Who do you trust: Peers or Technology?
A conjoint analysis of computational reputation mechanisms

Gunnar Stevens
University of Siegen / Bonn-Rhein-Sieg University of Applied Science
gunnar.stevens@uni-siegen.de / gunnar.stevens@h-brs.de

Paul Bossauer
University of Siegen / Bonn-Rhein-Sieg University of Applied Science
paul.bossauer@uni-siegen.de / paul.bossauer@h-brs.de

Abstract. Peer-to-peer sharing platforms are taking an increasingly key role in the platform economy due to their sustainable business model. By sharing private goods and services, the challenge arises to build trust between peers online, primarily without a physical presence. Peer rating has been proven to be an important mechanism. In this paper, we explore the concept called trust score, a computational rating mechanism adopted from car telematics, which can play a similar role in carsharing. For this purpose, we conducted a conjoint analysis in which 77 car owners chose between fictional user profiles. Our results reveal that, in our experiment, the telemetric-based score slightly outperforms the peer rating in the decision process, and the participants perceived the peer rating to be more helpful in retrospect. Further, we discuss potential benefits regarding existing shortcomings of user ratings and various concerns that should be considered in concepts such as the telemetric-based reputation mechanism that supplements existing trust factors, such as user ratings.
Introduction

Consumption behavior has changed in recent years, and consumers have increasingly given up individual ownership for demand-based access to goods. Accordingly, business models in which consumers simply pay for access to services or goods are rising (Wilhelms, Henkel, and Merfeld 2017). Benefiting from this development in society and from the commitment to a more sustainable business model, carsharing has gained public attention for several years (Münzel et al. 2018). In Germany, the number of carsharing customers rose to 2.46 million in the last year. In addition to car rental and commercial carsharing (station-based and free-floating), platforms, such as Turo, Getaround, or Getaway, allow private car owners to lend their cars to previously unknown individuals for an agreed period and price (Wilhelms, Henkel, and Merfeld 2017).

The Internet allows most transaction steps to be carried out online, which lowers the transaction costs but makes the collaboration more anonymous. Lack of trust in other “buyers” or “sellers” is one of the most frequent reasons for rejecting peer-to-peer (P2P) sharing platforms (Owyang, Tran, and Silva 2013). As a result, transparency, reputation, and trust are considered essential for success in the P2P sharing economy, which is also often called the share economy or shareconomy (R. Belk 2007; Botsman and Rogers 2011; Hawlitschek, Teubner, and Weinhardt 2016).

Sharing platforms typically use various, complementary trust-building mechanisms, such as showing self-provided information (e.g., pictures or personal information) and peer-provided information (e.g., consumer reviews and ratings) (Bente, Baptist, and Leuschner 2012). The Internet of Things (IoT) could serve as a third trust-building mechanism, using computationally provided information (Stevens et al. 2018). In P2P carsharing, for instance, the driving behavior of a car renter could be monitored. In principle, this information could be used by car owners to evaluate the driving behavior of a person who is currently renting their car.

So far, such information is used only by car insurance companies (Merzinger and Ulbrich, n.d.) and not by reputation systems for carsharing platforms. Hence, it is unclear whether users would trust such IoT-based ratings. This leads to the question regarding whether a computational rating has the same relevance as peer ratings, which are currently viewed as the gold standard among various reputation mechanisms in the sharing economy (Hawlitschek, Teubner, and Weinhardt 2016; Teubner et al. 2016).

To answer this question, we conducted a conjoint analysis in which people had to choose between two fictional profiles of people who sent a request to rent a car. The profiles included information about the peer rating, number of ratings, telematic-based driving scoring, and how many kilometers the person has driven with cars from the platform.
The major finding was that telematic-based scoring not only had a significant influence on the selection behavior, but the effect was slightly higher than the corresponding peer rating. More interestingly, at the end of the online survey, we also asked the participants about the relevance of the provided information. In this case, the peer rating was perceived as the most helpful and slightly better than the telematic-based scoring. This indicates an action-perception gap, in which we trust more in technology in making our decisions, whereas we trust more in other people in reflection.

Related Work

Carsharing

The sharing economy describes an “economic model enabled by modern ICT [information and communications technology], based on the sharing of digital content, physical goods, or the participation in commercial, cultural or social projects to access underutilized assets for monetary or non-monetary benefits” (Richter, Kraus, and Syrjä 2015). The basic idea is a joint consumption in the sense of “using instead of owning” (Pakusch et al. 2016). Products are not acquired by the consumer. Instead, they only obtain the temporary right to use the service or good, normally for a certain fee (R. Belk 2007). Among others, economic motivations, such as cost-savings, reducing the burden of ownership, or increasing access to resources, play an important role (Hamari, Sjöklint, and Ukkonen 2016). Digitalization further lowers both search and transaction costs, making it economically attractive to monetize under-used resources. A comparably old example is carsharing. It provides customers with on-demand access to any vehicle without having to buy and maintain it themselves. Payment is made only for the respective period of use and/or the driven distance (Witzke 2016).

Over time, different forms of carsharing have emerged. For services of classic station-based providers, members can pick up a car at a certain point after it has been booked, use it for a certain period and then return it to the dedicated parking space (Witzke 2016). With so-called fully flexible (station-independent) free-floating carsharing systems, vehicles are distributed arbitrarily over a corresponding business area in public parking spaces. In contrast to the previously described system, no fixed carsharing stations exist for these vehicles. Members can use available cars spontaneously without a prior reservation and then park the cars elsewhere in the business area (Witzke 2016). Both station-based and free-floating carsharing vehicles are owned and provided by commercial providers. Hence, these two types of commercial carsharing have much in common with commercial car rental.
In this paper, we focus on P2P carsharing in which car owners lend their private vehicles for a short period (Hampshire and Gaites 2011). According to a study with over 10,000 participants by the Ford Motor Company (Ford Motor Company 2016), 48% of respondents in Germany would lend their car for a fee, but few studies have investigated using P2P carsharing (e.g., Wilhelms, Henkel, and Merfeld 2017; Lewis and Simmons 2012; Ballús-Armet et al. 2014; Nobis 2006). The typical P2P carsharing users are young, predominantly male, well educated, and urban and have a good job (Lewis and Simmons 2012). Various motivations exist for car sharing, but economic reasons (e.g., reducing mobility and vehicle costs) and situational-practical reasons (e.g., availability, convenience, and flexibility) are often mentioned (Ballús-Armet et al. 2014; Nobis 2006). Wilhelms et al. (2017) found that car owners want to reduce their vehicle costs and earn small additional amounts of money that they can spend on other purposes. They also like to enable other people without a vehicle to travel. However, compared to the Ford study, a large gap exists between the intention to lend a private car for a fee and the actual behavior.

One barrier is the practicability of carsharing offers. For instance, lending a private car via a P2P platform includes the effort of entering the availability of the car, arranging handover dates and a follow-up check to determine whether the car has been damaged. In addition, a general fear of sharing a good with strangers exists (Wilhelms, Henkel, and Merfeld 2017; Shaheen and Cohen 2013). In particular, people often have a personal and emotional bond with their cars (Gatersleben 2007), which increases the fear of loss (R. W. Belk 1988) because others might not treat the rented car with care or might have an accident or return it late or dirty.

