

# Trust, reputation, and the value of promises in online auctions of used goods

Rationality and Society

2023, Vol. 0(0) 1–33

© The Author(s) 2023



Article reuse guidelines:

[sagepub.com/journals-permissions](https://sagepub.com/journals-permissions)

DOI: 10.1177/10434631231170342

[journals.sagepub.com/home/rss](https://journals.sagepub.com/home/rss)**Judith Kas** 

Department of Sociology/ICS, Utrecht University, Utrecht, The Netherlands

**Rense Corten** 

Department of Sociology/ICS, Utrecht University, Utrecht, The Netherlands

**Arnout van de Rijt** 

Department of Sociology/ICS, Utrecht University, Utrecht, Netherlands; European University Institute, Fiesole, Italy

## Abstract

Buyers in online markets pay higher prices to sellers who promise a high-quality product in auctions of used goods, even though they cannot assess quality until after the sale. The principal argument offered in prior work is that reputation systems render sellers' 'cheap talk' credible by allowing buyers to publicly rate sellers' past honesty and sellers to build a reputation for being honest. We test this argument using both observational data from online auctions on eBay and an internet experiment. Strikingly, in both studies we find that unverifiable promises are trusted by buyers regardless of seller reputation or the presence of a reputation system, and sellers mostly refuse to take advantage. We conclude that the prevailing conception of markets in economic sociology as made possible by opportunism-curtailling institutions is "undersocialized": Reputation systems may be used to identify more reliable providers of a product, but that they would be needed to

---

### Corresponding author:

Rense Corten, Department of Sociology, Utrecht University, Sjoerd Groenman Building, Padualaan 14, Utrecht 3584 CH, Netherlands.

Email: [r.corten@uu.nl](mailto:r.corten@uu.nl)

prevent otherwise rampant deceit relies on a cynical assumption about human behavior that is empirically untenable.

### **Keywords**

Reputation systems, trust, communication, signalling theory, eBay, online platforms

## **Introduction**

Every day, millions of second-hand items are auctioned off, sent at an agreeable price to a satisfied new owner. While perhaps at first glance a mundane social fact, we contend that what enables such smooth cooperation between perfect strangers is poorly understood, yet is a matter that goes to the heart of the social sciences. Online auctions of used goods are characterized by great variability in the quality of products. The same seller sometimes has an item to offer that is in excellent condition while at other times something that is not in great shape but nonetheless of interest to some buyer at a reduced price. To differentiate between these situations, sellers communicate the state of the item before auctioning it off. Research on auctions for used goods shows that most sellers honestly report on defects, and when they make promises of high quality, their items are indeed sold for higher prices (Anderson et al., 2007; Lewis 2011; Neto et al. 2016; Rawlins and Johnson 2007). Why do buyers believe anonymous sellers' promises if they have limited recourse in case of deceit and why do sellers not take advantage?

An ostensibly naive answer is that most people have been raised as honest human beings and believe others have too. Such would rely on a theory of action some would consider "oversocialized" (Granovetter 1985; Wrong 1961). Yet, as we will show, this is precisely the conclusion to which our research leads us. The consensus view in the literature on online auctions instead is that most people are absolutely capable of deceit when this serves them, but buyers can trust sellers' claims because sellers care about their reputation (Lewis 2011; Ba and Pavlou 2002; Houser and Wooders 2006; Snijders and Zijdemans 2004; Neto et al. 2016; Rawlins and Johnson 2007; Bolton et al. 2004; Depken and Gregorius 2010; Melnik and Alm 2005). This view constitutes a special case of the more general perspective in economic sociology that institutions such as reputation systems are needed to resolve issues of distrust and malfeasance in markets (Nee and Swedberg 2020; Zucker 1986; Kollock 1999; North 1990; Erikson et al., 2006; Greif 1989; Cook et al. 2005; Powell 1990; Simpson and Willer 2015; Thiemann and Lepoutre 2017; Sabin and Reed-Tsochas 2020).

In the specific case of second-hand auctions, many platforms allow buyers to publicly report and look up cases of honest behavior or broken promises with minimal effort. Reputation systems for online auctions communicate not only the quality of past goods or services delivered by a seller, but also their honesty about that quality. The quality of a product may be low, but a buyer will not complain about it unless it was unexpected. On eBay, reviews may contain comments like: ‘Just what I ordered!’ or ‘Item description was very accurate’. Buyers are more satisfied with their purchase if the product matches the product description than when the seller overstated the quality of the product. More satisfied buyers will provide better reviews, and good reviews are important for sellers, because they allow their products to be sold at high prices in the future. By providing clues from the past about the present intentions of the sellers, reputation systems would help buyers distinguish between sellers who intend to provide a high-quality good and sellers who do not intend to do that (Gambetta, 2009; Przepiorka and Berger, 2017; Przepiorka and Diekmann, 2013; Raub, 2004). Reputation systems would thus attach a cost to a false promise (Lewis 2011; Ba and Pavlou 2002; Houser and Wooders 2006; Snijders and Zijdemans 2004; Neto et al. 2016; Bolton et al. 2004; Depken and Gregorius 2010; Melnik and Alm 2005).

Reputation systems are therefore regarded as instrumental for trust in anonymous online markets (Teubner et al., 2017; Teubner et al. 2017; Cook et al. 2005; Przepiorka et al. 2017; Raub and Weesie 1990; Resnick et al., 2000; Tadelis 2016; Jiao et al., 2021). After all, in the counterfactual scenario of no reputation system the situation would be much like Akerlof’s (1970) classic “market for lemons” (Diekmann et al., 2014). Sellers’ promises should then not affect buyers’ decisions, as dishonest information can be provided without consequence (Akerlof 1970; Duffy and Feltovich 2006; Baldassarri 2015) when a buyer and seller will not meet again (Molm et al. 2000). The information asymmetry would therefore pose a risk of adverse selection that would drive high quality products out of the market (Akerlof 1970; Kollock 1994; Macy and Skvoretz 1998; Buchan et al. 2002). Buyers would not be willing to pay a high price for a product if they did not believe the seller was honest about its quality.<sup>1</sup> Sellers, in turn, could not profit from the sale of a high quality good. Sellers should then only provide low-quality products, and buyers should only pay low prices. Such market failure would be prevented by reputation systems.

While this institutional argument has repeatedly been made to account for buyers’ trust in auction sellers’ quality assurances, we are not aware of any empirical study in which it is evaluated. We tested the argument both in an observational study as well as in an online experiment. The observational data we draw on were collected by (Lewis, 2011; 2019) and contain information on a large number of auctions completed on eBay. The online experiment allows a controlled test of the argument by systematically

varying the presence of institutions that allow sellers to signal quality and buyers to publicly post evaluations. As it turns out, both studies refute the theoretical consensus: We find that unverifiable promises improve seller outcomes *regardless* of seller reputation or the presence of a reputation system. Reputation systems do not appear to add any weight to quality assurances. We conjecture that buyers correctly believe that most sellers are too decent to lie even when they can get away with it, so that a good reputation, while reassuring, is not instrumental to building trust.

## Theory

Here we flesh out in greater analytical depth the oft-made argument that we end up falsifying, that in auctions for second-hand products reputations render seller claims about product quality credible (Lewis 2011; Ba and Pavlou 2002; Houser and Wooders 2006; Snijders and Zijdemans 2004; Neto et al. 2016; Bolton et al. 2004; Depken and Gregorius 2010; Melnik and Alm 2005).

Sellers who sell second-hand products in auctions sometimes have high-quality products and sometimes only low-quality products. Buying second-hand products online involves several risks for buyers, due to defection by sellers or problems beyond sellers' control (Axelrod and Dion 1988). First, the quality of most products deteriorates over time, so the quality of used products will vary both between and within sellers. Second, after the buyer has paid for the product, the seller may not send the product, send it too late or send an inferior product. A classic example of this situation is Akerlof's (1970) market for second-hand cars in which buyers do not know if they are buying a good or a bad car. In many of these markets, sellers and buyers are strangers to each other and only interact once, and they have no incentive to invest in a good relationship (Lo Iacono and Sonmez, 2020; Yamagishi et al., 1998). This lack of dyadic embeddedness (Granovetter 1985; Uzzi 1996) is also typical for modern-day exchanges that are mediated by online platforms. The internet increases the radius of the trust problem: by lifting geographical barriers the internet allows interactions beyond the locally embedded social circles (Kuwabara 2015; Delhey et al. 2011). Online markets are therefore widely regarded as posing trust problems between buyers and sellers and have been widely studied as such (see Jiao et al., 2021 for a recent meta-analysis of more than one hundred studies of trust in online markets).

