegeldorf et al., 2014), and smart devices (Silverio-Fernández et al., 2018), it has recently become easier for companies and organizations to access information about customers as new sources of information emerge and become available (e.g., wearables). For instance, it is quite common in many countries by now that health insurers monitor their customers' physical condition with the help of health data collected by fitness apps and wearables, with the goal to reward the healthy and active customers.

Obviously, new opportunities to collect, store, analyze, and use personal data are very promising for a company's marketing activities since a much deeper understanding of consumers can result from these opportunities. Nevertheless, from a consumer's perspective, new data-driven business practices are associated with an increase in nontransparency. It is difficult for consumers to grasp the amount of data collected by the multitude of new data sources, and to understand how interconnected devices and technologies allow for the combination of data to create new information about them (Cumbley & Church, 2013).

Due to this increasing nontransparency, consumers' concerns about data collection, storage, and use are increasing, too. For example, in a 2015 survey on EU citizens' privacy concerns, only 35% of respondents agreed that providing personal information was not a big issue for them (European Commission, 2015). In a more recent study with American consumers, only 10% stated that they feel they have complete control over their personal information, and only 25% believed that most companies handle their sensitive personal data responsibly (PricewaterhouseCoopers, 2017). Due to changes in the European law on data protection and privacy (General Data Protection Regulation, 2016), the will of consumers becomes more powerful and consumers regain more control over their personal information.

The practice to use the consumers' data without their consent and against their will is no longer permitted and can lead to high fines and loss of reputation for companies (GDPR.EU, 2019). Consequently, companies and organizations must be more transparent about the data they collect and therefore, become more dependent on the consumers' willingness to share (WTS) their personal data. Both from a theoretical and a practical perspective, it is thus essential to better understand which factors determine a consumer's WTS her or his personal data with companies.

To advance the state of knowledge in this respect, previous research has primarily focused on consumers' attitudes and concerns towards data sharing or the univariate effects of specific situational factors in isolation as potential determinants of WTS. However, in real-world settings, there are often a multitude of factors that are simultaneously and interactively affecting a consumer's WTS. To take account of this complexity, we complement and extend previous research by (i) simultaneously addressing several contextual factors that companies can largely control themselves, (ii) comparing their relative impacts on WTS, and (iii) explicitly examining interactions between these contextual factors in addition to their specific univariate effects. In the subsequent theory section, we explicate the rationale for our study in more detail and elaborate on the importance of contextual factors for fostering a better understanding of WTS.

2 | THEORETICAL BACKGROUND

2.1 | WTS and privacy concerns

Over the last decades, different constructs that may be related to a person's intention to disclose personal data have been discussed in the fields of marketing and consumer research, psychology, and information systems (Li, 2011). In particular, consumers' *privacy concerns* have been intensely investigated as these perceptions and attitudes precede consumers' intentions and behavior, and are therefore an important mediator between companies' practices and consumers' behavior (Beke et al., 2018; Phelps et al., 2000; Smith et al., 1996). The more concerned consumers are about how personal data are collected, used, and stored by companies, the less willing they should be to share personal information with companies (Bansal et al., 2016). Although consumers' privacy concerns are well studied, the construct has some limitations. Evidently, privacy concerns capture only the negative side of consumers' perceptions and they cannot explain why consumers are sometimes willing to disclose personal information, although they perceive certain risks of doing so. Weighing benefits and risks from data sharing against each other, the so-called privacy calculus (Dinev & Hart, 2006), offers a broader explanation, because it assumes that consumers determine for themselves whether they perceive the consequences of the collection, storage, and use of personal information as beneficial enough to outweigh the corresponding risks. For instance, a person may decide to use a fitness tracker and share information on her fitness activities because she considers the benefits of keeping track of her physical condition more important than the risk of her data being passed on to third parties. A second conceptual limitation of the construct of privacy concerns is that many authors consider it a stable personal disposition and belief (Malhotra et al., 2004; Smith et al., 1996). One well-known example is Westins' Privacy Segmentation Index, which assumes that consumers can be clustered in three groups named “Fundamentalists,” “Pragmatics,” and “Unconcerned” (Kumaraguru & Cranor, 2005). However, such dispositional approaches assume that people's behavior is stable over time and not responsive to situational characteristics. Consequently, such approaches are limited in their power to explain how individuals respond to increasing complexity arising from emerging new sources of data or new data handling practices, for instance.

In the present paper, we propose that the intention to disclose private data depends on contextual factors on top of personal characteristics. In line with this assumption, Yun and colleagues note in a recent review paper that “researchers have called for research that investigates contexts of PIP [personal information privacy] concerns” (Yun et al., 2019, p. 571). The present paper aims to contribute to answer these calls.

2.2 | Contextual determinants of WTS

Several authors have postulated that situational or contextual factors may have a decisive impact on consumers' WTS personal data (Bansal et al., 2016; Kayhan & Davis, 2016; Marwick & Hargittai, 2019; Roeber et al., 2015; Taylor et al., 2015). To date, several approaches to categorize such contextual factors have been proposed (Anderson & Agarwal, 2011; Beke et al., 2018; Milne et al., 2017; Phelps et al., 2000; Yun et al., 2019). However, research has often focused on one or two of conceivable contextual factors only, neglecting potential interactions between them. The categories that are most commonly proposed are the following. Type of personal information/data; the collecting company/organization; information use, such as personalization or third-party use; compensation/rewards offered; and the degree of (perceived) control. In the following, we elaborate the state of research for the just mentioned categories and we derive our research hypotheses.

2.2.1 | Type of personal information

One example of contextual factors refers to different types of personal information that a person may share with a company (Lim et al., 2018; Milne et al., 2017; Phelps et al., 2000). Different types of requested information evoke different degrees of privacy concerns (Marwick & Hargittai, 2019; Phelps et al., 2000). For instance, consumers may perceive disclosing financial information as a larger threat to privacy than sharing purchase histories or demographic information, such as gender and age. In particular, concerns regarding health-related data, medical information, or geolocation data, have gained more and more attention from researchers in the last years (Bearth & Siegrist, 2020; Yun et al., 2019). Moreover, Roeber et al. (2015) found that people perceive financial data and health record data as particularly sensitive private information. Further types of personal information are media usage, shopping behavior, location data, personal interests, and interactions on social network sites, for instance (Roeber et al., 2015; Treiblmaier & Pollach, 2007). Hence, in accordance with the current state of research in this area, we state the following first hypothesis:

> **Hypothesis 1.** Consumers' willingness to share personal data with a company depends on the type of requested data.

2.2.2 | Data collector (industry sector) and intended use of data

From a consumer's perspective, information on who collects a particular type of data and what purpose these data are used for may also be relevant with respect to privacy concerns. To date, several studies have investigated people's privacy attitudes and behavior in the context of a specific industry branch, such as finance (Bansal et al., 2016), healthcare (Anderson & Agarwal, 2011), or retail and e-commerce (Jai & King, 2016). However, it is yet unclear to which extent companies from different industry sectors evoke different degrees of privacy concerns in consumers. For instance, are consumers differentially willing to share personal data with a bank in comparison to a retailer?