Trust

[Trust is] the willingness of a party to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to the trustor, irrespective of the agility to monitor or control that other party. (Mayer, Davis, and Schoorman 1995)

Trust is of particular importance in potentially risky and uncertain transactions where parties are interdependent (McKnight and Chervany 2001). Such situations are typical in the sharing context, as the key steps of P2P transactions are conducted online. Therefore, transparency, reputation, and trust are considered essential (R. Belk 2007; Botsman and Rogers 2011; Hawlitschek, Teubner, and Weinhardt 2016). This creates a paradoxical situation in which trust is essential because it is an efficient method of reducing transaction costs in such social exchanges (Huurne et al. 2017). However, building and sustaining trust online is more difficult because common trust factors, such as physical interaction or personal knowledge, do not exist (Hawlitschek, Teubner, and Weinhardt 2016; Möhlmann 2015).

As a research topic, trust has been studied as a multifaceted topic and complex phenomenon in many disciplines since the 1950s (Corritore, Kracher, and

On the individual level, the question arises regarding how trust forms and influences actions and decision-making processes. Due to the openness of the future, a decision can always turn out to be wrong in retrospect, even if it seemed to be the right one in the situation. Hence, the outcome of a situation is always uncertain and can never be guaranteed; therefore, we must trust and hope for the best. In developmental psychology, Erikson (1993) stressed the human’s ability to develop a sense of basic trust rather than distrust during infancy.

This basic trust refers to generalized trust as a learned trust behavior, which a person has built through numerous experiences during their lives and which has been condensed into a stable, long-term disposition. Generalized trust must be distinguished from particularized and situated trust, which is shaped by the situation and particular circumstances (Bjørnskov 2007). When deciding whether and how much to trust, people search for cues (e.g., trustworthiness attributes) helping them to interpret the situation (M. K. Lee and Turban 2001). The generalized trust of a person magnifies or reduces the signals that the cues provide and vice versa. People rely on previous experiences that contribute to how they interpret the individual cues and the whole situation in terms of trust. In addition, establishing trust is an interactive process based on positive feedback that reinforces the initial trustworthy behavior, so that trust gradually increases (Treek 2017). Hence, the outcome of the situation contributes to whether we link trust or distrust to such situations and its actors in the future.

Trusting a situation always means interpreting the situation. From this stance, we reformulate Blumers’ (1969) well-known premises of symbolic interactionism as follows:

1) Humans act toward things on the basis of the meanings resp. trust they ascribe to those things.
2) The meaning resp. trustworthiness of such things is derived from, or arises out of, the social interaction that one has with others and the society.
3) The meaning resp. trustworthiness of such things is handled in, and modified through, an interpretative process used by the person in dealing with the things he/she encounters.

The first point refers to the fact that things (physical objects, actions, concepts, etc) have both personal meaning and personal trustworthiness. The second point refers to the fact that this trustworthiness is derived from our previous experiences with the physical objects, actions, and concepts and our social interactions that provide additional information about the physical objects, actions, and concepts.

On a social level, one can assess the role of trust in the functioning of societies and economies. For instance, Luhmann (2018) stressed that trust presents an important resource for societies because it provides a means to decrease complexity
by reducing the number of options to consider in a given situation. The promotion of trust is therefore a duty for the functioning of society (Corritore, Kracher, and Wiedenbeck 2003). In economics, this view is reflected in the principal-agent theory (Sappington 1991; Ensminger 2001) and the transaction cost theory (Schneider 1987; Williamson 1979). Generally, trust can reduce control costs and transaction costs, thus contributing to economic prosperity. Trust has been studied in various areas, particularly concerning social media (Ridings, Gefen, and Arinze 2002), online privacy (Jakobi et al. 2019; Dinev and Hart 2006), online shopping (Gefen 2002), and the sharing economy (Hawlitschek et al. 2016; Hawlitschek, Teubner, and Weinhardt 2016). The studies reveal that trust has a positive effect on exchange relationships, which is expressed by the willingness to participate, disclose information, make a transaction, and so on.

**Sectoral trust.** In principle, general and situated trust are not a dichotomous category but a continuum. Moreover, people have a kind of basic trust that relates to different areas. This kind of trust can also be understood as sectoral trust, which mediates the particular situation and generalized trust. As a result, when dealing with a situation, person, organization, and so on, we use trust at different levels. Sectoral trust is more stable than situated trust but less stable than generalized trust.

Sectoral trust can be different in varying areas. For example, it may be that someone has a low level of trust in an individual but a high level of trust in the government as a whole (or vice versa). In the P2P sharing economy, products or services are usually offered by private individuals. Thus, users must trust other peers, the platform, and the offered products and services (Hawlitschek et al. 2016; Hawlitschek, Teubner, and Weinhardt 2016). In addition, using algorithmic reputation mechanisms, three sectoral trust areas are of special interest: *trust in people, trust in organizations, and trust in technology.*

**Interpersonal trust.** Trust in other people is probably one of the most elementary forms of trust, and some authors equate it to generalized trust. For instance, Erikson’s (1993) developmental psychological considerations about basic trust focus on interpersonal relationships with parents and people from the immediate environment. Rotter (1967) defined interpersonal trust as the generalized expectancy that the verbal statements of others can be relied upon. Various scales have been suggested in the literature for measuring interpersonal trust. Interpersonal trust depends on several factors, such as perceived competence, dependability, benevolence, empathy, and familiarity (Evans and Revelle 2008; Rotter 1967; Beierlein et al. 2012).

In the P2P sharing economy, interpersonal trust comes into play in two ways: First, one must trust the person who asks for a good or the person who offers a good. Second, interpersonal trust plays a key role in user ratings. If a person does not trust others, this person would also not trust the user ratings. While the first point is essential for a sharing economy to work, we focus on the second point in this paper. In our study, we adopt the short scale by Beierlein et al. (2012) to
measure the participants’ sectorial trust in other people and their peer-rating competence and benevolence.

**Organizational trust.** People have a certain amount of trust in companies and organizations when they buy products or use a service. In the case of online shopping, trust covers two dimensions: trust in the seller and trust in the goods offered. In particular, buyers must have a certain level of confidence in the seller’s integrity, benevolence, and competence (Gefen 2002; Bart et al. 2005; Urban, Amyx, and Lorenzon 2009). Competence refers to the ability of the vendor to fulfill promises (e.g., deliver a product on time). Integrity is shown by the company acting consistently, reliably, and honestly (e.g., if a fair cancellation is emphasized in advertising, the customer can simply return goods if he or she is not satisfied). Benevolence means that a seller puts customer interests above its own and displays a sincere concern for the well-being of the customers (e.g., by making the customer aware of better products, even if the profit margin is smaller; (Chen and Dhillon 2003). The range of application of this basic trust in the integrity, benevolence, and competence can vary. It can apply to the entire economy, a specific sector, or a particular company.