The prices buyers are willing to pay depend on product quality: They are willing to pay a higher price for a high-quality good than for a low-quality good. They want to avoid paying a high price for a product of inferior quality. Both buyers and sellers benefit more when a high-quality product is sold for a high price, than when a low-quality product is sold for a low price. However, without the prospect of future interactions, rational and selfish sellers have interests that are opposed to those of buyers: they benefit more

when they provide a low-quality product for a high price. In those markets, buyers are not willing to pay a high price for a product if they are not confident about its quality, and sellers will not provide a high quality good for a low price. This situation leads to adverse selection: it drives high quality products out of the market (Akerlof, 1970). To achieve better market outcomes, it would thus be beneficial if buyers could tell the difference between high- and low-quality products.

In online auctions, the final price is determined by the highest bidder. eBay is currently the largest platform for online auctions. Both individuals and businesses can sell their products via this platform. Sellers can list their products, choose a duration for their auction to run, set a starting price and provide a description of the products. Sellers can also add a reserve price, which is the minimum price that must be bid for the product to be sold. These reserve prices are not communicated to bidders, so they do not function as a signal of quality.<sup>2</sup> Bidders make bids that are publicly available. If the highest bid at the end of the auction is higher than the seller's reserve price, the buyer with the highest bid pays the price they offered and receives the product.<sup>3</sup>

Credible information may help a buyer estimate the probability that a seller will provide a high quality good. Different types of information can be communicated to buyers; one can distinguish between cheap and costly signals. Costly signals are *type separating*, i.e. sellers who intend to provide a high-quality good will produce the signal, while sellers who intend to send a low-quality good do not produce the signal (Przepiorka and Berger 2017; Gambetta 2009). When the net benefits of producing a signal are positive for sellers who intend to provide a high-quality good, they are negative for sellers who do not have that intention. In that case only sellers who really intend to send a high-quality product will produce the signal, while other sellers will not produce the signal. When a signal is not type-separating, it can easily be faked, and it is considered a *cheap signal*. The net benefits of producing a signal depend on the costs of producing the signal and on the expected (future) benefits of sending that signal.

One way for sellers to convince buyers of the high quality of their products, is to communicate this to buyers, for example through product descriptions. However, these descriptions are non-binding and can be produced at trivial cost by sellers who intend to provide a high-quality good and by sellers who intend to send a low-quality good alike. Hence, a description need not reflect a seller's true intentions and should be considered a cheap signal.

When a seller promises to provide a low-quality product, the best response for the bidder is to offer a low price, to avoid paying a high price for a low-quality good. When a seller promises to deliver a high-quality product, the best response of the bidders depends on whether they believe that the

seller will really provide a high-quality product. If they believe the message, the best response is to offer a high price. Otherwise they can best offer a low price. In the short term, sellers benefit more when buyers offer higher prices. This means that there is no reason for any rational and selfish seller to describe the product as a low-quality product, since they can in that case be sure that the buyer will not pay a high price. In the absence of a reputation system, under the assumption of goal-directed and self-interested behavior, sellers will always promise that they will provide a high quality good, so product descriptions are cheap signals.

However, empirical research shows that these product descriptions *do* affect buyers' decisions (Lewis 2011; Anderson et al., 2007; Rawlins and Johnson 2007; Neto et al. 2016; Snijders and Zijdemans 2004; Heijst et al., 2008; Sena et al. 2005; ter Huurne et al., 2021a). For example, Lewis (2011) found that sellers on eBay who auction second-hand cars receive higher prices when they include in the product description that the car does not have scratches, dents or rust. Sellers who mention any of these deficits receive lower prices than sellers who do not mention these deficits.

While such observations seem in line with the general observation in experimental research that communication tends to enhance cooperation in social dilemmas (Balliet, 2010), it is not straightforward that these effects translate to the online context. For instance, Bicchieri and Lev-On (2007) argue that computer-mediated communication in general tends to be less effective than face-to-face communication in fostering cooperation, and that this effectiveness crucially depends on features of the online social environment that allow the establishment of a social norm of promise-keeping. The presence of a reputation system is such a feature that may explain buyers' trust in product descriptions.

Reputation systems collect, aggregate and distribute feedback about sellers to buyers (Resnick et al., 2000). In markets in which sellers repeatedly sell a single product of fixed quality, ratings left behind by previous customers help buyers differentiate between sellers of good and bad products (Buskens, 2003; Cook et al., 2005; Kuwabara, 2015; Raub and Weesie, 1990; Sorenson, 2014; Weigelt and Camerer, 1988; Yamagishi and Yamagishi, 1994). Studies indeed show that positive ratings positively affect seller choice and price in homogenous markets (Boero et al., 2009; Charness et al., 2011; Duffy et al. 2013; Fehrer and Przepiorka 2013; Bolton et al. 2004; Diekmann et al., 2014; Simpson and Willer 2015; Resnick and Zeckhauser 2002; Frey and Van De Rijt 2016).

In contrast to such homogenous markets, sellers of second-hand items sometimes have high-quality products and sometimes low-quality products for sale. Each product is sold in a dedicated online auction. Prior to the auction the seller describes the state of the product. At this stage a selfish and dishonorable seller could willingly portray their product as better than it

really is. Reputation systems may help buyers determine how accurate quality claims made by seller about a specific product are (Lewis 2011; Ba and Pavlou 2002; Houser and Wooders 2006; Snijders and Zijdemans 2004; Neto et al. 2016; Melnik and Alm 2005; Depken and Gregorius 2010). Reputation systems may thus help overcome Akerlof's adverse selection problem when there is heterogeneity in product quality between different offerings by the same seller.

Through numeric ratings and reviews, online reputation systems often contain information about the quality of past products sold by the seller *and* about the accuracy of a seller's past quality claims. Reputation systems thus do not only allow buyers to select sellers on the basis of the quality of the products they have delivered in the past. They also allow buyers to assess the extent to which sellers have been honest in the past. Sellers who have lied in the past about the quality of their products are likely to have worse ratings and reviews than sellers who have not lied in the past. When assessing the accuracy of a product description, buyers can thus use the information in the reputation system. Sellers who have lied in the past can reasonably be expected to do so again in the future. Sellers previously found honest who promise to deliver a high-quality product are therefore expected to receive higher prices than sellers previously found dishonest when making the same promise. Buyers should then be more likely to believe information if that information is provided by more honest sellers.

From this theoretical argument, we now derive a number of hypotheses. First:

**H1.** The better the reputation of a seller for being honest is, the more likely buyers are to trust claims about quality from this seller.

If one assumes that reputation systems also contain information about the accuracy of past product descriptions, reputation systems do not only provide an incentive to deliver high-quality products, but also an incentive to be honest in describing products when they are of low quality. In case of a product of inferior quality, it may be tempting to selfish sellers to lie about the quality of the products in order to attract a higher price. However, in the presence of a reputation system the long-term benefits of communicating honestly may outweigh the short-term costs. In the long run, sellers may benefit from the reputational gains from honest communication in the past. If the expected future reputation gains are larger than the short-term costs, sellers are expected to provide a high quality good if they can, and to communicate honestly when they sell a product of inferior quality. The reputation system thus provides an incentive for honest communication, which makes that communication more credible. While under assumptions of goal-directed and self-interested behavior, sellers who promise to deliver a

high-quality product in the absence of a reputation should not be trusted, the reputation system creates the conditions for these messages to be credible. Combining communication with reputation may thus make cheap talk costly. According to this argument then, sellers who want to keep up a good reputation need to be honest about the quality of their products to buyers. In the presence of a reputation system, sellers should then be more likely to do as promised.

**H2.** The correlation between promises made by sellers and their actual behavior is stronger in the presence of a reputation system.

Because the information that is provided by sellers about the quality of their products is expected to be more truthful when there is a reputation system, buyers should rely more on that information when there is a reputation system.

**H3.** The effect of promises by sellers to honor trust on trust of the buyers is stronger in the presence of a reputation system.

When sellers and buyers interact only once and when there is no reputation system, noncooperative behavior has no direct disadvantage for individual sellers, but it leads to negative outcomes for the entire market. In the presence of a reputation system, and with the possibility to communicate, sellers can credibly communicate the quality of their products, and buyers can adjust the prices they offer to the quality of the products. They are expected to pay higher prices to goods of better quality, and lower prices to products of inferior quality. Without the reputation system, communication between buyers and sellers is cheap talk, and buyers are not expected to pay higher prices when sellers promise to deliver better quality products.