To date, companies can use consumer data in numerous ways. For instance, consumer data can be used for the purpose of personalizing communication or segmenting and profiling consumers for customized offers (Treiblmaier & Pollach, 2007). However, in the age of digitization, the ways companies can utilize consumer data go beyond classical marketing tasks and extend to possibilities such as selling the data to third parties, using data to predict future behavior or improve customer insights in general. In this regard, Roeber et al. (2015) point out that contextual factors, such as the sector an organization collecting data is operating in, and how the data is going to be used, have rarely been investigated so far. Addressing this research gap, we state our second and third hypotheses as follows.

> **Hypothesis 2.** People's willingness to share their personal data with a company depends on the industry sector that company belongs to.

> **Hypothesis 3.** People's willingness to share their personal data with a company depends on the purpose for which the data is used.

2.2.3 | Type of compensation/reward

There is initial evidence that consumers are willing to disclose personal information if they receive a reward in return (Gabisch & Milne, 2014; Jai & King, 2016; Li et al., 2010). However, research differentiates between monetary and nonmonetary incentives. In studies about monetary incentives, where consumers receive a financial reward for their personal data (Roeber et al., 2015), it has already been shown that consumers' willingness to accept monetary incentives in return for their data is dependent on the particular context (Acquisti et al., 2013). In contrast, evidence on the effect of nonmonetary incentives on consumer's WTS data is less conclusive to date. However, Gabisch and Milne (2014) found that nonmonetary incentives were not perceived by consumers as a "payment" and therefore had no positive effect on consumers' WTS. To address the question to what extent different types of incentives differentially affect consumers' WTS personal data with companies, we state our fourth hypothesis as follows.

> **Hypothesis 4.** People's willingness to share personal data with a company depends on the type of reward provided in return.

2.2.4 | Control and data anonymity

Phelps et al. (2000) assume that people's privacy concerns and their WTS personal data are affected by the control people perceive to have over their personal data, such that a higher level of perceived control over the use of personal information is supposed to attenuate privacy concerns. Consequently, the higher the perceived control, the more likely consumers are to share their personal information, even in the case of very sensitive data (Brandimarte et al., 2013). Companies and organizations can implement several practices to increase consumers' perceived degree of control over their data to encourage disclosure behavior. For example, Roeber et al. (2015) found that the right to delete personal data on request increased people's WTS. Another way of providing consumers with the opportunity to keep control over their data and retain privacy is to collect and use anonymized data (Hoffmann, Novak, & Peralta, 1999). Hence, we state our fifth hypothesis as follows.

> **Hypothesis 5.** People's WTS personal data with a company depends on the degree to which that data is anonymous, meaning that the shared data does not allow for personal identification.

2.2.5 | Interaction effects

While previous studies often investigated specific contextual determinants (e.g., reward, type of data) in the context of a particular industry sector (e.g., social media networks; health data; e-commerce), the present study aims to develop and test a model that can account for people's WTS across a variety of different contexts. This broader and more holistic view extends previous research and allows for making comparisons between different contextual factors and hence contribute to a better understanding of the interplay between them. For example, are people more willing to share a particular type of data that will be used for a particular purpose when the company requesting that data stems from industry sector "A" rather than "B" (e.g., retailer vs. bank)? Moreover, to what extent does the type of compensation provided for sharing personal data matter on top and would the opportunity to stay anonymous moderate the corresponding effects? To address these questions, we explicitly investigate the statistical interactions between all of the contextual factors under consideration. However, as this part of our analysis is exploratory, we do not state specific research hypotheses.

The rest of the paper is structured as follows. In the next section, we describe the method we used to recruit participants and collect data, and further explain our experimental design. Subsequently, we present the results, discuss these with respect to practical and theoretical implications, and close with concluding remarks.

3 | METHOD

To address our research question, we conducted a large-scale online experiment. The online experiment consisted of several scenarios, each describing a case in which a particular company requests a particular type of data from its customers. For each scenario a participant was confronted with, the participant simply had to indicate whether he or she would be willing to share the corresponding type of data with the corresponding company by making a yes-or-no decision.

The scenarios were constructed on the basis of five scenario variables: (1) industry sector, (2) type of requested data, (3) intended use of data, (4) type of compensation for shared data, and (5) granting of anonymity. Each scenario variable comprised 2–10 attributes as outlined in Table 1. As for the scenario variables (2), (3), and (4), the corresponding attributes were selected based on qualitative research on most commonly requested types of data, typical classes of data usage and prominent forms of compensation observed in the Swiss industry. The scenario variable "granting of anonymity" was included due to the body of evidence indicating that anonymity is a relevant criterion for the decision to share data in general. Further, the four industry sectors were selected due to their relevant market shares.

TABLE 1 Scenario variables and corresponding attributes as independent variables capturing situational characteristics; IVs = independent variables; levels denoted by "a" are used as reference levels in all regression models

Scenario variables/IVs	Attributes/levels
1. Industry sector	a. Telecommunication
	b. Health insurance
	c. Retail
	d. Banking sector
2. Type of requested data	a. Data on online behavior (e.g., search history)
	b. Geolocation data (e.g., GPS)
	c. Data on social communication (e.g., conversation in a chat-room)
	d. Data on fitness (e.g., number of steps)
	e. Medical data (e.g., heart-rate)
	f. Data on payment behavior (e.g., payment modality)
	g. Data on receipts and expenditures (e.g., bank statement)
	h. Data on purchase behavior (e.g., products, basket)
	i. Data on online media consumption (e.g., online videos, online music)
	j. Data on social media activity (e.g., comments, likes)
3. Intended use of data	a. Improvement of offerings
	b. Improvement of customer insights
	c. Prediction of future purchase behavior
	d. Passing data to third parties
4. Type of compensation for shared data	a. No compensation
	b. Virtual reward points (e.g., in the context of a loyalty program)
	c. Direct financial compensation
	d. Discounts
	e. Donation to charity
5. Granting of anonymity	a. Anonymity not granted
	b. Anonymity granted

Each scenario that was presented to the participants comprised a particular combination of the five scenario variables' attributes. In total, the experiment involved 1600 scenarios as resulting from all possible attribute combinations ($4 \times 10 \times 4 \times 5 \times 2$). To give an example, the scenario consisting of each of the five scenario variables' corresponding first attributes -as indicated in Table 1—confronts subjects with the following situation: "Your telecommunication provider would like to use data on your online behavior in order to improve its offerings. If you are willing to share this type of data with your telecommunication provider, you will receive virtual reward points that you can later exchange for benefits. The data does not allow for personal identification, i.e., you will remain anonymous". After reading a particular scenario's description, a subject had to give a yes-or-no answer to the following question: "Would you be willing to share the requested data with that company under the given specifications?" The answers to this question constitute the dependent variable, namely the subjects' "willingness to share data" (WTS), while the scenario variables constitute the independent variables capturing situational characteristics.