The role of trust in providers of a P2P sharing platform was investigated in various studies. In particular, Hawlitschek et al. (2016) pinpointed that this trust model can also apply to the sharing economy. Despite its importance, however, this trust propensity is out of the scope of our research, as we focus on the effects of reputation systems and not the effect providers have.

**Technology trust.** Systems must be trusted to be used. To be trustworthy, it is important that the system is reliable and meets user expectations. Donick (2019) referred to the fear of flying as a paradigmatic case of technology (dis)trust: *This fear often has nothing to do with high altitude - a lot of people with high anxiety have no problem looking out of the window of an airplane. Fear of flying is rather due to the unfamiliar technology that you rarely have to deal with and that you have to expose yourself to for several hours.*

The reverse case of faith in technology also exists, where people trust something because it is the result of scientific or technical procedures that do not suffer from human weaknesses, such as prejudices, attention deficits, limited processing capacity, and so on. For example, in the multiplication of large numbers, one might use a calculator because this technology is more trustworthy than mental arithmetic.

Regarding system design, Lee and See (2004) also emphasized the danger of both under-trusting and over-trusting. If users place too little trust in technology, the system capabilities are neglected, leading to disuse. If users place too much trust in technology, users could become gullible, perform risky actions, or be tempted toward other forms of misuse. Therefore, trust must be calibrated so that trust matches the system capabilities, leading to appropriate use. Interactions such as obtaining feedback and experiences are helpful to calibrate trust.
Due to the complexity, technology trust has recently been researched, especially regarding artificial intelligence and autonomous systems. Concerning human-automation interaction in general, Lee and See (2004) emphasized that trust depends on the performance, process, or purpose of an automated system. Palmer et al. (2016) outlined a list of dimensions needed for creating system trust in autonomous systems:

- such as perceived competence referring to the belief that the system can perform the task in question,
- benevolence referring to the belief that the system supports the mission,
- understandability referring to the belief that the conclusions that a system reaches can be understood,
- reliability referring to the belief that the system has only a small chance of failing during a mission, and
- the false-alarm rate referring to the fact that certain error rates are known and acceptable.

Davis (2020) suggested a scale measuring the propensity to trust autonomous systems. The scale consists of four subscales:

- capability measuring the general trustworthiness of autonomous systems using the capability of humans as a baseline for comparison,
- legitimacy measuring a person’s perception that autonomous systems make legitimate significant decisions independent of human direction,
- collaboration addressing that decision-making should integrate both humans and autonomous systems, and
- transparency addressing the transparency of autonomous systems and how easy it is for a person to judge whether an autonomous system is trustworthy.

To the best of our knowledge, no study measures consumer trust in telematic technologies, especially concerning P2P carsharing. Hence, we adopt general considerations about trusting autonomous systems to measure the participants’ sectorial trust in technology.

Reputation Systems

As an important substitute for trust mechanisms in the offline world, reputation systems play an important role in the online world (Huurne et al. 2017; Ert, Fleischer, and Magen 2016). The basic idea is that the parties rate each other (e.g., after concluding a transaction) and derive a trust or reputation score from the aggregated ratings of a particular party that can help other parties decide whether
to interact with that party in the future. Reputation systems also create an incentive for good behavior and therefore tend to have a positive effect on market quality. They could also be described as collaborative sanctioning systems to reflect their collaborative nature. Reputation systems are related to collaborative filtering systems (Schafer, Konstan, and Riedl 1999) because the systems use the opinions of a community of users or customers to help individuals more effectively identify relevant content from a potentially overwhelming set of choices. In their seminal work, Resnick and Zeckhauser (2002) provided a functional definition of a reputation system; it must do the following:

1. provide information that allows peers to distinguish between trustworthy and untrustworthy peers,
2. encourage peers to be trustworthy, and
3. discourage participation from those who are not trustworthy.

A vast amount of information exists that is consciously and unconsciously used by people as indicators regarding whether someone is trustworthy, such as appearance, behavior, and what others say about someone. Here, the trustworthiness of the source from which the information originates plays an important role. People typically use information from different sources to come to a decision. If the information from different sources points in the same direction, the information is usually reinforced. However, if the information is contradictory, the information is weighted, where a weighted mean is formed considering the trustworthiness of the information and its source.

In the literature on the P2P sharing economy, various sources of information are mentioned that contribute to building trust. We classify them in four categories; yet, they are neither exhaustive nor mutually exclusive.

**Self-provided information.** Peers could provide information about themselves and the goods or services they offer. For instance, Airbnb, eBay, and BlaBlaCar allow their users to create a profile adding personal attributes and describing their products and services in their own words. Repschläger et al. (2015) pinpointed that personal attributes, such as name, age, address, and uploaded pictures serve as trust factors. Concerning the service description, Teubner et al. (2016), through the analysis of Airbnb offers, showed that the number of photos, for example, has a positive effect on the willingness to pay. However, compared to peer ratings, the effect is comparably low.

**Platform-provided information.** Self-provided information is always suspected of being embellished. This trust dilemma can be partially resolved by other external sources, especially by the platform provider as an intermediary. For instance, some P2P platforms have procedures in place to verify the authenticity of the information, such as name, age, and driver’s license (Repschläger et al. 2015). However, these procedures contribute to building trust only if the procedure itself and the platform providers are trustworthy. A kind of spillover effect exists in which the reputation of the platform provider is passed on to the platform user.
Frauds, for instance, take advantage of this spillover effect by setting up fake shops on eBay, Walmart, or the Amazon marketplace to profit from the popularity of and trust in the platform (Doyle 2017).

**Peer-provided information.** Another important source is peers, where peer and consumer ratings and reviews serve as a proxy for word of mouth recommendations (Zhu and Zhang 2010). Peer reviews are usually written to either recommend the product to others or to warn others to stay away from the product (Hennig-Thurau, Walsh, and Walsh 2003; Sen and Lerman 2007). Users tend to write reviews mostly for products that they perceive to be exceptionally good or exceptionally bad (Dellarocas and Narayan 2006). Generally, consumers read reviews to evaluate products and reduce the risk of making a wrong purchase decision (King, Racherla, and Bush 2014). Online reviews influence the willingness to pay and often serve to build trust and customer loyalty (King, Racherla, and Bush 2014).

Customer review management is a common activity of online retailers (Zhu and Zhang 2010). The positive effects of online reviews are also accompanied by the risk of purchased reviews, where products seem to be better than they are or where competitors are devalued with purchased poor ratings (Zhu and Zhang 2010). Consumers typically pay more attention to negative reviews than positive reviews. Further, the depth, length, and quality of reviews (Mudambi and Schuff 2010) and the number of reviews and ratings (Li and Hitt 2008) have an effect. In particular, in the early stages of product introduction, when only a few consumer reviews and ratings existed, the recommendations seem more susceptible to bias (Li and Hitt 2008). Teubner et al. (2016) uncovered that the average rating score has the most influence on the price. Compared to this, the review counts have less influence (and counterintuitively, a negative one).