When product quality within sellers is sufficiently heterogenous, reputation alone is not enough to solve the trust problem without a communication system for providing feedback on promises kept and broken. When there is only a reputation system, but no communication between buyers and sellers, the reputation system only contains information about the quality of the product, not about the accuracy of past product descriptions. However, when the quality of the products offered by the same seller varies greatly over time, the quality of products delivered in the past are no guarantee for the future. Consequently, in markets with heterogeneity in product quality within and between sellers neither communication nor a reputation system are expected to be a sufficient condition for better market outcomes. There should then be no difference in trust between cases where only reputation building or only communication are possible or where neither is possible, and that trust increases only when both reputation building and communication are enabled. Both communication and a reputation system would be necessary to increase trust.



**H4.** Buyers place more trust in sellers when there is a reputation system and a communication system than when there is a) only communication, b) only reputation or c) neither communication nor reputation.

## **Empirical strategy: an observational study and an experiment**

To the best of our knowledge, the consensus argument that communication between buyers and sellers affects buyers' decisions because there is a reputation system is never tested in any of the studies that advance it (Lewis 2011; Ba and Pavlou 2002; Houser and Wooders 2006; Snijders and Zijdemans 2004; Neto et al. 2016; Rawlins and Johnson 2007; Bolton et al. 2004; Depken and Gregorius 2010; Melnik and Alm 2005). We are aware of one study on the interaction between seller reputation and claims made by sellers: studying auctions of baseball cards on eBay, Jin and Kato (2006) do not find evidence that buyers rely more on claims made by more reputable sellers. However, their dataset contained both auctions of graded and ungraded cards. While the quality of ungraded cards is very difficult to detect online, this is not true for graded cards. For the latter category, claims and ratings of sellers have a different effect (Jin and Kato 2006). Moreover, the rating of the seller and the claims about the quality of the cards made by the seller were found to be correlated with the price of the cards, as well as with whether the card was graded or not. Based on the pooled results for graded and ungraded cards we cannot conclude that the effect of claims is equally strong for reputable and for un reputable sellers of ungraded cards.

We test H1 in Study 1 by using data from eBay auctions of second-hand cars. When using data from real platforms one is bound to the institutional design of that platform. eBay allows its sellers to communicate with buyers, and to build a reputation. However, to test H2-H4, we will need to systematically vary the presence of a reputation system and the opportunity for sellers to communicate with buyers. We are aware of a few studies in which the combined effect of reputation systems and communication systems is tested (Brosig-Koch and Heinrich 2018; Cason and Mui 2014; Denant-Boemont et al. 2011; Duffy and Feltovich 2006; Wilson and Sell 1997). However, none of these studies implement the Akerlof problem of quality heterogeneity characteristic of online auctions and are therefore of little use for understanding the problem at hand. In Study 2 we use an online experiment with quality heterogeneity within and between sellers in which we test if the presence of a reputation system changes the effect of communication on trust (H2-H4), while also testing H1 once more.<sup>4</sup>

### *Study 1: observational data from an online auction platform*

For Study 1, we use publicly available<sup>5</sup> data collected by (Lewis, 2011; 2019), containing information on 146,734 auctions of cars offered between

March and October 2006 on eBay. Paying for a car involves a major trust problem, because cars are expensive items and the buyer cannot know whether the car is of high quality. If buyers offer a (higher) price for a car, this can thus be regarded as an expression of trust in the seller. Trust was therefore operationalized as the number of bidders, the highest bid and the probability that the car was sold. We analyzed whether the seller's reputation and claims about the quality of the car affected these dependent variables in accordance with H1. Following [Lewis \(2011\)](#), we excluded auctions of new cars, cars that were built before 1950, incomplete or incorrect observations (e.g. with a negative mileage) and salvaged cars. We also excluded cars that could be bought immediately for a specified price and cars with a warranty, because the trust problem is less severe in those cases. A single outlier with a bidding price of \$860,100 was excluded (this is more than \$600,000 above the next highest price, and more than 80 standard deviations above the average). The remaining dataset contains 65,243 observations.

Positive seller reputation was operationalized as the number of positive ratings the seller obtained, and negative reputation was defined as the number of negative ratings. Because these distributions were highly skewed, with many sellers having no or a few ratings and only very few sellers with many ratings, the measures we include in the analysis are the natural logarithm of these numbers, after first adding 1 to avoid  $\ln(0)$ .<sup>6</sup>

In our measurement of claims about product quality made by sellers we follow [Lewis \(2019\)](#) who identified words in product descriptions that were relevant for the value of the car (e.g. rust, dent and scratch) and were frequently used. He then identified for each car if the seller mentioned the word and in which context. Sellers either did not mention a problem, informed the buyer that the car had the problem ('car has rust'), or explicitly mentioned that the car did not have the problem ('car is rust free'). We used this information to construct two variables counting the number of problems acknowledged and denied by the seller.

To test H1, we regressed measures of buyer trust on measures of seller reputation. We used the reputation of the seller and claims made by the seller as independent variables. We used three different dependent variables. First, we took the natural logarithm of the price offered by the highest bidder. Listings that did not receive any bids are excluded from this specific analyses. We used linear regressions for this variable. For the second dependent variable (whether the car was sold or not) we used logistic regression. The third dependent variable is the number of bidders, which is a count variable. We used negative binomial regression for this variable.

For every dependent variable, we first constructed a model with the independent variables and all control variables, as well as three sets of fixed effects, for car model, year of production and week of the auction (models 1, 3, and 5). Some cars were auctioned by the same seller, so we included

random intercepts for sellers. In a second model (models 2, 4, and 6) we added the interaction between the reputation variables and the number of problems denied and acknowledged by the seller to test if bidders rely more on claims made by sellers with a better reputation (H1).

In all models we controlled for several car characteristics: logged mileage, logged length of the product description (in bytes), whether the car was a featured listing, and the number of options a car had (categorized as none, one or two, or more than two, because about half of the cars has up to two options). We also controlled for the duration of the auction (in days) and for several seller characteristics: whether the seller provided a phone number and address and whether the seller was a power seller. The histogram of mileage shows a relatively normal distribution, except for a high peak at the maximum of 500,000 miles. The results do not change when these cars are excluded from the analyses.

Table A1 in the online appendix contains the descriptive statistics for the variables used in the analyses. The average seller has many more positive reviews than negative reviews. This skewness of rating scores has been reported for various kinds of platforms (Teubner and Glaser 2018; Kas et al., 2021; Bridges and Vásquez 2018) and suggests that dishonest seller behavior is rare. The descriptive statistics also show that on average, sellers acknowledge more problems than they deny.

## Results of study I

Table 1 displays the results of the mixed effects regressions, excluding control variable effects. The full table including the results for the control variables can be found in the online appendix, Table A2. Our regression results replicate the main effects of reputation reported in Lewis (2011). The negative main effects of positive and negative reviews on price (model 1) are likely due to sellers with time selling lower quality products (Lewis 2011).<sup>7</sup> This arguably makes our price variable less suitable as measure of buyer trust. Sellers with a larger number of positive reviews were more likely to sell their car (model 3), while negative ratings did not significantly impact sales chances. Sellers with more positive ratings attracted more bidders (model 5), while those with more negative ratings drew fewer bidders (model 5).

We also replicate the main effects of denying and acknowledging problems reported in Lewis' (2011). On average, denying problems did not affect prices (model 1), but had a positive effect on the probability that the car was sold (model 3) and on the number of bidders (model 5).<sup>8</sup> Acknowledging problems had a negative effect on prices (model 1), but a positive effect on the number of bidders (model 5) and the probability that the car was sold (model 3). It is plausible that sellers who acknowledge problems set lower secret reserve prices, so that they are more likely to sell their car

Table 1. Multilevel regression results (control variables omitted).