Subjects were recruited via a large Swiss household pool, containing sociodemographic features or related proxy variables (e.g., gender, age, education, political affiliation) of a large proportion of Swiss consumers, and invited to participate in the online experiment by email. The email invitation containing the link to the online survey informed the participants that the experiment was part of a research project on "Digital Media and Consumer Behaviour" and that two vouchers worth 250 Swiss francs for an online shop of their choice would be drawn among all participants. Concretely, the invitation was sent in two waves between December 2016 and February 2017 to more than 300,000 randomly selected adults (18 years or older) in the German and French speaking part of Switzerland. Out of this pool, a total of $N = 20{,}508$ participants completed the online experiment, which was programmed with the survey tool Unipark (www.unipark.com), and hence constitute our sample. To collect subjects' WTS across all 1600 scenarios, we employed an incomplete block design as experimental setup with 160 sets, such that each participant was presented with only 10 scenarios. Further, to minimize participants' cognitive load, we kept the independent variable "intended use of data" constant per participant. The incomplete block design assured that—given these specifications—the independent variables' levels were varied across subjects in a way that allows for evaluating main effects and interaction terms.

Since the dependent variable (WTS) is binary, the data were analyzed using a generalized linear logistic mixed effects model. Scenario variables and demographic information on the participants (age and gender) were included as fixed effects. Since all participants are presented with 10 scenarios, the responses for each participant are not independent. Hence, a random intercept per subject was included to account for unobserved factors that affect a person's general WTS. The model parameters were estimated using the restricted maximum likelihood approach as implemented in glmer() function of the lme4 package for R (Bates et al., 2015).

4 | RESULTS

4.1 | Sample description and descriptive statistics

The gender distribution in the sample showed a proportion of 47.7% (9789) males and 50.7% (10,406) females, while 1.5% (313) of subjects did not indicate their gender or showed a missing value. Participants were at least 18 years old and the median age category spanned between 45 and 49 years. Though the randomly invited participants self-selected into the experiment, the sample was approximately representative of the Swiss population regarding gender (females: 50.7% [sample] vs. 51.6% [population]; males: 47.7% [sample] vs. 48.4% [population]). With respect to age, the sample distribution partially deviates from the population distribution, such that people between 40 and 64 years are overrepresented (55.1% [sample] vs. 43.8% [population]) and people over 65 years are underrepresented (14.5% [sample] vs. 26.8% [population]). The proportion of people between 18 and 39 in the sample approximately matches the population proportion (30.3% [sample] vs. 29.4% [population]).[1]

As for the main variable of interest, considering all 205,080 data points, participants indicated to be willing to share their data in 38,706 out of 166,374 cases, that is, 18.87% of the time.

4.2 | Generalized linear (logistic) mixed effects model

To test our research hypotheses and to explore potential interactions, a generalized linear (logistic) mixed effects model was applied to predict the subjects' decisions to share their private data. The scenario and demographic variables were considered as fixed effects (without interactions) and a random intercept per subject was included as random effect (see Appendix S1). To explore interactions between the scenario variables, two-way interactions between all scenario variables were added to the model. For all factor variables, a dummy coding scheme was used, with the first level—corresponding to the levels denoted by "a" in Table 1—as the reference level. Hence, all coefficients are log odds ratios from the respective levels to the corresponding reference level. Note that the coding scheme does not affect the significance tests for the whole factor, and that all other (log) odds ratios between arbitrary pairs of levels can be calculated from the coefficients.

A version of the model without the random term was also fitted for comparison. However, as expected, the random effect is highly significant as can be seen in Table 2, which shows an ANOVA table indicating a comparison between two models without interaction terms, with one model taking into account the random effects for subjects (2) and the other model (1) not taking into account random effects. The table also shows a comparison of the model with random effects and all second-order interaction terms between scenario variables (3) to the model without interaction terms (2).

To evaluate the statistical significance of fixed effects (i.e., hypothesis tests) and of the interaction effects, a likelihood-ratio test was used. To assess the variables' quantitative impacts, we consider the odds ratios as is natural in logistic regression. All variables are factor variables and we computed the maximum (estimated) odds

[1]The population percentages refer to that part of the Swiss population not including children and adolescents under 18 years at the time when the experiment was conducted.

TABLE 2 ANOVA table indicating a comparison of a model taking into account random effects but no interactions (2) to a model that does not take into account random effects and interactions (1) and a comparison between a model with random effect and interactions (3) and model (2) ($N_{decisions}$ = 205,080)

Model	npar	AIC	BIC	LRT	*df*	Pr(Chi)
1. Model without random effects and without interaction terms	35	179,347	179,704			
2. Model without interaction terms	36	157,473	157,841	21875.7	1	<.001
3. Model with interaction terms	178	153,549	155,368	4207.8	142	<.001

Abbreviations: AIC, Akaike information criterion; BIC, Bayesian information criterion; df, degrees of freedom; LRT, likelihood ratio test; npar, number of parameters; Pr(Chi), p-value.

Types of variables	Factors	# of levels	LRT	Pr(Chi)	OR_max
Scenario variables	Intended use of data	4	368.11	<.001	1.86
	Industry sector	4	1378.62	<.001	1.91
	Type of data	10	9376.43	<.001	9.51
	Type of compensation	5	1017.97	<.001	1.83
	Anonymity	2	4652.68	<.001	2.45
Demographic variables	Gender	2	29.20	<.001	1.15
	Age categories	14	1005.78	<.001	4.62

TABLE 3 Likelihood-ratio-test indicating the statistical significance of fixed effects without interaction terms ($N_{decisions}$ = 205,080, LRT = likelihood ratio test, Pr(Chi) = p-value, OR_max = maximum odds ratio)

ratio between factor levels as a simple measure of effect size. That is, for each factor, we compute the ratio of (a) the odds of a positive WTS-decision for the factor level showing the highest odds to (b) the odds of a positive WTS-decision for the factor level showing the lowest odds. For factors with only two levels, this reduces to the normal odds ratio, with the level showing the higher odds for a positive WTS-decision constituting the numerator. Additional visualizations were used to interpret significant fixed and interaction effects.

4.3 | Results of hypothesis tests: Effects of scenario variables on WTS

Table 3 shows the results of the likelihood-ratio test evaluating the statistical significance of fixed effects. These calculations are based on the model with just main effects and a random effect for the subject, corresponding to Model 2 in Table 2. Detailed regression output including all coefficients is given in Table S1 in Appendix S1. All scenario variables are significant and therefore in support of Hypotheses 1–5. In addition, the demographic variables also had a significant effect on the subjects' WTS. The corresponding maximum odds ratios are indicated as “OR_max” in Table 3.