**Technology-provided information.** This category includes all information that is not provided by the entity or the peers but is automatically collected by devices or sensors. Users consider such information to help identify trustworthy peers. This includes information collected by the platform regarding user activity or membership duration. Such information can serve as an indicator of trustworthiness (Repschläger et al. 2015; Vanderveld et al. 2016). The benefit of the computationally provided information is that the information is not collected manually; therefore, no additional effort is necessary for the user. Second, the information is based on objective measures and is more difficult to manipulate than self- and peer-provided information (or at least tampering requires additional effort). Despite these benefits, the current reputation systems make little use of such information.

**Using Telematic Technology to Build Trust**

Examining other domains, we find a more detailed example in car telematics, which are based on computational scoring using IoT-based information to build
trust (Ma et al. 2018). In the literature, we find various terms, such as telematic traffic (Ma et al. 2018), usage-based insurance (Soleymanian, Weinberg, and Zhu 2019), or pay-how-you-drive auto insurance (Kantor and Stárek 2014). In all of these, insurance companies target the pricing to the actual driving behavior of their customers (Soleymanian, Weinberg, and Zhu 2019). To offer individually targeted price discounts based on each consumer’s driving behavior, telematic devices measure some key elements of interest, such as the current time, acceleration, and position while driving with the help of sensors to detect speed, speed violations, braking, mileage, and traveling direction and to use this information to evaluate and build trust in a person’s driving behavior (Ma et al. 2018; Soleymanian, Weinberg, and Zhu 2019). Typically, the data are not only used by the insurance companies but are also provided as feedback to the drivers (e.g., via a smartphone app; (Merzinger and Ulbrich, n.d.; Soleymanian, Weinberg, and Zhu 2019).

Studying the effect on driving behavior, Soleymanian, Weinberg, and Zhu (2019) mentioned that telematic tariffs have led to an improvement in driving behavior, resulting in safer drivers. In particular, they found that younger drivers improve their usage-based insurance scores faster than older drivers after usage-based insurance adoption and that females exhibit more improvement than males. They noted that economic incentives lead to greater improvements in driving behavior. However, in a study using telematic technology to measure driving performance, Choudhary et al. (2019) focused on nonfinancial incentives. In collaboration with the industry, they launched a smartphone app called DrivePower that aims to nudge safe driving using feedback on the driving behavior. In their field experiment, they found that even such nudges improved the average driving performance.

To a certain extent, the telematic system can be interpreted as a reputation mechanism in the sense of the work by Resnick and Zeckhauser (2002). In the case of car insurance, the price expresses the company’s trust that the policyholder will not suffer an accident. The telematic system provides helpful information that the insurance company uses to determine whether clients are trustworthy drivers. The major aim of our work is to explore whether a telematic system can be used to build trust in the sharing economy. For this reason, we considered the possible benefits and concerns regarding using a telematic system compared to other reputation mechanisms. Next, we explore whether people might trust such a new mechanism, which is represented by the extent the information from a telematic system might influence their decisions.

Potential Benefits

Using telematic systems as reputation systems competes with other trust mechanisms, especially user ratings, as one of the most important mechanisms today. In the following paragraphs, we outline various benefits emerging from the use of telematic systems as a reputation mechanism. We focus especially on how telematic-based scoring systems can compensate for known weaknesses in user ratings.
Inflation of positive feedback. Even if no ground truth exists, user ratings often seem implausibly rosy (Filippas, Horton, and Golden 2018). Several authors have stressed that the bilateral reputation mechanism is often extremely positively biased (Nosko and Tadelis 2015; Dellarocas and Wood 2008; Resnick and Zeckhauser 2002). Horton et al. (2018; 2015) reported a similar phenomenon concerning the reputation system of a large online labor market. In addition, they indicated that the public rating scores increase strongly over time so that it is likely that the informativeness of reputation systems is eroding over time.

Several explanations exist concerning this discrepancy. Several authors have pointed out that the rating bias is the outcome of a reciprocity effect (Dellarocas and Wood 2008; Bolton, Greiner, and Ockenfels 2013; Horton and Golden 2015). Dellarocas and Wood (2008) assumed reciprocity expressed by a tit-for-tat strategy. This strategy represents an important driver of reporting behavior, fostered by the bilateral reputation system of eBay at that time. Bolten et al. (2013) pinpointed that reciprocal positive ratings help both buyers and sellers because it increases their reputations and reduces their transaction costs. By the same token, a negative rating harms the seller’s reputation, triggering possible retaliation by also providing a negative rating. This leads to additional costs for both buyers and sellers. In addition, the whole platform suffers because low average ratings decrease the overall trust and increase the overall transaction costs.

Horton and Golden (2015) noted a second reason for reputation inflation. They found that positive feedback and not providing negative feedback take market penalty into account, which is associated with bad ratings. For instance, The Guardian reported that Uber bans drivers with bad ratings (Paul 2019). If a rating has a dramatic consequence for the person concerned, it increases the risk of retaliation, which could also have negative consequences for the rater. To avoid such consequences or express solidarity with the gig worker, not rating the worker or giving the worker a positive rating might be a better option.

The use of a telematic system could counteract the inflation of positive feedback because it is resistant to reciprocity and solidarity for better or worse. Moreover, automatic scoring can be an emotional relief for peers because the technology serves as a scapegoat. In the case of doubt, the blame for a bad rating is ascribed to the technology instead of a person.

Fake, paid, and promotional reviews. Fake, paid, and promotional reviews can deceive consumers into making suboptimal decisions and can increase mistrust in the entire reputation mechanism (Mayzlin, Dover, and Chevalier 2014). Therefore, this problem has recently attracted significant interest in the mass media, consumer protection organizations, and academia (Wu et al. 2020). Fox Business, for instance, reported that e-commerce platforms are flooded with fake five-star reviews (Henney 2019), The New York Times reported on a case in which a company posted fake reviews for years (Maheshwari 2019). Consumer protection organizations warn against trusting peer reviews. In addition, they often advise on what consumers should look for in peer reviews. The consumer center in Germany, for instance, has recommended relying on professional and independent testers that are funded with tax money and work according to transparently defined criteria (Verbraucherzentrale 2018). In academia, most research focuses on how fake
reviews can be avoided or at least identified. For instance, Ott et al. (2011) proposed a linguistic model to identify opinion spam presenting inappropriate or fraudulent reviews. Akoglu et al. (2013) suggested a network-based framework for opinion fraud detection in online reviews. Heydari et al. (2016) proposed a time-series based model for detecting review spam. Hooi et al. (2017) suggested a graph-based model to detect fraud.

Despite these efforts, fake and paid reviews will likely still be a problem in the future. Hence, additional reputation mechanisms are helpful even if frauds also try to exploit the vulnerabilities in telematic systems. As a second independent mechanism, it provides additional protection, which makes it more difficult for frauds to falsify both the telematic scoring and user ratings.

Rating fatigue. The rating ratio, defined as ratings per usage represents the desire of the user to rate the product or service. In the case of eBay, Bolten et al. mentioned that 70% of the traders leave a rating. Resnick and Zeckhauser (2002) noted a similar ratio of about 50% for buyers and 60% for sellers. In the case of YouTube, Chatzopoulou et al. (2010) found a rating/viewing ratio of 0.25% on average. No comparable studies on rating ratios for P2P platforms were found, but we expect a ratio somewhere between that for eBay and that for YouTube.