	Logged price			Sold		
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
<b>Main effects</b>						
Log positive ratings	-0.01 <sup>***</sup> (0.00)	-0.01 <sup>***</sup> (0.00)	0.33 <sup>***</sup> (0.01)	0.36 <sup>***</sup> (0.01)	0.03 <sup>***</sup> (0.00)	0.04 <sup>***</sup> (0.00)
Log negative ratings	-0.02 <sup>***</sup> (0.01)	-0.04 <sup>***</sup> (0.01)	0.03 (0.03)	0.01 (0.04)	-0.02 <sup>*</sup> (0.01)	-0.02 <sup>*</sup> (0.01)
Denied	-0.00 (0.00)	-0.02 <sup>**</sup> (0.01)	0.19 <sup>***</sup> (0.02)	0.27 <sup>***</sup> (0.04)	0.02 <sup>***</sup> (0.00)	0.07 <sup>***</sup> (0.01)
Acknowledged	-0.10 <sup>***</sup> (0.00)	-0.11 <sup>***</sup> (0.01)	0.26 <sup>***</sup> (0.02)	0.39 <sup>***</sup> (0.04)	0.02 <sup>***</sup> (0.00)	0.07 <sup>***</sup> (0.01)
<b>Interactions</b>						
Positive * denied	—	0.00 (0.00)	—	-0.02 <sup>*</sup> (0.01)	—	-0.01 <sup>***</sup> (0.01)
Negative * denied	—	0.00 (0.01)	—	0.02 (0.03)	—	0.00 (0.01)
Positive *	—	0.00 (0.00)	—	-0.04 <sup>***</sup> (0.01)	—	-0.02 <sup>***</sup> (0.00)
Acknowledged	—	0.03 <sup>***</sup> (0.00)	—	0.02 (0.03)	—	0.01 <sup>*</sup> (0.01)
Negative *	—	0.00 (0.00)	—	-1.29 (0.74)	—	1.96 <sup>***</sup> (0.18)
Acknowledged	—	9.70 <sup>***</sup> (0.15)	-1.22 (0.74)	—	1.98 <sup>***</sup> (0.18)	—
Constant	9.70 <sup>***</sup> (0.15)	9.70 <sup>***</sup> (0.15)	-1.22 (0.74)	-1.29 (0.74)	1.98 <sup>***</sup> (0.18)	1.96 <sup>***</sup> (0.18)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
FE model/year/week	Yes	Yes	Yes	Yes	Yes	Yes
Seller intercept	Yes	Yes	Yes	Yes	Yes	Yes
Observations	56,923	56,923	65,243	65,243	65,243	65,243

Note: 95% Confidence Interval in parentheses. <sup>\*\*\*</sup> $p < .001$ , <sup>\*\*</sup> $p < .01$ , <sup>\*</sup> $p < .05$ . Random intercepts for sellers. Models 1 and 2: Multilevel linear regression. Models 3 and 4: Multilevel logistic regression. Model 5 and 6: Multilevel negative binomial regression.

despite lower prices, and when auctions stay low they likely attract more bidders.

Our main interest is in the interaction effects between reviews and claims about car quality. To test H1 that buyers are more likely to believe claims made by sellers who have been more honest in the past, we interacted the reputation variables with the claims made by the sellers. The key test of H1 lies in the interaction effect of problems denied (that is, promises of quality) and reputation. Contrary to theoretical expectations, the significant interactions between the number of problems denied and the number of positive reviews on the probability that the car was sold (model 4) and on the number of bidders (model 6) suggest that sellers with more positive reviews benefit less from denying problems than sellers with fewer positive reviews. Further contradicting H1, the effect of positive quality claims on prices did not differ across sellers with many or few negative reviews.

We furthermore find that the effect of acknowledging problems on sales probability and number of bidders is more strongly negative for sellers with more positive reviews, while sellers with more negative reviews suffer less from acknowledging problems: the negative effect of acknowledging problems on the price is smaller for sellers with more negative reviews, and the positive effect of acknowledging problems on the number of bidders is larger for them. This suggests that, because low-reputation sellers are less likely to be trusted in the first place, acknowledging problems is most damaging to high-reputation sellers who have “more to lose” in terms of trust. While this latter result is in itself not inconsistent with the theory, overall, the results do not consistently show that quality claims more likely to be believed when sellers have a better reputation, so we reject H1.<sup>9</sup>

In defense of the theory of reputation, one could offer as potential explanation for why we did not find clear support for the predicted effect that we could not distinguish between having a reputation for honesty and a reputation for providing high-quality products. Reputation was operationalized by the numeric ratings provided by other users on the platform. These ratings were the results of a users’ (mental) aggregation of the evaluation of different aspects of the interactions. The data do not provide any details on what these aspects were and their weighing, so we do not know if these ratings were the result of providing a high-quality product or of honest communication or both. Whereas rating scores tend to be overly positive, written reviews have been found to be more detailed and nuanced (Bridges and Vásquez 2018). The written reviews could thus provide more insight here, but these were not available in the data.

The observational nature of the data also prevents us from ruling out confounding effects of omitted variables. Both reputation scores and trust in promises made by sellers may be caused by other, unobserved factors. For example, earlier studies on discrimination in the platform economy show

that demographic characteristics such as ethnicity affect an individual's chances of being trusted (Edelman and Luca 2014; Kas et al., 2021; Tjaden et al. 2018). Given that on most platforms, reviews can only be written after a completed transaction, these initial differences in the chances of being trusted may be reflected in the number of reviews an individual has managed to accumulate. Main effects of reputation and interaction effects between reputation and quality claims may then spuriously reflect main and interaction effects involving ethnicity. Similarly, individual personality differences may theoretically lead to spurious effects. For instance, traits like "agreeableness" (Ashton and Lee, 2007) may be correlated with behaviors in interpersonal interactions that affect the likelihood of being trusted, which again could be reflected in numbers of reviews. To address these issues, we conducted a controlled experiment in Study 2.

## **Study 2: an online experiment**

We conducted a second test of the theory in an online experiment in which we systematically varied the presence of a reputation system and the possibility for communication between buyers and sellers. This allowed us to test H1 in a controlled environment, eliminating the influence of omitted variables through randomization, and to test H2, H3 and H4 regarding the presence of a reputation system.

The experiment consisted of 36 rounds of an adaptation of the Trust Game (Dasgupta 1988; Kreps 1996) widely used to study effects of reputation. Subjects in a session were divided into groups of six or eight participants. At the beginning of each round, half of the participants in a group were assigned to role A (buyer), the other half was assigned to role B (seller). The game was divided into three blocks of 12 rounds. In each of these blocks, every subject played six rounds in role A and six rounds in role B, but the order in which they played in the different roles was determined randomly (cf. Charness et al., 2011).<sup>10</sup> In every round, each player A was randomly and anonymously matched with a player B.

The pairs played one of two versions of a Trust Game. The left part of Figure 1 shows the first version of the game, which was a regular Trust Game. In that game, player A could choose to place trust ('pay a high price') or to not place trust ('pay a low price') in player B. If player A did not place trust, the game ended and both players received a payoff of two points. If player A chose to place trust, player B could choose between honoring ('deliver a high-quality product') or abusing trust ('deliver a low-quality product'). If player B honored trust, both players received a payoff of 3 points. If player B abused trust, player A received 0 points and player B received 4 points. The second version of the game was similar to the first version. The only difference was that player B did not have the option to

honor trust if player A placed trust. This version therefore represents a situation in which a seller (player B) only has a low quality product. Together, the two versions capture the seller heterogeneity that is crucial to our theoretical approach: without it, a negative signal would always mean that the trustee is choosing to cheat in a situation where they could have honored trust and announces this to trustor ahead of time. The right panel of Figure 1 shows the structure of one round of this game without a reputation system. In this figure P, R, S and T represent the material payoffs of the actors. Figure 2 contains the numbers used in the current experiment.

In each round, the pairs played version 1 or two of the game with equal probability ( $\pi = 1/2$ ). Players B were informed about which version of the game they were assigned to at the beginning of the round. Players A never

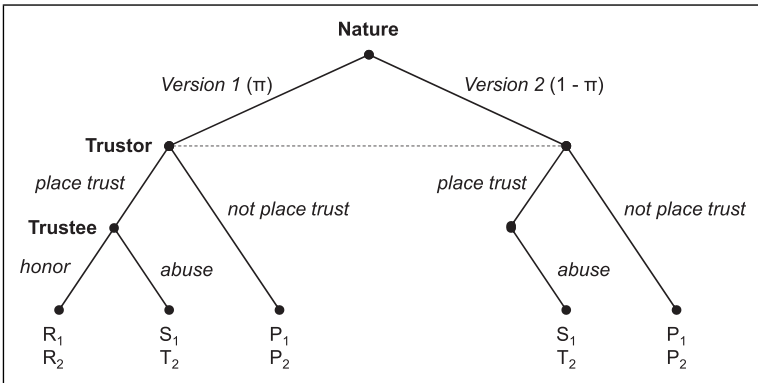


Figure 1. One round of the game used in the experiment, where  $T_2 > R_i > S_1$ .

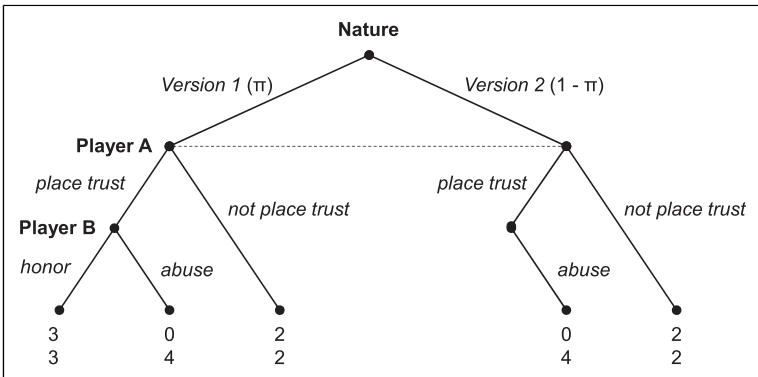


Figure 2. Numerical example of the game used in the experiment.

knew which version they were playing. This heterogeneity introduced uncertainty for players A, since they did not know when a player B had the opportunity to honor trust and when not. This captures the key element characteristic of second-hand markets that a potential buyer does not know whether a seller can deliver a high quality product or only a low quality product. Even when a seller is very trustworthy and always chooses to send a high quality product when they have one available, they may frequently have only low quality products to sell.