As for the demographic variables, the model indicates that both gender and age significantly affect WTS-decisions. However, the corresponding maximum odds ratios indicate that the gender effect is rather small while the effect of age is quite pronounced. Concretely, the odds of a positive WTS-decision are estimated to be only 1.15 times higher for males as compared to females, while the odds of a positive WTS-decision are estimated to be 4.62 times higher for people in the lowest age category as compared to people in the highest age category.

Concerning the scenario variables, the value for data type stands out in particular, showing a maximum odds ratio of 9.51 between the factor level yielding the highest odds (i.e., sharing data on payment behavior) and the factor yielding the lowest odds (i.e., sharing data on social communication). In other words, the odds of a positive WTS-decision are estimated to be 9.51 times higher when a company requests data on payment behavior as compared to data on social communication. Thus, among the five scenario variables we have tested, the type of data a company asks for appears to have the largest effect on consumers' WTS in general. However, the other four scenario variables all appear to have a significant and relatively pronounced effect as well. For instance, the odds of a positive WTS-decision are estimated to be 2.45 times higher when the data asked for is anonymous as compared to when it allows for personal identification. Furthermore, as indicated by maximum odds ratios of 1.83, 1.86, and 1.91, consumers' WTS data also depend significantly on the type of compensation offered for the data, what purpose the data are intended to be used for, and what industry sector a corresponding company belongs to, respectively. Figure 1 summarizes these findings by visualizing the predicted relative univariate impact of the scenario variables' factor levels on the probability of a positive WTS-decision in comparison to the baseline (i.e., overall) probability of a positive WTS-decision as indicated by the horizontal line.To summarize, we can reject all five associated null hypotheses, and state that our data support the following five alternative hypotheses. Consumers' WTS personal data with a company depends on:

1. The type of data the company is requesting.
2. The industry sector that the company belongs to.
3. The purpose for which the company will use the data.

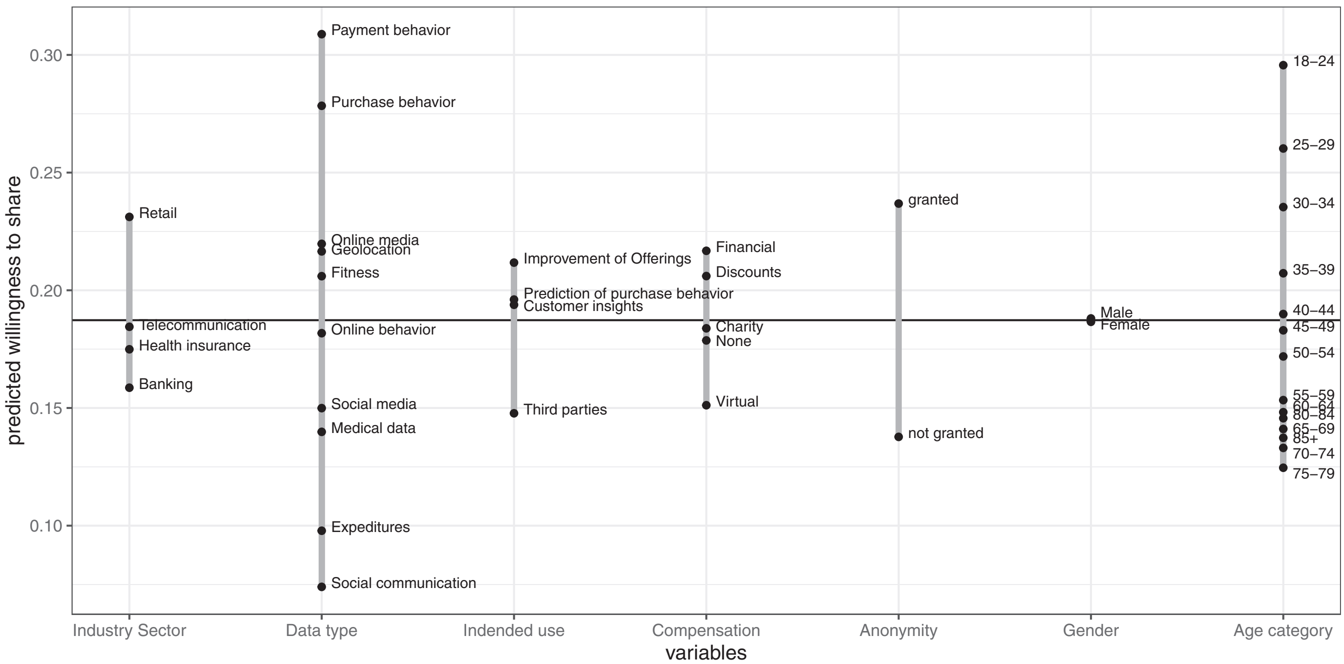

FIGURE 1 Visualization of the scenario variables' effect sizes. The y-axis indicates the proportion of positive WTS-decisions relative to the baseline (horizontal line) as predicted by the model

TABLE 4 Interaction effects between scenario variables as evaluated by a generalized linear mixed effects model ($N_{decisions} = 205{,}080$, df = degrees of freedom, LRT = likelihood ratio test, Pr(Chi) = p-value, OR_max = maximum odds ratio)

Interaction terms	*df*	LRT	Pr(Chi)	OR_max
Intended use × Industry sector	9	33.20	<.001	1.26
Intended use × Data type	27	218.73	<.001	2.22
Intended use × Compensation	12	27.28	<.01	1.36
Intended use × Anonymity	3	88.16	<.001	1.43
Industry sector × Data type	27	3597.22	<.001	11.66
Industry sector × Compensation	12	13.75	.32	1.20
Industry sector × Anonymity	3	13.14	<.01	1.16
Data type × Compensation	36	99.71	<.001	2.12
Data type × Anonymity	9	12.13	.21	1.15
Compensation × Anonymity	4	31.75	<.001	1.27

4. The type of reward provided in return for sharing the data.
5. The degree to which the shared data does allow for personal identification.

4.4 | Results of the interaction effects

Clearly, interactions between scenario variables may matter substantially. For instance, it may make quite a difference whether the company requesting medical data is a health insurer or a bank.

Table 4 shows an overview of the results from the generalized linear mixed effects model analysis incorporating two-way interactions between all scenario variables, corresponding to Model 3 in Table 2. This model offers a significant improvement over the main-effects-only model, and hence is retained as the best model. The full table with parameter estimates per factor level is shown in Appendix S2. While all but two interaction terms are statistically significant, most effects are only small in size and do not show changes in the rank order of the corresponding odds for sharing.