Several reasons exist for the users’ unwillingness to report feedback. First, the rating means additional time and mental effort without any immediate benefit. Second, the evaluation also means the disclosure of personal information. Dellarocas and Wood (2008) also hypothesized that peers remain silent because they are afraid that reporting a negative experience will lead to retaliation by negative feedback. These reasons are reinforced by the fact that consumers are increasingly asked to evaluate products, services, locations, and so on, which can lead to rating fatigue.

Computer-based scoring can counteract such rating fatigue because it does not require any additional effort from users. Regarding the inflation of positive user feedback, for carsharing, considering all of the rides of a person, computer-based scoring becomes more balanced and reliable.

Discrimination. Ongoing evidence about everyday forms of ethnic and racial discrimination exists, especially regarding the labor market and housing market, but it also occurs in the service sector. Goddard et al. (2015) indicated that drivers pass black pedestrians more than twice the average rate, and black pedestrians wait longer than white pedestrians.

Concerning the sharing economy, Calo and Rosenblat (2017) stated that both service providers and service consumers face racial and other discrimination. Edelman (2017) found evidence that Airbnb guests with distinctively African American names were rated significantly less positive. They concluded that this penalty is consistent with the racial gap found in contexts ranging from labor markets and online lending to classified ads and taxicabs. Analyzing the data on a European carpooling platform using fictional profiles to ask for a ride, Carol et al. (2019) reported a similar result of ethnic and gender discrimination. Similarly, Tjaden et al. (2018) uncovered that drivers with an Arab-, Turkish-, or Persian-sounding name attracted significantly less interest in their offers (fewer clicks on the offer) than drivers with typical German names.
They concluded that these findings result from both taste-based and information-based discrimination. Regarding taste-based discrimination, the authors assumed that some consumers may not choose an offer because of stereotypes and prejudices. Regarding information-based discrimination, the authors assume that other consumers may not choose an offer because of security and safety issues. As there were no objective indicators on the platform, consumers might use the name of the driver as a cue for the perceived safety of a ride.

Intelligent systems are not free from bias, unfairness, and discrimination (Hacker 2018). To our best knowledge, however, telematic systems are not subject to the same types of discrimination as those described above. Therefore, telematic scoring might help reduce discrimination, especially information-based discrimination, because it can give an additional cue used by people to judge driving competence and the safety of a ride.

**Sectorial trust of heterogeneous user groups.** Although users of P2P platforms likely have basic interpersonal trust, it is conceivable that some might trust peer ratings more, whereas others have greater confidence in technical scoring. Providing an additional reputation mechanism can better consider this heterogeneity among users.

Moreover, because user ratings and telematic scoring examine the same subject from slightly different perspectives, people can triangulate these trust cues. If both systems measure the same directions, the two trust cues reinforce each other. An open question, however, is what happens when scores are opposing. In this case, it is reasonable that users create a kind of individual weighting of the scores, which reflects their personal sectoral trust. People who are more likely to trust other people are more likely to give a higher weighting to a user rating. In contrast, those who tend to trust technology are more likely to give higher weight to telematic-based scoring. Thus, telematic-based scoring should not replace the user rating but rather supplement it to consider the heterogeneity of the user group and support the individual, weighted triangulation of both trust cues.

**Additional added value.** In addition to these issues, a further added value of telematic-based scoring can be identified, which goes beyond the existing reputation assessment. Bossauer et al. (2020) examined the tension between trust and privacy using telematic systems in P2P carsharing by interviewing potential rentees and car owners. In contrast to user ratings, which provide an informational basis for decision-making, monitoring the car during the journey especially seems to promote trust on the car owner side. For example, some car owners specify the maximum speed of the car, which then can be controlled and reported by the telematic system in the event of a violation.

To build trust, the study further indicates that both parties should have access to the recording of the driving behavior and can use this data under the same conditions. For instance, in the case of litigation, telematic data can enable a more objective evaluation in the event of a dispute. In addition, under certain conditions, renters are willing to disclose information in specific situations if this promotes trust. For instance, the renter may be willing to share the current location because,
due to a traffic jam or other reasons, a renter may not manage to return the car in time.

Potential Concerns

Inappropriate trust. Telematic systems have so far only been used by insurance companies. Their scoring procedures are assumingly based on data-based risk analyses. However, to our knowledge, no independent evaluation of the procedures exists. Therefore, the extent telematic scoring reflects the true safety risk of a driver and accordingly conveys appropriate trust is unclear.

Moreover, the question about appropriate trust can only be answered from the ascribed meanings of such a telematic-based score. Due to the novelty of the concept, no knowledge yet exists on how much users trust such a score and whether it results in a lack of trust (not trusting a safe driver) or over-trusting (trusting an unsafe driver). Further studies are necessary to calibrate the trust cues to reconcile the perceived trust with the objective risk.

Privacy concerns. Telematic-based scoring requires that the renter’s driving behavior is recorded by a telematic box or similar device. The driving data are used to determine personal characteristics (e.g. how safely people are driving). Hence, an implementation of the concept must be compliant with the General Data Protection Regulation (GDPR). The collecting, processing, and storing of the data are only used for the specified purpose, and the informed consent of the renter is needed. A qualitative study (Stevens et al. 2018; Bossauer et al. 2020) revealed that renters are willing to disclose such data in principle, if they obtain advantages (e.g. a more favorable rent and a higher probability of borrowing a car). However, it is an open question regarding the extent a telematic-based scoring concept would be accepted by users and how it could increase the likelihood to rent a car via a P2P platform, such as Getaroud or Turo.

System accountability and the right to an explanation. The GDPR includes the right to “be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her” (GDPR §21.1). Whether concepts such as telematic-based scoring fall under this paragraph is debatable because computational scoring should inform but not replace human decision-making. Nonetheless, for various reasons, computational scoring and reputation systems should provide a “right to [an] explanation” (Edwards and Veale 2017), making the scoring accountable (Jakobi et al. 2018; 2019) for both parties: the person being scored and the person who uses the score. Regarding this, the algorithms behind telematic-based scoring must be reviewable and traceable. Computational scoring should reflect common sense about what is considered a “good driver.” In addition, the scoring system must not lead to unfairness and discrimination (Hacker 2018). We also need appropriate metaphors and visualization to make the computational
ratings comprehensible for both renters and car owners to increase trust in computational trust factors.

**Ethical concerns.** In addition to privacy and legal issues, computational reputation systems also raise the ethical question of how much decisions should be delegated to computers. The current development of social scoring in China, for example, and the dark side of the sharing economy (Calo and Rosenblat 2017) have raised the fundamental question of what influences reputation systems in general (both human and computational scoring) should have on social and economic systems. Moreover, in the sharing economy, online reputational mechanisms could but should not replace traditional service and work regulations, anti-discrimination laws, and inspection by public authorities (Ranchordás 2019). Even if someone voluntarily agrees to information disclosure, it remains an open question regarding whether computational scoring is ethically and socially acceptable for liberal societies (Landwehr, Borning, and Wulf 2019). For instance, the danger of unequal treatment concerning privacy exists, where rich people can own a private car, while poorer people must give up their privacy to rent a car.