The experiment had a two (No communication versus Communication) by two (No reputation versus Reputation) between-subjects design. Each session was assigned to one of four conditions. We refer to the treatment with both communication and reputation as the ‘Combined’ condition. In the two Communication treatments, players B (sellers) could send a binary, non-binding message at no cost to players A (buyers) in which they revealed what their choice would be if player A were to place trust. In the two Reputation conditions, reputation was operationalized as information about previous decisions made by players B. Before choosing to place trust or not, players A were informed about the actual decisions made by player B in the rounds in which player B acted in role B. In the Combined condition players A were also informed about promises made by players B in the past. The complete instructions for the participants can be found in the online [appendix \(B\)](#).

## Procedures

260 subjects participated the experiment in May2020.<sup>11</sup> Nine subjects quit the experiment before the end and were excluded from the analyses.<sup>12</sup> Ten sessions that each lasted for about 75 min were conducted. The numbers of groups and group sizes per condition are shown in [Table 2](#).<sup>13</sup> Subjects were recruited via Prolific.<sup>14</sup> Only subjects with the British nationality who were 18 years and older could sign up. Subjects could participate only once and were paid according to their decisions and decisions made by others in the experiment. For every 15 points earned in the game, the subjects received £1.00.

**Table 2.** Numbers of groups and group sizes per condition.

Condition	Group size (round 1)			Total
	4	6	8	
Control	1	2	6	9
Reputation	0	1	7	8
Communication	1	1	7	9
Combined	3	0	7	10
Total	5	4	27	36



The subjects received on-screen instructions. After reading the instructions they completed some comprehension questions. They could only continue with the experiment after they answered all questions correctly. The experiment was computerized with the software oTree (Chen et al. 2016), commonly used in online networked experiments. Figure B2 in the online appendix displays an example screen participants saw during the experiment. Participation was anonymous and no deception was used.<sup>15</sup> After playing the game, the subjects completed a short survey asking about some demographic characteristics as well as about their experience with game theory and their idea about the goal of the study. 74% of the subjects was female, 20% male, one subject indicated 'other gender' and the remaining subjects did not indicate their gender. The age of the subjects ranged from 18 to 79 (mean = 39). 12% of them indicated to be a student. 8% of the subjects had no degree, or only completed elementary school. About one third completed high school, another third obtained an undergraduate degree and 16% held a graduate degree.

## Variables

Two treatment dummies were constructed that indicated whether a subject participated in a game with or without a reputation system, and whether player B could communicate with player A. The main dependent variable in most of the analyses was the decision of player A whether or not to place trust in player B.

The message sent by player B was operationalized by two dummy variables. These dummies indicated whether player B promised to honor trust or informed player A that they would abuse trust. The reference category was not sending a message. Honest communication was defined as promising to abuse trust, or as promising to honor trust and actually honoring trust. Dishonest communication entailed promising to honor trust, but then abusing trust. When player B promised to honor trust, but did not get the opportunity to prove that they were honest because player A did not place trust, players A could not judge whether that promise was honest or dishonest. Rounds in which the promise could not be classified as honest or dishonest have missing values for this variable. Note that all players B always had the possibility to be honest: either by promising to honor trust and actually honoring trust, or by warning in advance that they would not honor trust.

Player B's reputation was measured by four variables that were constructed in two different ways. The first method included all previous decisions made by a player B in role B in the reputation variables. The first pair of variables counted the number of times player B honored and abused trust in the past. The second pair of reputation variables was only constructed for

subjects in both conditions with communication and counted the number of times player B communicated honestly or dishonestly in the past. The second method accounts for the fact that although players A had access to all previous decisions of player B, they had to scroll down in the history window when they wanted to see more than the last six rounds (see [Figure B2](#) in the online appendix). For this second method, the four reputation variables were constructed in same way, except that only the last six (more readily observable) rounds in which player B acted in role B were included. In every analysis, only one pair of reputation variables is included, to avoid severe multi-collinearity.

## Results of study 2

### *Descriptive statistics*

[Table 3](#) shows the distribution of outcomes across conditions and versions. Trust was much harder to achieve than in the standard trust game used in prior studies. After all, in 50% of the cases (version two of the game) the trustee had no choice but to abuse trust. Moreover, with the exception of the Combined condition, trustees could not build a reputation for honesty. It should be no surprise then that trust levels were overall lower than typically found in regular trust games ([Boero et al., 2009](#); [Buskens et al. 2010](#); [Charness et al., 2011](#)). Placement of trust was comparable between the Control (23%) and Reputation (21%) conditions and between the Communication (33%) and Combined (37%) conditions, but more common in the latter two than in the former two. Similarly, placed trust was honored much more often in the Communication (59%) and Combined (52%) conditions than in the Control (34%) and Reputation (26%) conditions. That is, the ability to build a reputation impacted trust very little but the ability to make a promise boosted trust. Contradicting theoretical predictions, the percentage of cases in which a trustee honored trust was highest in the Communication condition (59%), higher than in the Combined condition (52%). In the Communication condition, most promises were made when the option to honor was indeed available (version 1), and in 136 of the 163 cases in which trust was placed and it could be honored it was in fact honored. For the majority of trustees a promise was a promise, even though it could be broken without consequence.

### *Group-level tests*

H2 states that sellers are more likely to be honest in the presence of a reputation system. [Table 3](#) shows no increase in the fraction of honored promises when going from the Communication condition to the Combined

**Table 3.** Distribution of outcomes across conditions.

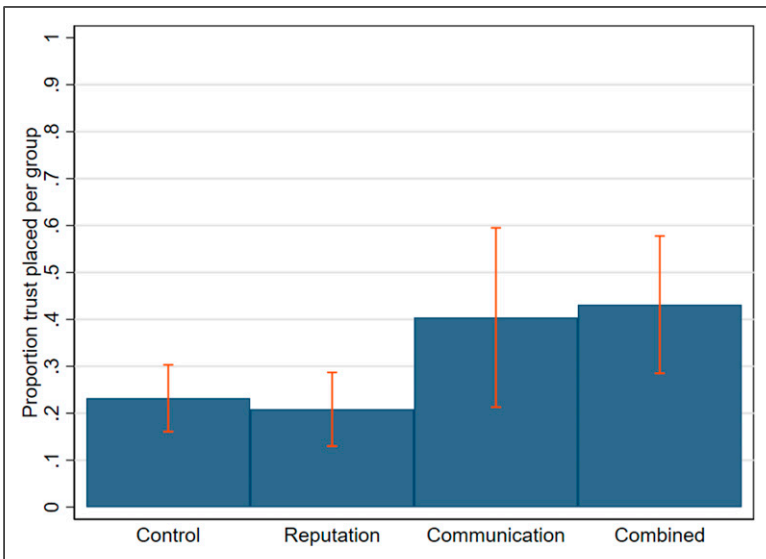
Condition version	Control		Reputation		Communication		Combined	
	1	2	1	2	1	2	1	2
All	523	524	495	445	414	432	508	492
Promised					332	173	401	193
Trust placed	124	116	102	98	163	69	204	87
Trust honored	82		51		136	2	152	5

Note: Cells are empty when outcomes are precluded in the experimental design. The grey numbers in the 'Trust placed' row in the communication and combined conditions correspond to separate cases where trust was placed even though no promise was made or a warning of abuse was issued. [Table C1](#) in the online appendix contains a further breakdown of these numbers.

condition. A rank-sum test confirms that also at the level of (independent) groups there was no significant difference in the fraction of honest messages between the two conditions ( $M_{\text{Communication}} = 0.73$ ,  $M_{\text{Combined}} = 0.71$ ,  $z = 0.089$ ,  $p = 0.929$ ,  $N = 18$ ).<sup>16</sup> We therefore also do not find support for H2.

H4 predicts that trust will be placed more often if both a reputation system and the opportunity to communicate are present. To test H4, we compare the trust level of the Combined condition with each of the other conditions (Figure 3). The trust level is indeed the highest in the Combined condition, although the difference with the Communication condition is very small. For a robust test of the differences, we use three rank-sum tests at the group level. These show that players A placed more trust in groups in the Combined condition than in groups in the Control condition ( $z = 3.136$ ;  $p = .002$ ,  $N = 19$ ) and groups in the Reputation condition ( $z = 2.887$ ,  $p = .004$ ,  $N = 18$ ), but not when compared with groups in the Communication condition ( $z = 0.572$ ,  $p = .567$ ,  $N = 19$ ). We do not find support for H4.