The interaction between industry sector and data type is, however, quite pronounced, and the rank order of the odds for sharing relative to data type is different for each industry sector. As visualized in Figure 2, the data participants are most likely to share with a bank are data on payment behavior, while data on purchase behavior are most likely shared with a retailer, fitness data are most likely shared with a

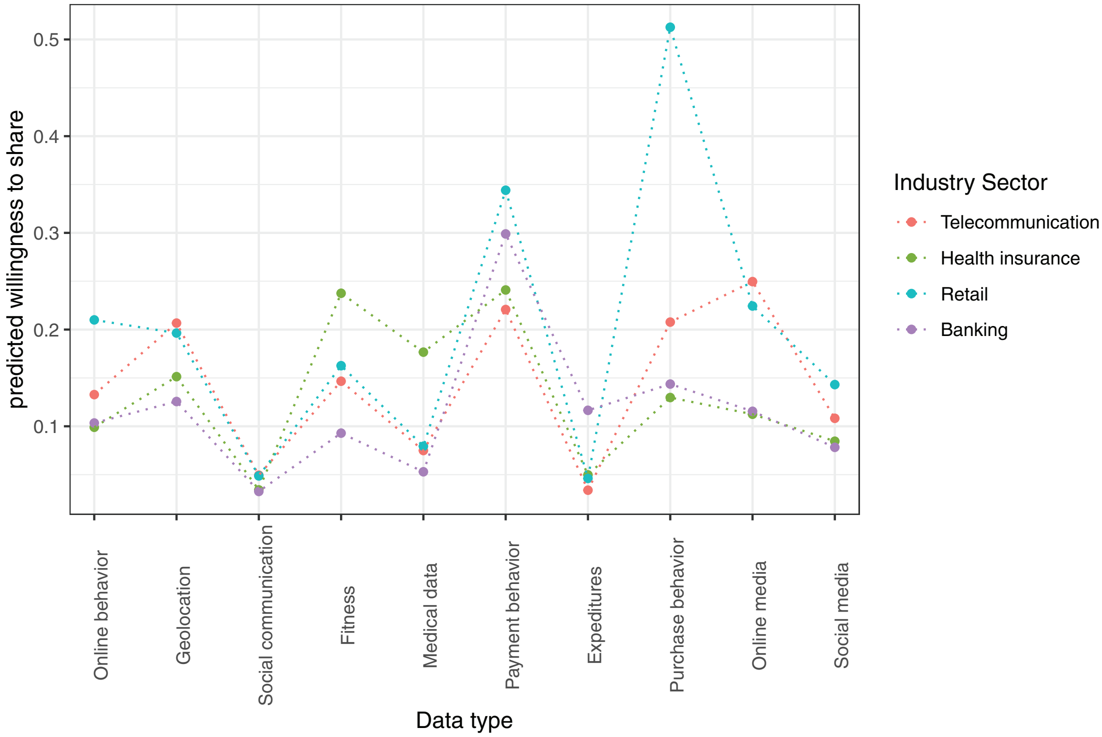

FIGURE 2 Interaction between industry sector and data type. The *y*-axis indicates the proportion of positive WTS-decisions as predicted by the model [Colour figure can be viewed at wileyonlinelibrary.com]

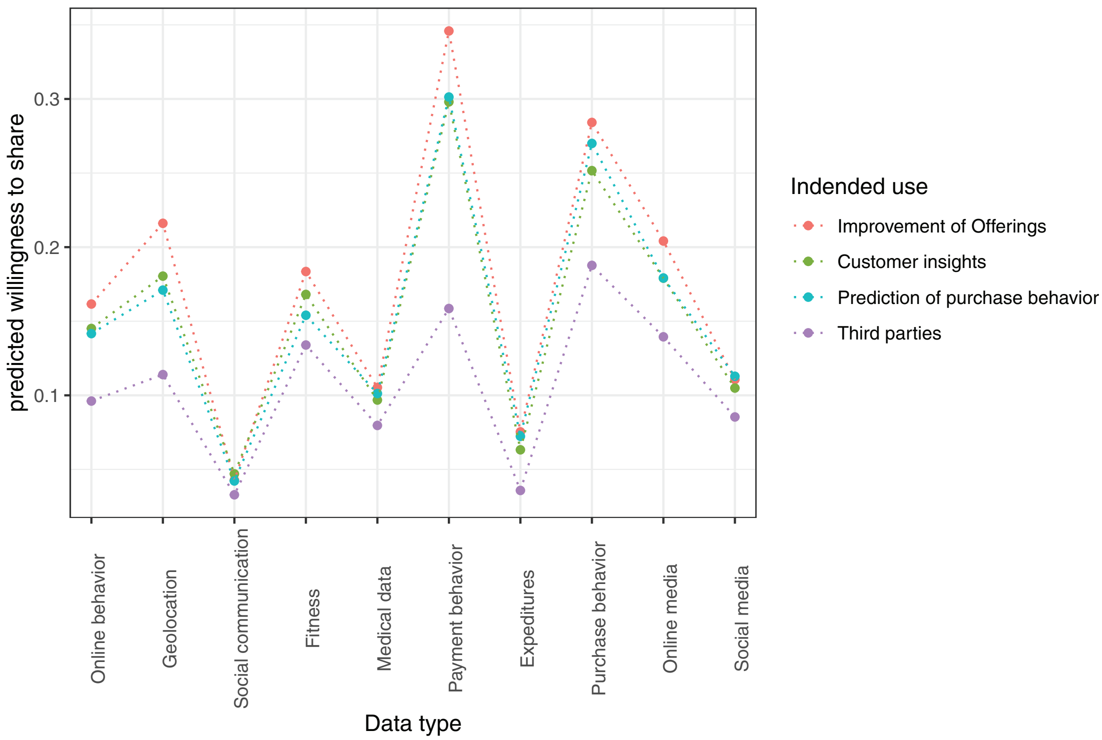

FIGURE 3 Interaction between data type and intended use of data. The *y*-axis indicates the proportion of positive WTS-decisions as predicted by the model [Colour figure can be viewed at wileyonlinelibrary.com]