## Methodology

The previous section revealed that using telematic systems as a reputation mechanism would be far from perfect in various situations. However, these systems have the potential to compensate for some of the imperfections in other mechanisms, such as user ratings and user reviews. However, besides technical feasibility, users must accept and trust such new reputation mechanisms.

An acceptance indicator is that people trust in the information provided and that this information influences their decision-making behavior regarding identifying trustworthy persons (Resnick and Zeckhauser 2002). To verify this, we conducted a conjoint analysis experiment in which we compared the influence of a fictional telematic score with that of a user rating. As the computational scoring is quite novel and does not consider “soft” issues, such as friendliness or cleanliness, we expect that telematic-based scoring does not influence the behavior of the car owner as strongly as the user rating.

## Procedure

We used a conjoint analysis as a common market research technique to measure customer preferences for new or changed features or prices for products and services and to guessimate unconscious decision processes (Dobney, Ochoa, and Revilla 2017; Rao 2014). In contrast to identifying the importance of individual product attributes, conjoint analysis measures the acceptance of complete products, which are regarded as a bundle of attributes and their importance. This way, the
analysis can specifically uncover how respondents develop preferences and help to develop new products.

For the conjoint analysis, a special focus must be placed on the choice and definition of factors (also called attributes) and their levels (also called values) (Rao 2014). As we aimed to analyze the effect of computationally provided ratings compared to the established peer rating, our conjoint analysis focused on these factors but neglected other, self-provided information on the person, such as gender, age, photo, and so on. We also did not include peer reviews because it is difficult to define the manifold examples in a controlled way.

Table 1 displays the factors, levels, and corresponding scale values that we presented to the participants. In a pre-test, we determined that all of them are crucial for the decision of the user. The coinage of the term trust score was an evolutionary process. Various terms used in the insurance industry indicate that no term has been established to communicate the telematic concept to the customer. Our goal was to use a term that does not lead to under-trusting or over-trusting in the score. Initially, we used the telemetric score as a technical but neutral term. However, in our pre-testing, we became aware that people did not understand what we meant, so they tended to under-trust the concept. Therefore, we explained the term in the sense that we speak of a telematic score that aims to support trust in the driving skills of a renter. In these discussions, the term trust score has become established to express the goal of the concept instead of the technology to reach it. In our conjoint analysis experiment, we used this term, defining the rating of such a telematic system as the trust score, which, in principle, could also be used in reputation systems for other areas of the sharing economy.

<table>
<thead>
<tr>
<th>Factor</th>
<th>Levels</th>
<th>Scale Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>User ratings</td>
<td>1 of 5 stars</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>3 of 5 stars</td>
<td>3.0</td>
</tr>
<tr>
<td></td>
<td>4.5 of 5 stars</td>
<td>4.5</td>
</tr>
<tr>
<td>Number of ratings</td>
<td>3</td>
<td>1.1</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>3.0</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>5.3</td>
</tr>
<tr>
<td>Trust score</td>
<td>1 of 5 stars</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>3 of 5 stars</td>
<td>3.0</td>
</tr>
<tr>
<td></td>
<td>4.5 of 5 stars</td>
<td>4.5</td>
</tr>
<tr>
<td>Kilometers driven</td>
<td>50</td>
<td>1.6</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>3.9</td>
</tr>
<tr>
<td></td>
<td>5,000</td>
<td>6.2</td>
</tr>
</tbody>
</table>

For the user rating, we used a five-star scale because this is common in platforms like Airbnb, Amazon, or Getaround. We chose three levels: 1 star, 3 stars, and 4.5 stars. The same scale and levels were used for the trust score to increase comparability. The star ratings for the factors user rating and trust score were
designed identically to make them easier to assess. Moreover, both factors used the same star visualization to avoid any effects caused by the presentation and not by the semantics of both factors.

We also included the number of ratings, which is common information in peer-based reputation systems, and correspondingly included kilometers driven as an additional factor in the computational reputation system (Table 1). For both factors, we used a log-linear scale to consider the psychometrical fact that a noticeable change is only perceived when the value doubles. In other words, a notable difference exists if someone has a driving experience of 100 km instead of 50 km, whereas the difference between 5,450 km and 5,000 km is negligible. We include three levels to represent low (three ratings [log: 1.1]/50 km [log: 1.6]), middle (50 ratings [log: 3.0]/500 km [log: 3.9]), and high (200 ratings [log: 5.3]/5,000 km [log: 6.2]) values.

The structure of the questionnaire has been divided into five parts. First, the participants receive an introduction and explanation of the trust score (Figure 1) to become familiar with the case.

<table>
<thead>
<tr>
<th>What is the Trust Score?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Many peer-to-peer carsharing platforms use telematics. For this purpose, car owners are provided with a small box free of charge, which is installed in the vehicle.</td>
</tr>
<tr>
<td>- The telematics box records the driving style of the tenants whenever the vehicle is booked and used by the platform.</td>
</tr>
<tr>
<td>- In addition to braking behavior, speed and acceleration, external factors such as weather, nighttime, location are also used to reliably record driving behavior - careful driving is rewarded with points via intelligent algorithms.</td>
</tr>
<tr>
<td>- If the tenant drives carefully, they receive a good trust score on the platform.</td>
</tr>
<tr>
<td>- If the tenant tries to manipulate the box, this is immediately recognized automatically and reported to the car owner.</td>
</tr>
</tbody>
</table>

Figure 1. Explanation of the trust score presented to the survey participants.

Second, the participants decided between two fictional profiles (Figure 2).

Now you should decide for a potential tenant of your car. You will always see two profiles at the same time. Which of the two applicants would you prefer?

<table>
<thead>
<tr>
<th>User ratings</th>
<th>Number of ratings</th>
<th>Trust score</th>
<th>Kilometers driven</th>
</tr>
</thead>
<tbody>
<tr>
<td>⭐⭐⭐⭐⭐</td>
<td>20</td>
<td>⭐⭐⭐⭐</td>
<td>5,000</td>
</tr>
<tr>
<td>⭐⭐⭐⭐⭐</td>
<td>20</td>
<td>⭐⭐⭐⭐⭐⭐</td>
<td>50</td>
</tr>
</tbody>
</table>
Third, respondents were asked three questions about their trust in user ratings (interpersonal trust) and three questions about their trust in the trust score (technology trust), each based on the Likert scale. Fourth, to uncover the perceived helpfulness of the trust factors, the participants rated the helpfulness of the provided information from 1 (not helpful) to 4 (helpful). At the end of the questionnaire, general data about car ownership and demographical items were gathered.

Sample and Acquisition

During the period from May 6, 2019, to June 15, 2019, the survey was electronically shared via email at the Hochschule Bonn-Rhein-Sieg University of Applied Science in Germany and via social networks. Moreover, 77 people participated, including 43 women and 34 men. The average age was 30.6. In this sample, 88.3% of the respondents had a car, and the median car value was between 15,000 and 25,000 EUR. The median car age was between 6 and 15 years. Most of the respondents had a positive bond with their cars (mean: 3.5; standard deviation: 1.2; 1 (I do not care) to 5 (high bond)).