We note, though, that the number of cases at the group level is relatively small. Indeed, an approximate power analysis suggests that with the given number of cases and the observed means and standard deviations, finding a significant difference would require a rather large true effect. It is thus possible that this null finding is the result of a lack of statistical power. For this reason, we also test the hypothesis at the individual level (below). So far,



**Figure 3.** Mean proportion of trust placed per group, by experimental condition ( $N = 36$ ). Error bars represent 95% confidence intervals.

however, the evidence suggests that the ability to make a promise, even if from a self-interest perspective such a promise would be empty, is key to trust, and not the ability to build a reputation for being honest.

It takes time for trustees to build a reputation, so the hypothesized effects should be stronger in the second half of a game. We repeated the rank-sum tests of differences between the level of trust between the conditions, this time only including the second half of each game. The results of these tests are no different from the main tests, suggesting that the lack of an effect of a reputation system cannot be explained by the limited number of reviews accumulated in the first half of a game.

### *Individual-level tests*

H1 and H3 do not have group-level analogues and must be tested at the level of individual decisions. We regressed individual decisions in game rounds. 18% of the variance was explained at the individual level and 2% at the group level.

To test H1 that claims by sellers with a reputation for being honest are trusted more, we estimated a logistic regression model with the decision to place trust in the Combined condition as dependent variable (Table 4, Model 1). To account for interdependencies between observations, we included random intercepts for individual subjects as well as for the group they played in. All analyses include controls for round number (interval measure) and group size. The two honesty-reputation variables that are based on all previous decisions of player B in role B were centered and interacted with two dummy variables that indicated whether player B promised to honor trust in the focal round. The large and significant coefficient of the variable 'Promised to place trust' shows that players A were more likely to place trust in players B who promised to honor trust. The interactions between the promises made by player B and B's reputation for honesty are not significant. Results do not change when the multiple interaction effects are estimated separately in different models. For models that include interaction effects, differences between logistic regression models and linear probability models may arise (Ganzach et al. 2000). In our case the results do not change when running a linear probability model. The results also remain the same when transforming the reputation scores of the trustees to dummy variables that indicate whether the trustee has at least once been (dis)honest in the past. These results do not provide support for H1.

As an additional robustness check, we repeated the analysis with a slightly different calculation of reputation scores. Each player A viewed player B's reputation in a history box on their screen. The last six rounds in which player B acted in role B were always visible, but to see earlier rounds, player A had to scroll down. When using reputation variables that are only based on the last six rounds in which the current player B also acted as player

**Table 4.** Logistic regression models with subject- and group-level random effects.

	Model 1	Model 2	Model 3	Model 4
	DV = trust	DV = honest	DV = trust	DV = trust
<b>Treatment variables</b>				
Reputation condition (Ref. Cat. = control condition)	—	0.73 (0.23)	2.19* (0.86)	0.89 (0.23)
Communication condition	—	—	—	2.10** (0.55)
Combined condition	—	—	—	1.41 (0.52)
<b>Promises made by player B</b>				
Promised to honor trust	7.01*** (1.63)	—	14.85*** (4.11)	—
Promised to abuse trust	0.91 (0.32)	—	1.01 (0.24)	—
<b>Reputation player B</b>				
# Honest	1.09 (0.12)	—	—	—
# Dishonest	0.78 (0.18)	—	—	—
<b>Interaction variables</b>				
Reputation honest * promise to honor	0.98 (0.11)	—	—	—
Reputation dishonest * promise to honor	1.14 (0.27)	—	—	—
Reputation honest * promise to abuse	0.95 (0.12)	—	—	—
Reputation dishonest * promise to abuse	1.18 (0.34)	—	—	—
Reputation treatment * promise to honor	—	—	0.48* (0.14)	—
<b>Control variables</b>				
Round number	0.95*** (0.01)	1.01 (0.008)	0.95*** (0.01)	0.96*** (0.00)
Group size	0.92 (0.09)	1.11 (0.97)	1.01 (0.08)	1.04 (0.06)
Constant	0.69 (0.52)	1.34 (0.94)	0.15** (0.10)	0.29** (0.14)
Number of decisions	1000	863	1849	3849
Number of subjects	61	117	119	238
Number of groups	10	18	19	36

Note: Odds ratios. Standard errors in parentheses. \*\*\* $p < .001$ , \*\* $p < .01$ , \* $p < .05$ .

B, the results do change. The interaction between the number of times player B has been honest and the promise to honor trust turns positive and significant (odds ratio = 1.254,  $z = 2.04$ ,  $p = .041$ ). This means that players A were significantly more likely to place trust in a player B who promised to honor trust when that player B had been more honest in just the last six rounds.

Model two in [Table 4](#) shows the results of a logistic regression of both communication conditions with the honesty of player B as the dependent variable. The dummy variable that indicated if player B promised to honor trust was interacted with the dummy variable for the reputation treatment. The logistic regression results corroborate the findings from the group-level rank-sum test of H2. We find that players B were not more likely to make an honest promise in the presence of a reputation system than when there was no reputation system. The results did not change when fitting a linear probability model. We reject H2.

Model 3 in [Table 4](#) contains the results of a logistic regression of both communication conditions with the decision of player A as the dependent variable. We interacted the variable that indicated if player B promised to honor trust with the reputation treatment variable to test H3. Contrary to theoretical expectations, players A attached less value to B's promise to honor trust when there was a reputation system than when there was no reputation system. When running a linear probability model, we find that the interaction turns insignificant. Neither model supports the hypothesis that the effect of promises to honor trust is stronger in the presence of a reputation system. We also reject H3.

To evaluate H4 at the individual level, we estimated a logistic regression with the variables that indicated to which treatment the subjects were assigned as the independent variables and with the decision of player A as the dependent variable. Model 4 contains the results of this regression. Corroborating the finding from the Wilcoxon rank-sum test at the group level, we find that more trust was placed when players B could make promises to players A than when they could not. The presence of the reputation system did not affect the probability with which player A placed trust. Running a linear probability model instead does not alter the results. We reject H4.

To summarize, support for H1 was weak and inconsistent while H2, H3, and H4 were unequivocally rejected. The results from the experiment do not support the theory.

## Discussion

We are left with a puzzle. Unverifiable assurances about product quality made by sellers in second-hand markets increase the prices buyers are willing to pay (Lewis 2011; Anderson et al., 2007; Rawlins and Johnson 2007; Neto et al. 2016; Snijders and Zijdemans 2004; Heijst et al., 2008;

Sena et al. 2005). Buyers mostly report to be satisfied with sellers' products; strangers thus cooperate successfully in this context. We have firmly rejected the going explanation that this is enabled by reputation systems that provide a necessary rational incentive for sellers to behave honorably so that buyers can trust them. Both in observational data from eBay auctions and controlled data from an original experiment we found that buyers trust quality claims no less when coming from sellers who lack a good reputation and just as much when sellers cannot even build a reputation. Why then do buyers trust promises made by sellers when these commitments can be reneged on without consequence, and why do sellers not take advantage? How was second-hand market platform Craigslist able to reduce solid waste by a third of a pound per capita (Fremstad 2017; Dhanorkar 2019) without a reputation system?

We contend the puzzle arises from an *undersocialized* conception of human behavior that assumes away essential normative behavior in markets. The middle ground between generalized morality and economic utilitarianism (Granovetter 1985) that the new institutionalism and economic sociology have settled on is one that firmly replaces atomism with networks and perfect-with bounded rationality, but often still implicitly maintains economics' self-interested behavior as theory of action except in embedded relations (Granovetter and Swedberg 1992; Fourcade and Healy 2007; Krippner and Alvarez 2007; Zelizer 2012; Hitlin and Vaisey 2013; Nee 2005; Nee and Swedberg 2020; Bandelj 2020). Just as in the new institutionalism in economics, markets are in essence still understood as functioning properly only because individuals' natural inclination to lie and cheat when this pays off is reined in by effective institutions that reward self-control over those urges. We do not wish to bend the stick too far the other way and deny the possibility of occasional deceit by bad actors or suggest that institutions like reputation systems or network embeddedness would never be important for market functioning. Rather, we suggest that there are identifiable limits to what most people are willing to do for selfish gain and that this is common knowledge so that problems that go beyond those limits are solved without institutions, e.g. through socialization into self-respecting individuals. Once identified, these limits will imply conditions under which institutions can be expected to be instrumental for market functioning.