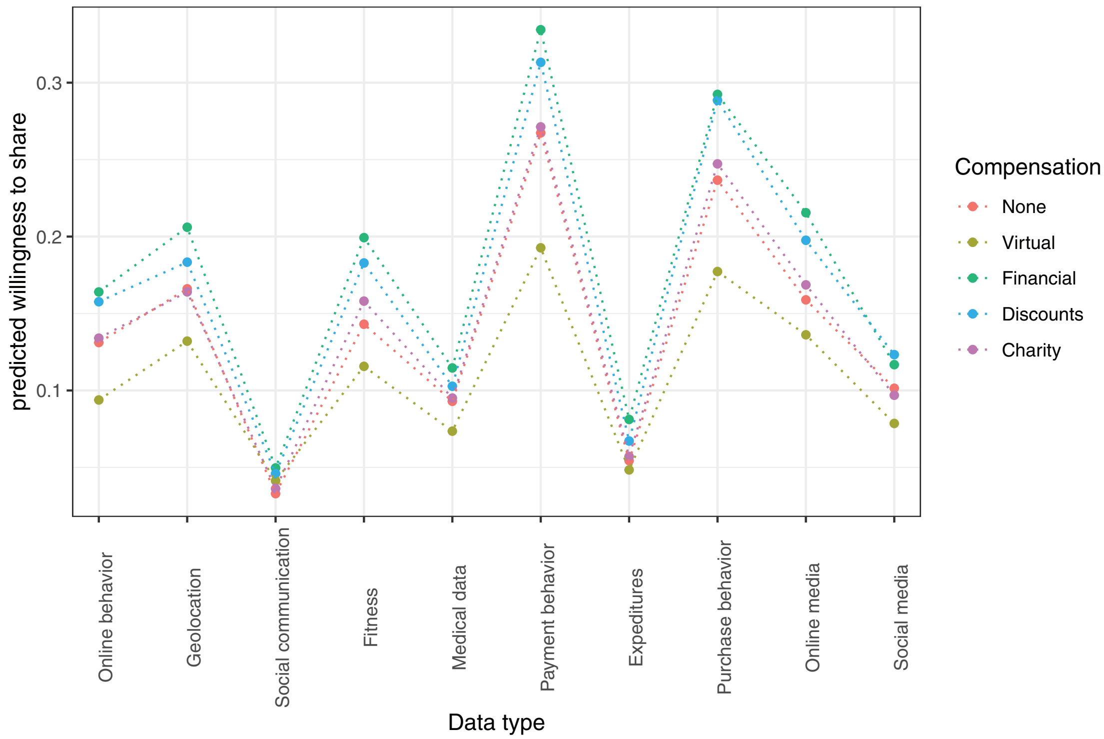

FIGURE 4 Interaction between data type and type of compensation. The *y*-axis indicates the proportion of positive WTS-decisions as predicted by the model [Colour figure can be viewed at wileyonlinelibrary.com]

health insurer and data on online media consumption are most likely shared with a telecommunication provider. As will be discussed later on, it seems that consumers' WTS data with a company is affected by the degree to which there is a match between the type of data requested and the corresponding company's core business.

Two other interaction effects appear noteworthy as they show a maximum odds ratio of more than 2, namely (i) the interaction between data type and intended use of data and (ii) the interaction between data type and type of compensation. However, as can be seen in Figures 3 and 4, these interaction effects are predicated on differences in the relative impact that particular factor levels have on the proportion of positive WTS-decisions, while the corresponding rank orders do hardly change. For instance, consumers are most likely to share data with a company when these data are used for improving offerings and are least likely to share data when these data can be passed on to third parties, irrespective of what type of data is requested. However, on the one extreme, the predicted probability of sharing data on payment behavior differs by more than 15 percentage points depending on whether these data will be used for the improvement of offerings or will be passed to third parties. On the other extreme, the predicted probability of sharing data on social communication varies by less than 5 percentage points depending on what these data will be used for. Analogously, while the type of compensation appears to have a considerable effect on consumers' decisions to share data on their payment behavior, the type of compensation appears to have a negligible effect on consumers' decisions to share data on their social communication. On the aggregate, it appears that the more sensitive a particular type of data is perceived to be, the less impact do other factors have on corresponding WTS-decisions.

5 | DISCUSSION

5.1 | Conclusions

Our data clearly indicate that peoples' WTS personal data with companies depends on contextual characteristics. That is, while participants indicated to be willing to share their data in about 19% of all the cases investigated in this study, the proportions of positive WTS-decisions vary widely across the different corresponding scenarios. For instance, the odds of sharing data on payment behavior are almost 10 times higher than the odds of sharing data on social communication.

Furthermore, situational characteristics do not only affect WTS-decisions in isolation, but moreover do so in combination. That is, our data indicate several interaction effects between situational characteristics on WTS-decisions. As exemplified above, the most pronounced moderation of this kind concerns the interaction between the industry sector a company belongs to and the type of data asked for by the corresponding company. The pattern of results in this respect can most succinctly be summarized by stating that a better intuitive match between the core business a company is engaged in—as indicated by the industry sector it belongs to—and the type of data asked for, results in higher proportions of people willing to share the corresponding data with the corresponding company. For instance, data related to health, such as fitness or medical data, are more likely shared with a health insurer than a retailer, bank, or telecom provider. Analogously, data on purchase behavior are relatively more likely shared with a retailer, data on receipts and expenditures are relatively more likely shared with a bank, and data on online media consumption are relatively more likely shared with a telecom provider as compared to companies belonging to the respective other industry sectors.

A second pattern that the analysis of interaction effects revealed is that the more sensitive a particular type of data is perceived to be, the less impact do other factors have on corresponding WTS-decisions. In other words, consumers will be very unlikely to share private data that they perceive as very sensitive, irrespective of what type of compensation they are offered in return or the degree of anonymity that is granted to them. However, regarding data that is not perceived as very sensitive, other factors, such as what compensation is offered and whether the data allow for personal identification, for instance, will likely have a considerable impact on individual decisions to share these data.

5.2 | Practical implications

Our data indicate that decision-makers in organizations need to take several implications into account if they want to motivate customers and consumers to share their data with them.

First, our results show that people are more likely to share their data with a company when there is an intuitive match between the type of data the company is requesting and the core business that company is engaged in. This suggests that people are more likely willing to share a particular type of data when they feel that they understand why a particular company is requesting that data and correspondingly what the data will be used for. Hence, companies may be well advised to make transparent and explain to their customers why they are requesting a particular type of data to increase their customers' understanding und consequently increase their WTS regarding these data—especially when there is no intuitive or obvious connection between the requested data type and the company's general business model.

Second, the pattern of results suggests that the more sensitive a particular type of data is perceived to be, the less impact do other factors have on corresponding WTS-decisions, and vice versa. This indicates that incentives, monetary, or nonmonetary, will likely not be very effective at motivating customers to share data that they perceive as particularly sensitive. Therefore, especially in such cases, transparency and explication may be the most promising pathways for eliciting WTS. Conversely, incentives are likely effective means for increasing customer's WTS when the data at hand is not perceived as particularly sensitive—even when an intuitive match between data type and core business may be lacking.

Third, our results show that the most effective single factor for evoking WTS is to grant anonymity. Clearly, data that allow personal identification of customers and consumers are likely the most valuable for companies in most cases. Nonetheless, it may be advisable for decision-makers to weigh the benefits of collecting a lot of anonymous data against the benefits of collecting much less data that would allow personal identification.

5.3 | Research implications

Our study design was set up in a way to bolster the external validity of the findings and to allow for clear insights about the causal relationship between contextual factors and WTS to derive practical implications. While existing studies were often based on the survey methodology with cross-sectional designs and homogeneous samples (Kayhan & Davis, 2016; Taylor et al., 2015), we applied an experimental approach with a very large sample. Thus, the study complements a few other studies that applied experimental approaches or conjoint analysis methods to quantify the impact of different factors on privacy-relevant decisions (Acquisti et al., 2013; Roeber et al., 2015). However, while these studies mainly focused on assessing the "value" or "price" of privacy, our study focused on the impact and interplay of contextual factors.