Results

We analyzed the choice-based conjoint data using the multinomial logit model and the maximum likelihood method. We did not encode the levels by dummy variables (Baier and Brusch 2009) to estimate the part-worth utilities for each factor level but interpreted the factors as a metric. For different factor levels, we used the scale values in Table 1. We estimated the factor coefficient using multinomial logistic regression (Starkweather and Moske 2011). The calculations were conducted using the mlogit package for R (Croissant 2012).

Table 2 summarizes the findings of the logistic regression. Both the likelihood-ratio test and McFadden’s $R^2$ (Table 2) indicate a good model fit and predict significantly better than the null model. Respondents were only able to make their decision on the information provided in the profile. In a real-world situation, other factors that were not considered in this study also play a role. However, this indicates that the factors are relevant to the decision process and that the linear and log-linear scales approximated the “true” part-worth values quite well. This is also reflected by the fact that all trust factors are significant with a p-value of $< .001$ for the trust score and user rating, $< .02$ for the number of ratings, and $< .04$ for the kilometers driven.

1 To cross-check the results, we also estimated the part-worth utilities using the dummy encoded method. This led to a similar result, especially for the same ranking of relevance of trust factors.
The result reveals that the trust score has a significant influence on the decision of the car owner regarding who should be allowed to borrow the owner’s car. Moreover, our findings indicate that the coefficients of the trust score and user rating are quite close. This holds even when we consider the standard deviation of the coefficients. Compared to this, the trust factors of kilometers driven and the number of ratings are five times lower, which indicates that they are not more important than the other factors.

Our findings can best be interpreted by the odds ratio\(^2\) given in the right column of Table 2, which is calculated by taking the exponential value of the estimated coefficients. Each star in the trust score increases the ratio of the probabilities (odds ratio) by a factor of 2.91. For instance, two users who want to rent a private car have the same profile except that user A has one more trust score star than user B. In this case, user A has a 2.91 times higher chance to be selected as a trustworthy person than user B. This is slightly more than the effect of an extra star in the user rating, where the odds ratio is 2.67.

Table 2. Coefficients and odds ratios of the trust factors used in the conjoint analysis study

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>SD</th>
<th>Odds Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.01</td>
<td>1.00</td>
</tr>
<tr>
<td>Trust score</td>
<td>1.07***</td>
<td>2.91</td>
</tr>
<tr>
<td>User rating</td>
<td>0.98***</td>
<td>2.67</td>
</tr>
<tr>
<td>Kilometers driven</td>
<td>0.21**</td>
<td>1.23</td>
</tr>
<tr>
<td>Number of ratings</td>
<td>0.20***</td>
<td>1.23</td>
</tr>
</tbody>
</table>

Observations: 395
R\(^2\): 0.49
Log-Likelihood: -140.97
LR-test: 265.45*** (df = 5)

Note: *p<0.1; **p<0.05; ***p<0.01

Table 3 illustrates how helpful the provided information was perceived as being by the participants to select a fictional profile. First, the table reveals that all information was rated as helpful (a value above 3 means helpful or very helpful). This confirms the results of the conjoint analysis finding that all factors are relevant to the selection decision. However, there are also key differences. For example, the differences in the rating of information usefulness are much smaller than the coefficients of the conjoint analysis. One explanation could be that they were all considered helpful and relevant. However, the four-point scale we used provides too little possibility to further specify the degree of helpfulness. Hence, the ratings become remarkably close to each other.

Table 3. Perceived helpfulness of the provided information to select the fictive profile

\(^2\) For metric product properties, the odds ratio provides, *ceteris paribus*, the change in the odds when the metric property is increased by one unit of scale (m -> m+1) (Baier and Brusch 2009).
Table 4 Measurement model for the sectoral trust

<table>
<thead>
<tr>
<th>Factor</th>
<th>Items</th>
<th>Loadings&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Cronbach's alpha&lt;sup&gt;b&lt;/sup&gt;</th>
<th>AVE&lt;sup&gt;c&lt;/sup&gt;</th>
<th>CR&lt;sup&gt;d&lt;/sup&gt;</th>
<th>Rho A&lt;sup&gt;d&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peer Trust</td>
<td>PT1</td>
<td>.925</td>
<td>.502</td>
<td>.651</td>
<td>.784</td>
<td>.652</td>
</tr>
<tr>
<td></td>
<td>¬PT2</td>
<td>.668</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PT3</td>
<td>too small</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technology Trust</td>
<td>TT1</td>
<td>.810</td>
<td>.816</td>
<td>.731</td>
<td>.891</td>
<td>.816</td>
</tr>
<tr>
<td></td>
<td>¬TT2</td>
<td>.877</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>¬TT3</td>
<td>.876</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes:
- All item loadings > .5 indicates reliability.
- Cronbach’s alpha > .7 for a factor indicates reliability.
- Average variance extracted (AVE) > .5 for a factor indicates convergent reliability.
- Composite reliability (CR) and rho A > .7 for a factor indicate internal consistency.

We measured the sectoral trust regarding peer trust and technology trust using three items, each using a four-point Likert scale. Items PT2, T2, and T3 measured the factor negatively; hence, the scale was reversed before analysis. The factor loading of item PT3 was too small; thus, the item was dropped. The reliability of the remaining items was tested using composite reliability (CR); however, Cronbach’s alpha for peer trust is quite low at 0.502. Therefore, the results must be handled with care (Table 4).

We estimated the influence of sectorial trust on the perceived usefulness of the information using structural equation modeling (SEM), which is illustrated in Figure 3.

It is also noticeable that, in the perceived helpfulness, the trust score did not reach the first rank as in the conjoint analysis. Moreover, it only placed at the third rank, behind even the number of ratings. Although it played only a minor role in the selection process, the participants consider it to be the second-most important information. In summary, it seems that a kind of perception-action bias exists in that, in the deliberate reflection, the relevance of peer-provided information seems to be rated higher than the actual influence on the decision-making.
Figure 3 Influence of the sectoral trust on the perceived usefulness of the trust cues provided. Structural equation model together with the estimated path coefficients and $R^2$ (Calculations were carried out with SmartPLS v3.2.8. Ringle, Wende, and Becker 2015).

According to Henseler (2009), one of the main assessments of a structural model comprises the evaluation of the $R^2$ of the latent variables. Regarding the provided trust cues, $R^2$ indicates that the sectorial trust in peers and technology explains 31.2% of the variance of the perceived usefulness of the user rating and 50.3% of the variance of the perceived usefulness of the telematic-based scoring.

Sectoral trust influences the perceived usefulness to varying degrees. While peer trust loads strongly on the peer rating at 0.572, it loads only weakly on the telematic-based scoring at 0.093. This result can be interpreted to mean that the user rating is a more helpful reputation system for people with a high level of trust in others. The opposite is the case with technology trust. At 0.709, the technology trust loads strongly on the trust score but negatively affects the user rating. Hence, the trust score seems to be a more helpful reputation system for people with a high level of trust in technology.

This preliminary result indicates that telematic-based scoring should not replace but should supplement the traditional user rating because the scores address different user groups. Nonetheless, these findings must be handled with care because the findings are based on a small sample size, and our model neglects any moderating effects, such as the existence of generalized trust that might affect both interpersonal and technology trust.