One concrete limit suggested by the present paper is one drawn between simple selfish behavior and being a liar. There is overwhelming evidence that in homogenous markets where suppliers always offer the same product or service, reputation systems do precisely what a theory of self-interest would predict (Boero et al., 2009; Bolton et al. 2004; Charness et al., 2011; Duffy et al. 2013; Fehrler and Przepiorka 2013; Frey and Van De Rijt 2016; Jiao et al., 2021; Teubner et al. 2017; Buskens et al. 2010; Ba and Pavlou 2002). In homogeneous markets, the challenge is to find a seller who will reliably



deliver a good product. Failure to deliver a product or not producing a very good product can happen to a well-intended person. They also happen to sellers who in order to save cost choose not to invest in a better product or a more reliable delivery service and are not thinking too hard about disappointed customers. The problem is one of avoiding such lousy transaction partners. Reputation systems help buyers find sellers who perform better for the same price (Yang 2008; Weigelt and Camerer 1988)

We contrast this with the heterogeneous markets we have studied in this paper. In online auctions of second-hand goods, the challenge is instead to avoid sellers who do something most of us would feel ashamed of: to make a promise with the intention of breaking it. We posit that for this reason reputation systems are much less needed here. Few sellers are willing to go so low and because buyers know this, they trust sellers at their word even if they do not have a reputation for keeping their promises. In the eBay study, quality assurances coming from second-hand car sellers without a good reputation seemed to be believed just as much and we found the vast majority of reviews to be highly positive, also when given to a seller lacking a reputation. Even in the online experiment, where the context of mere game play makes deceit much less depraved, we found that reputations have no noticeable impact on buyers' willingness to trust promises and that a majority of sellers honor their promises.

While the multiple research methods and data sources used in the present study have complementary strengths and as such provide a robust basis for drawing these conclusions, evidence on mechanisms is weak. We cannot rule out that the reason buyers do not combine reputation information and promises made by sellers when making decisions is that they act habitually rather than strategically. Perhaps they simply do not stop to think and realize sellers could lie and get away with it. To anticipate the future behavior of sellers who in turn anticipate future behavior of other buyers, current buyers need to reason multiple steps ahead and believe sellers do too. Similarly, it may be that many sellers do not anticipate the reputational consequences or lack thereof when lying or speaking the truth. Also, when buyers do not condition their behavior on the seller's reputation for honesty, there is no reason for sellers to communicate more honestly, which could also explain why we did not find a difference in the communication of the sellers between the experimental conditions. While failure to reason strategically is an implausible account for behavior in the experiment, as subjects practiced strategic play in the instructions phase prior to the experiment, we cannot rule out that many sellers and buyers in everyday markets for used goods would act normlessly if not for a widespread oblivion about incentives, opportunism, and the strategic importance of reputations. A more qualitative approach including interviews with buyers and sellers could help differentiate between our conjecture of internalized norms, bounded rationality and yet other

behavioral logics that could undergird non-reliance on reputations in the making, believing and honoring of promises in second-hand markets.

### **Acknowledgements**

We thank Vincent Buskens, Werner Raub, and two anonymous reviewers for helpful comments and suggestions.

### **Declaration of conflicting interests**

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### **Funding**

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Dutch Research Council (452-16-002).

### **Data Availability Statement**

The data collected by the authors for this article as well as all code necessary to reproduce the results reported here are available at <https://osf.io/edw25/>

### **ORCID iDs**

Judith Kas  <https://orcid.org/0000-0002-6124-0249>

Rense Corten  <https://orcid.org/0000-0001-8154-9393>

Arnout van de Rijit  <https://orcid.org/0000-0002-4208-3452>

### **Supplemental Material**

Supplemental material for this article is available online.

### **Notes**

1. While it is true that online platforms sometimes offer some protection to buyers (e.g. eBay. “Vehicle purchase protection” <https://pages.motors.ebay.com/buy/purchase-protection/index.html>, accessed 25 August 2022), such protection is typically limited in scope and thus still leaves a considerable risk to the buyer. Even if protected, a rational buyer will still prefer to avoid buying a “lemon”, given high transaction costs involved with reimbursement.
2. eBay. “Selling through Auctions.” <https://www.ebay.com/help/selling/listings/selling-auctions?id=4110>. accessed 19 February 2020.
3. eBay. “Setting a Reserve Price.” <https://www.ebay.com/help/selling/listings/selling-auctions/reserve-prices?id=4143>. Accessed 19 February 2020.

4. The data collected by the authors for this article as well as all code necessary to reproduce the results reported here are available at <https://osf.io/edw25/>.
5. The data collected by Lewis are available at <http://doi.org/10.3886/E112439V1>.
6. The results as reported in [Table 1](#) remain substantively the same when we used constants other than 1.
7. This is what models might show when a variable ‘time’, or number of prior transactions by a seller would be included in the models. However, with the current dataset that is not possible, because it contains auctions that were online during a period of 6 months, so there is limited information on auctions by the same seller at different points in time.
8. [Lewis’ \(2011\)](#) results on the effect of denying problems were mixed: he found that denying certain problems led to higher prices, while denying other problems did not affect prices.
9. Concerns about the interpretation of logistic regression (e.g., [Mood, 2010](#)) have led some to prefer alternatives such as linear probability models (e.g., [Huang, 2022](#)). As an alternative to the logistic regression model (model 4), we also estimated a linear probability model. The results are qualitatively in line with the results of model 4, with the exception of the interaction between positive ratings and problems denied, which is no longer significant. This does not alter our conclusion with regard to Hypothesis 1.
10. This procedure ensured that the number of times all subjects acted in the roles of trustor and trustee did not vary too much at any point in time, but at the same time that subjects did not know for sure what role they would play in the next round. Only when a subject quit the experiment in the middle of a session it was sometimes necessary for the other subjects to play more than six times in each role in a block of 12 rounds.
11. Each session was scheduled at a fixed date and time. Subjects could sign up for a specific session through a link that was posted on Prolific. Subjects who finished the instructions, but who could not be assigned to a group received a show-up fee of £2.50.
12. A total of 260 subjects started the experiment. Eight of them chose to quit the experiment before the end and nine lost the connection with the server. If a participant quit the experiment, the participant they were matched with also had to stop, because the experiment is conducted in groups with an even number of subjects. All data from subjects who decided to quit are excluded from the experiment. For the other subjects the data from the rounds that they completed were included. If the remainder of the group consisted of only two members, the data of the remainder of the rounds are excluded from the analyses.
13. Because the number of subjects per group could potentially change due to dropout, [Table 3](#) shows the starting sizes in round 1.
14. See [www.prolific.com](http://www.prolific.com). The experiment was initially designed to be conducted in a laboratory, but due to the Covid19-crisis we diverted to an online solution. While laboratory experiments allow for more control over the behavior of the

subjects during the experiment, the online setting provided a higher level of anonymity, since subjects did not see each other before and after the experiment. Since the theory assumes that buyers and sellers are strangers, the online setting provides a cleaner test of the theory. Moreover, the subject pool of Prolific is more varied in terms of a number of demographic characteristics compared to the average laboratory student pool.

15. In order to pay the subjects according to their earnings in the game, we had to collect their Prolific participant IDs. After completing the cleaning of the data, these IDs are stored separately from the experimental data.
16. In one group in the Communication condition there were not enough messages sent for it to be included in this analysis.

## References

- Akerlof G (1970) The market for lemons: qualitative uncertainty and the market mechanism. *Quarterly Journal of Economics* 84(3): 235–251.
- Anderson ST, Friedman D, Milam GH, et al. (2007) Seller strategies on ebay: does size matter? *International Journal of Electronic Business* 5(6): 643–669.
- Ashton MC and Lee K (2007) Empirical, theoretical, and practical advantages of the HEXACO model of personality structure. *Personality and Social Psychology Review* 11(2): 150–166.
- Axelrod R and Dion D (1988) The further evolution of cooperation. *Science* 242(4884): 1385–1390.
- Ba S and Pavlou PA (2002) Evidence of the effect of trust building technology in electronic markets: price premiums and buyer behavior. *MIS Quarterly* 26(3): 243–268.
- Baldassarri D (2015) Cooperative networks: altruism, group solidarity, reciprocity, and sanctioning in ugandan producer organizations. *American Journal of Sociology* 121(2): 355–395.
- Balliet D (2010) Communication and cooperation in social dilemmas: a meta-analytic review. *Journal of Conflict Resolution* 54(1): 39–57.
- Bandelj N (2020) Relational work in the economy. *Annual Review of Sociology* 46: 251–272.
- Bicchieri C and Lev-On A (2007) Computer-mediated communication and cooperation in social dilemmas: an experimental analysis. *Politics, Philosophy & Economics* 6(2): 139–168.
- Boero R, Bravo G, Castellani M, et al. (2009) Reputational cues in repeated trust games. *The Journal of Socio-Economics* 38(6): 871–877.
- Bolton GE, Katok E and Ockenfels A (2004) How effective are electronic reputation mechanisms? an experimental investigation. *Management Science* 50(11): 1587–1602.
- Bridges J and Vásquez C (2018) If nearly all airbnb reviews are positive, does that make them meaningless? *Current Issues in Tourism* 21(8): 2057–2075.