Moreover, our research contributes to research on the "irrational" side of consumers' decisions to share personal information (Acquisti, 2009). The privacy calculus described earlier assumes that consumers weigh potential risks against benefits when deciding to share their personal data. This assumption implies that consumers make a conscious, deliberate decision. However, it seems plausible that consumers, facing high complexity, often do not have sufficient time and mental resources to collect all relevant information and make trade-offs between their need for privacy and other goals (Kim et al., 2015). For instance, research has shown that consumers exhibit inconsistencies in their privacy-related decisions, such that they are more willing to share data in practice than they indicated to be when asked theoretically—a phenomenon termed "privacy paradox" (Gerber et al., 2018; Norberg et al., 2007). Focusing on the role of contextual factors can reveal heuristic strategies that consumers apply when deciding whether to share personal information. For instance, it has been shown that perceptions and decisions are strongly affected by the context or the situation when people are in a heuristic mode of decision-making (John et al., 2009). Even though our methodology only allowed us to elicit stated rather than revealed preferences, the stimuli we confronted participants with in the experiment correspond closely with situations they could indeed face in the wild, we argue. Also, we varied contextual factors across experimental conditions rather than asking participants to rate or evaluate particular factors in isolation, for instance. These methodological features are in favor of our findings' external validity.

However, we would also like to address an important limitation of our research in this respect. Considering well-known phenomena indicating that stated preferences do only match revealed preferences to a limited extent, such as the intention-behavior gap (see e.g., Sheeran, 2001) and the privacy paradox (see, e.g., Norberg et al., 2007), our data likely underestimate people's WTS in the real world. Also, our study was not designed to provide reliable estimates of absolute WTS rates in the population in the first place. However, we argue that the ordinal structure of the patterns we observe in our data likely do have sufficient external validity, such that the *relative* impact of the situational characteristics, as well as the interactions between them, are expected to hold in the field, too. For example, we would argue that people are relatively more likely to share medical data with a health insurer than a bank also in real world settings, even though the corresponding absolute proportions of positive WTS decisions in the field may deviate from the proportions our data indicate.

ACKNOWLEDGMENT

This research has in part been funded by Innosuisse—Schweizerische Agentur für Innovationsförderung (19006.1PFES-ES).

CONFLICT OF INTEREST

The authors have no conflict of interest to declare.

DATA AVAILABILITY STATEMENT

Data available upon request.

ORCID

Kurt Alexander Ackermann https://orcid.org/0000-0003-1156-3880

REFERENCES

Acquisti, A. (2009). Nudging privacy: The behavioral economics of personal information. *IEEE Security Privacy*, *7*(6), 82–85. https://doi.org/10.1109/MSP.2009.163

Acquisti, A., John, L. K., & Loewenstein, G. (2013). What is privacy worth? *The Journal of Legal Studies*, *42*(2), 249–274. https://doi.org/10.1086/671754

Anderson, C. L., & Agarwal, R. (2011). The digitization of healthcare: Boundary risks, emotion, and consumer willingness to disclose personal health information. *Information Systems Research*, *22*(3), 469–490. https://doi.org/10.1287/isre.1100.0335

Bansal, G., Zahedi, F. M., & Gefen, D. (2016). Do context and personality matter? Trust and privacy concerns in disclosing private information online. *Information & Management*, *53*(1), 1–21. https://doi.org/10.1016/j.im.2015.08.001

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bearth, A., & Siegrist, M. (2020). Psychological factors that determine people's willingness-to-share genetic data for research. *Clinical Genetics*, *97*, 483–491. https://doi.org/10.1111/cge.13686

Beke, F. T., Eggers, F., & Verhoef, P. C. (2018). Consumer informational privacy: Current knowledge and research directions. *Foundations and Trends in Marketing*, *11*(1), 1–71. https://doi.org/10.1561/1700000057

Brandimarte, L., Acquisti, A., & Loewenstein, G. (2013). Misplaced confidences privacy and the control paradox. *Social Psychological and Personality Science*, *4*(3), 340–347.

Cumbley, R., & Church, P. (2013). Is "Big Data" creepy? *Computer Law & Security Review*, *29*(5), 601–609. https://doi.org/10.1016/j.clsr.2013.07.007

Dinev, T., & Hart, P. (2006). An extended privacy calculus model for e-commerce transactions. *Information Systems Research*, *17*(1), 61–80. https://doi.org/10.1287/isre.1060.0080

European Commission. (2015). Special Eurobarometer 431: Data protection—ecodp.common.ckan.site_title. https://data.europa.eu/euodp/en/data/dataset/S2075_83_1_431_ENG

General Data Protection Regulation. (2016). OJ L 119, 4.5.2016, p. 1–88 (BG, ES, CS, DA, DE, ET, EL, EN, FR, GA, HR, IT, LV, LT, HU, MT, NL, PL, PT, RO, SK, SL, FI, SV). https://eur-lex.europa.eu/eli/reg/2016/679/oj

Gabisch, J. A., & Milne, G. R. (2014). The impact of compensation on information ownership and privacy control. *Journal of Consumer Marketing*, *31*, 13–26. https://doi.org/10.1108/JCM-10-2013-0737

GDPR.EU. (2019). 59,000 breaches reported in first eight months of new GPDR requirements. GDPR.Eu. https://gdpr.eu/gdpr-requirements-data-breach-reporting/

Gerber, N., Gerber, P., & Volkamer, M. (2018). Explaining the privacy paradox: A systematic review of literature investigating privacy attitude and behavior. *Computers & Security*, *77*, 226–261. https://doi.org/10.1016/j.cose.2018.04.002

Hoffman, D. L., Novak, T. P., & Peralta, M. A. (1999). Information privacy in the marketspace: Implications for the commercial uses of anonymity on the Web. *The Information Society*, *15*(2), 129–139.