Conclusion

Trust plays a key role in the sharing economy in general and in sharing a car with a stranger in particular. Reputation systems have become an established means to overcome the trust dilemma in the sharing economy. Today, reputation systems primarily rely on trust factors based on peer-provided information, such as peer ratings and peer reviews. In this paper, we suggested using additional trust factors that use computationally provided information because such information can be
collected automatically and evaluated using objective measures. We explored the adoption of telemetric-based scoring of driving behavior, which is currently used only by car insurance companies.

Our study aimed to determine (1) whether such a trust score influences the behavior of car owners in distinguishing between trustworthy and untrustworthy peers and (2) the extent of this effect. Both the findings of the conjoint study and the perceived helpfulness present the first indicator that telemetric-based scoring can be effective. Moreover, some evidence indicates that such scoring might have a critical role in future reputation systems and play the same role as user ratings. Thus, computational reputation systems using telemetric technologies have the potential to enhance the overall carsharing experience.

However, several limitations must be considered. First, all measured effects are statistically significant, but the sample is still relatively small with $N = 77$. In addition, the sample is formed by young people who are, on average, 33.6 years old. While this reflects a common finding in the literature that the sharing economy is particularly popular with younger people, the aim should be to motivate older people as well. It can be assumed that older people are more likely than younger people to own a car that they could make available to others. Therefore, the results should not be over-generalized.

In addition, our research design focused on the main effects only, and any moderation effects were excluded for the sake of simplicity (such as the interaction between the number of ratings and user rating and between the kilometers driven and trust score). This makes our model more parsimonious; however, in the future, moderation factors should also be considered.

Another important limitation is that most participants had no first-hand experience with the telematic-based scoring. It is quite usual for conjoint analysis to evaluate not-yet-existing product features; however, the fictional profiles limited the ecological validity of the experiment. Hence, the experiment reflects the projection of how people might interpret this factor in the future. Thus, our study demonstrates the potential of a telematic-based scoring for P2P carsharing, but to realize this potential, the socio-technical concept of using telematic technology for this purpose must be adequately implemented.

Moreover, two interrelated issues must be addressed critically. As outlined in the beginning, no calibration avoids under- or over-trusting in the information provided by a telematic-based reputation system. In our pre-testing, we became aware that the term *telemetric score* was too technical, and people had difficulty understanding the intention of the score. Thus, the danger that people could under-trust and ignore the score exists. In contrast, our study reveals that naming the construct *trust score* helps people consider this information in their decisions. However, future studies are needed to validate whether such a term avoids over-scoring and misinterpretation.
In addition, our study indicates observed perception-action where participants are unconsciously more influenced by the computational rating when they are aware of it. Hence, the danger of a non-reflective manipulation by algorithms exists, which is perhaps even worse because it takes place unconsciously and is largely concealed.

Our study has demonstrated the potential of a computational reputation system for the sharing economy. In the future, the concept should be elaborated in diverse ways. We require a better understanding of how decision-making works in practice when using the information provided by telemetric technology. For instance, we should study the situated actions of lending cars to others via carsharing platforms. An interesting question would be whether general interpersonal trust and technology trust influence trust factors. We should also explore the concept regarding other sharing areas (e.g. Does telematic-based scoring have a similar positive influence on the trustworthiness of Uber drivers?). Finally, to validate our findings, the concept should be implemented prototypically under real-world conditions. Regarding this, our work presents a pre-study showing both the practical relevance and theoretical foundation of using computer-based scoring as a mechanism that can compensate for some weakness in the traditional user rating.

Acknowledgement

We want to thank Katrin Igelmund, Elisa Kart, and Sarah Kretzer, students of the Bonn-Rhein-Sieg University of Applied Sciences, for supporting the empirical work. Further, we want to note that this research was funded by Federal Ministry of Education and Research; grant number 16SV8257.

References


Beierlein, Constanze, Christoph Kemper, Anastassyia Kovaleva, and Beatrice Rammstedt. 2012. ‘Kurzskala Zur Messung Des Zwischenmenschlichen Vertrauens: Die Kurzskala Interpersonales Vertrauen (KUSIV3)’. GESIS.


Choudhary, Vivek, Masha Shunko, Serguei Netessine, and Seongjoon Koo. 2019. ‘Nudging Drivers to Safety: Evidence from a Field Experiment’.


Report DST-Group-TR-3696. Joint and Operations Analysis Division
Defence Science and Technology Group.

Dellarocas, Chrysanthos, and Ritu Narayan. 2006. ‘What Motivates Consumers to
Review a Product Online? A Study of the Product-Specific Antecedents of
Online Movie Reviews’. In WISE, 1–6.

Dellarocas, Chrysanthos, and Charles A. Wood. 2008. ‘The Sound of Silence in
Online Feedback: Estimating Trading Risks in the Presence of Reporting
Bias’. Management Science 54 (3): 460–76.

Dinev, Tamara, and Paul Hart. 2006. ‘An Extended Privacy Calculus Model for

Dobney, Saul, Carlos Ochoa, and Melanie Revilla. 2017. ‘More Realism in
Conjoint Analysis: The Effect of Textual Noise and Visual Style’.


Doyle, Alicia. 2017. ‘Amazon Seller Scams: How Scammers Use Amazon
guides/amazon-seller-scams-how-scammers-use-amazon-marketplace-to-
rob-shoppers.

Edelman, Benjamin, Michael Luca, and Dan Svirsky. 2017. ‘Racial
Discrimination in the Sharing Economy: Evidence from a Field
22.

Edwards, Lilian, and Michael Veale. 2017. ‘Slave to the Algorithm: Why a Right
to an Explanation Is Probably Not the Remedy You Are Looking For’.

Ensminger, Jean. 2001. ‘Reputations, Trust, and the Principal Agent Problem’.


Ert, Eyal, Aliza Fleischer, and Nathan Magen. 2016. ‘Trust and Reputation in the
Sharing Economy: The Role of Personal Photos in Airbnb’. Tourism

Measurements of Interpersonal Trust’. Journal of Research in Personality

Inflation’. In Proceedings of the 2018 ACM Conference on Economics
and Computation, 483–84.

Ford Motor Company. 2016. ‘Ford-Umfrage: Viele Europäer wären bereit, ihr
Auto gegen eine Gebühr zu verleihen, sogar ihr Zuhause und ihren Hund’.

Gatersleben, Birgitta. 2007. ‘Affective and Symbolic Aspects of Car Use’. In
Threats from Car Traffic to the Quality of Urban Life: Problems, Causes
and Solutions, 219–33. Emerald Group Publishing Limited.

among Online Consumers’. ACM SIGMIS Database: The DATABASE for
Advances in Information Systems 33 (3): 38–53.


Ringle, Christian M., Sven Wende, and Jan-Michael Becker. 2015. ‘SmartPLS 3’. Boenningstedt: SmartPLS GmbH.


https://www.verbraucherzentrale.de/wissen/vertraege-reklamation/kundenrechte/was-ist-von-onlinebewertungen-zu-halten-13724.