Jai, T.-M., & King, N. J. (2016). Privacy versus reward: Do loyalty programs increase consumers' willingness to share personal information with third-party advertisers and data brokers? *Journal of Retailing and Consumer Services*, *28*, 296–303. https://doi.org/10.1016/j.jretconser.2015.01.005

John, L. K., Acquisti, A., & Loewenstein, G. (2009). The best of strangers: context dependent willingness to divulge personal information (SSRN Scholarly Paper ID 1430482). Social Science Research Network. https://papers.ssrn.com/abstract=1430482

Kayhan, V. O., & Davis, C. J. (2016). Situational privacy concerns and antecedent factors. *Journal of Computer Information Systems*, *56*(3), 228–237. https://doi.org/10.1080/08874417.2016.1153913

Kim, M., Ly, K., & Soman, D. (2015). A behavioural lens on consumer privacy (Behavioural Economics in Action Research Report Series). Toronto: Rotman School of Management, University of Toronto. https://inside.rotman.utoronto.ca/behaviouraleconomicsinaction/files/2013/09/ConsumerPrivacy-BEAR-2015-Final.pdf

Kumaraguru, P., & Cranor, L. F. (2005). Privacy indexes: A survey of Westin's studies. School of Computer Science, Carnegie Mellon University, Pittsburgh. https://doi.org/10.1184/R1/6625406.v1

Li, H., Sarathy, R., & Xu, H. (2010). Understanding situational online information disclosure as a privacy calculus. *Journal of Computer Information Systems*, *51*(1), 62–71. https://doi.org/10.1080/08874417.2010.11645450

Li, Y. (2011). Empirical studies on online information privacy concerns: Literature review and an integrative framework. *Communications of the Association for Information Systems*, *28*(1), 28. https://doi.org/10.17705/1CAIS.02828

Lim, S., Woo, J., Lee, J., & Huh, S.-Y. (2018). Consumer valuation of personal information in the age of big data. *Journal of the Association for Information Science and Technology*, *69*(1), 60–71. https://doi.org/10.1002/asi.23915

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

Marwick, A., & Hargittai, E. (2019). Nothing to hide, nothing to lose? Incentives and disincentives to sharing information with institutions online. *Information, Communication & Society*, *22*(12), 1697–1713. https://doi.org/10.1080/1369118X.2018.1450432

Milne, G. R., Pettinico, G., Hajjat, F. M., & Markos, E. (2017). Information sensitivity typology: Mapping the degree and type of risk consumers perceive in personal data sharing. *Journal of Consumer Affairs*, *51*(1), 133–161. https://doi.org/10.1111/joca.12111

Norberg, P., Horne, D. R., & Horne, D. A. (2007). The privacy paradox: Personal information disclosure intentions versus behaviors. *The Journal of Consumer Affairs*, *41*(1), 100–126.

Phelps, J., Nowak, G., & Ferrell, E. (2000). Privacy concerns and consumer willingness to provide personal information. *Journal of Public Policy & Marketing*, *19*(1), 27–41. https://doi.org/10.1509/jppm.19.1.27.16941

PricewaterhouseCoopers. (2017). Consumer intelligence series: Protect. me. PwC. https://www.pwc.com/us/en/services/consulting/library/consumer-intelligence-series/cybersecurity-protect-me.html

Roeber, B., Rehse, O., Knorrek, R., & Thomsen, B. (2015). Personal data: How context shapes consumers' data sharing with organizations from various sectors. *Electronic Markets*, *25*(2), 95–108. https://doi.org/10.1007/s12525-015-0183-0

Sheeran, P. (2001). Intention–behavior relations: A conceptual and empirical review. In *European review of social psychology* (pp. 1–36). John Wiley & Sons.

Silverio-Fernández, M., Renukappa, S., & Suresh, S. (2018). What is a smart device?—A conceptualisation within the paradigm of the internet of things. *Visualization in Engineering*, *6*(1), 3. https://doi.org/10.1186/s40327-018-0063-8

Smith, H. J., Milberg, S. J., & Burke, S. J. (1996). Information privacy: Measuring individuals' concerns about organizational practices. *MIS Quarterly*, *20*(2), 167–196. https://doi.org/10.2307/249477

Taylor, J. F., Ferguson, J., & Ellen, P. S. (2015). From trait to state: Understanding privacy concerns. *Journal of Consumer Marketing*, *32*(2), 99–112. https://doi.org/10.1108/JCM-07-2014-1078

Treiblmaier, H., & Pollach, I. (2007). Users' perceptions of benefits and costs of personalization. Proceedings of the Twenty Eighth International Conference on Information Systems, Montreal, 141.

Xia, F., Yang, L. T., Wang, L., & Vinel, A. (2012). Internet of things. *International Journal of Communication Systems*, *25*, 1101–1102.

Yun, H., Lee, G., & Kim, D. J. (2019). A chronological review of empirical research on personal information privacy concerns: An analysis of contexts and research constructs. *Information & Management*, *56*(4), 570–601. https://doi.org/10.1016/j.im.2018.10.001

Ziegeldorf, J. H., Morchon, O. G., & Wehrle, K. (2014). Privacy in the internet of things: Threats and challenges. *Security and Communication Networks*, *7*(12), 2728–2742. https://doi.org/10.1002/sec.795

AUTHOR BIOGRAPHIES

Kurt Alexander Ackermann is a Senior Lecturer in Marketing with a Specialization in Consumer Psychology and Behavioral Economics at the Centre for Behavioural Marketing at the Institute of Marketing Management within the School of Management and Law at Zurich University of Applied Sciences (ZHAW) in Switzerland. He earned his master's degree in Psychology at the University of Zurich (UZH) and his doctoral degree in Decision Theory and Behavioral Game Theory at the Federal Institute of Technology (ETH) in Zurich, Switzerland. His research is interdisciplinary and involves topics relating to consumer psychology, behavioral economics, and computer-human interaction, for instance.

Linda Burkhalter is a Senior Lecturer in Consumer Behavior at the Institute of Natural Resource Sciences at the Zurich University of Applied Sciences in Wädenswil, Switzerland. She received a diploma in psychology from Humboldt University of Berlin in 2006 and completed her PhD in business administration/marketing at the University of St. Gallen in 2011. Her research areas are sustainable consumer decision-making, nudging, behavioral change, and the impact of digitalization on consumers and climate.

Thoralf Mildenberger is a Senior Lecturer in Computational Statistics at Zurich University of Applied Sciences (ZHAW) in Winterthur, Switzerland. He earned his PhD in Statistics in 2011 from TU Dortmund University for a thesis on nonparametric curve estimation. His research focuses on various topics in statistics, machine learning, and data science with applications to data from different fields including social sciences, business analytics, traffic and transport, and manufacturing.

Martin Frey is a Lecturer in Statistical Data Analysis at Zurich University of Applied Sciences (ZHAW) in Winterthur, Switzerland. He earned his PhD in Environmental Sciences in 2009 and in 2011 a Postgraduate Degree in Applied Statistics both from ETH Zurich. His research focuses on data analysis and machine learning with applications to different fields including engineering, environmental science, sports analytics, business analytics, traffic, tracking data, and predictive maintenance.

Angela Bearth is a Senior Researcher at the Consumer Behavior Group at ETH Zurich. Her research focuses on the public's perceptions and acceptance of innovative technologies. She has studied Social and Health Psychology at the University of Zurich and has earned her PhD in Social Science at ETH Zurich.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Ackermann, K. A., Burkhalter, L., Mildenberger, T., Frey, M., & Bearth, A. (2021). Willingness to share data: Contextual determinants of consumers' decisions to share private data with companies. *Journal of Consumer Behaviour*, 1–12. https://doi.org/10.1002/cb.2012