- Brosig-Koch J and Heinrich T (2018) The role of communication content and reputation in the choice of transaction partners. *Games and Economic Behavior* 112: 49–66.
- Buchan NR, Croson RTA and Dawes RM (2002) Swift neighbors and persistent strangers: a cross-cultural investigation of trust and reciprocity in social exchange. *American Journal of Sociology* 108(1): 168–206.
- Buskens V (2003) Trust in triads: effects of exit, control, and learning. *Games and Economic Behavior* 42(2): 235–252.
- Buskens V, Raub W and van der Veer J (2010) Trust in triads: an experimental study. *Social Networks* 32(4): 301–312.
- Cason TN and Mui VL (2014) Coordinating resistance through communication and repeated interaction. *The Economic Journal* 124(574): 226–256.
- Charness G, Du N and Yang CL (2011) Trust and trustworthiness reputations in an investment game. *Games and Economic Behavior* 72(2): 361–375.
- Chen DL, Schonger M and Wickens C (2016) oTree—an open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance* 9: 88–97.
- Cook KS, Hardin R and Levi M (2005) *Cooperation without Trust?* New York: Russell Sage Foundation.
- Dasgupta P (1988) Trust as a commodity. In: Gambetta D (ed), *Trust: Making an Breaking Cooperative Relations*. Cornwall: Basil Blackwell. 49–72.
- Delhey J, Newton K and Welzel C (2011) How general is trust in “most people”? solving the radius of trust problem. *American Sociological Review* 76(5): 786–807.
- Denant-boemont L, Masclet D and Noussair C (2011) Announcement, observation and honesty in the voluntary contributions game. *Pacific Economic Review* 16(2): 207–228.
- Depken CA and Gregorius B (2010) Auction characteristics, seller reputation, and closing prices: evidence from ebay sales of the iphone. *International Journal of Electronic Business* 8(2): 170–186.
- Dhanorkar S (2019) Environmental benefits of internet-enabled C2C closed-loop supply chains: a quasi-experimental study of craigslist. *Management Science* 65(2): 660–680.
- Diekmann A, Jann B, Przepiorka W, et al. (2014) Reputation formation and the evolution of cooperation in anonymous online markets. *American Sociological Review* 79(1): 65–85.
- Duffy J and Feltovich N (2006) Words, deeds, and lies: strategic behaviour in games with multiple signals. *Review of Economic Studies* 73(3): 669–688.
- Duffy J, Xie H and Lee YJ (2013) Social norms, information, and trust among strangers: theory and evidence. *Economic Theory* 52(2): 669–708.
- Edelman B and Luca M (2014) Digital discrimination: the case of airbnb.com. harvard business school nom unit working paper. Available at: [http://www.west-info.eu/files/airbnb\\_research.pdf](http://www.west-info.eu/files/airbnb_research.pdf).

- Erikson E, Bearman P and Emily (2006). *American Journal of Sociology* 112(1): 195–230.
- Fehrler S and Przepiorka W (2013) Charitable giving as a signal of trustworthiness: disentangling the signaling benefits of altruistic acts. *Evolution and Human Behavior* 34(2): 139–145.
- Fourcade M and Healy K (2007) Moral views of market society. *Annual Review of Sociology* 33: 285–311.
- Fremstad A (2017) Does craigslist reduce waste? evidence from California and Florida. *Ecological Economics* 132: 135–143.
- Frey V and van de Rijt A (2016) Arbitrary inequality in reputation systems. *Scientific Reports* 6(1): 38304–38305.
- Gambetta D (2009) Signaling. In: *The Oxford Handbook of Analytical Sociology*. Oxford: Oxford University Press. 168–194.
- Ganzach Y, Saporta I and Weber Y (2000) Interaction in linear versus logistic models: a substantive illustration using the relationship between motivation, ability, and performance. *Organizational Research Methods* 3(3): 237–253.
- Granovetter M (1985) Economic action and social structure: the problem of embeddedness. *American Journal of Sociology* 91(3): 481–510.
- Granovetter M and Swedberg R (eds), (1992) *The Sociology of Economic Life*. Oxford: Westview Press.
- Greif A (1989) Reputation and coalitions in medieval trade: evidence on the maghribi traders. *The Journal of Economic History* 49(4): 857–882.
- Heijst D, Potharst R and van Wezelvan Wezel M (2008) A support system for predicting ebay end prices. *Decision Support Systems* 44(4): 970–982.
- Hitlin S and Vaisey S (2013) The new sociology of morality. *Annual Review of Sociology* 39: 51–68.
- Houser D and Wooders J (2006) Reputation in auctions: theory, and evidence from ebay. *Journal of Economics and Management Strategy* 15(2): 353–369.
- Huang FL (2022) Alternatives to logistic regression models in experimental studies. *The Journal of Experimental Education* 90(1): 213–228.
- Iacono SL and Sonmez B (2020) The effect of trusting and trustworthy environments on the provision of public goods. *European Sociological Review* 37(1): 1–14.
- JiaoJiao R, PrzepiorkaPrzepiorka W and Buskens V (2021) Reputation effects in peer-to-peer online markets: a meta-analysis\*. *Social Science Research* 95: 102522.
- Jin GZ and Kato A (2006) Price, quality, and reputation: evidence from an online field experiment. *The RAND Journal of Economics* 37(4): 983–1005.
- Kas J, Corten R and van de Rijt A (2022) The role of reputation systems in digital discrimination. *Socio-Economic Review* 20(4): 1905–1932. DOI: [10.1093/ser/mwab012](https://doi.org/10.1093/ser/mwab012).
- Kollock P (1994) The emergence of exchange structures: an experimental study of uncertainty, commitment, and trust. *American Journal of Sociology* 100(2): 313–345.

- Kollock P (1999) The production of trust in online markets. *Advances in Group Processes* 16(1): 99–123.
- Kreps DM (1996) Corporate Culture and Economic Theory. In: Alt J and Shepsle K (eds), *Firms, Organizations and Contracts*. Oxford: Oxford University Press. 175–221.
- Krippner GR and Alvarez AS (2007) Embeddedness and the intellectual projects of economic sociology. *Annual Review of Sociology* 33: 219–240.
- Kuwabara K (2015) Do reputation systems undermine trust? divergent effects of enforcement type on generalized trust and trustworthiness. *American Journal of Sociology* 120(5): 1390–1428.
- Lewis G (2011) Asymmetric information, adverse selection and online disclosure: the case of ebay motors. *American Economic Review* 101(4): 1535–1546.
- Lewis G (2019) *Replication Data for: Asymmetric Information, Adverse Selection and Online Disclosure: The Case of EBay Motors*. Nashville: TN: American Economic Association.
- Macy MW and Skvoretz J (1998) The evolution of trust and cooperation between strangers: a computational model. *American Sociological Review* 63(5): 638–660.
- Melnik MI and Alm J (2005) Seller reputation, information signals, and prices for heterogeneous coins on ebay. *Southern Economic Journal* 72(2): 305–328.
- Molm LD, Takahashi N and Peterson G (2000) Risk and trust in social exchange: an experimental test of a classical proposition. *American Journal of Sociology* 105(5): 1396–1427.
- Mood C (2010) Logistic regression: why we cannot do what we think we can do, and what we can do about it. *European Sociological Review* 26(1): 67–82.
- Nee V (2005) The new institutionalisms in economics and sociology. In: Smelser NJ and Swedberg R (eds), *The Handbook of Economic Sociology*. Princeton, NJ: Princeton University Press. pp. 49–74.
- Nee V and Swedberg R (eds), (2020) *The Economic Sociology of Capitalism*. Princeton, NJ: Princeton University Press.
- Frota Neto JQ, Bloemhof J and Corbett C (2016) Market prices of remanufactured, used and new items: evidence from ebay. *International Journal of Production Economics* 171(3): 371–380.
- North DC (1990) *Institutions, Institutional Change and Economic Performance*. Cambridge: Cambridge University Press.
- Powell WW (1990) Neither market nor hierarchy: network forms of organization. *Research in Organizational Behavior* 12: 295–336.
- Przepiorka W and Berger J (2017) *Signalling Theory Evolving: Signals and Signs of Trustworthiness in Social Exchange Social Dilemmas, Institutions and the Evolution of Cooperation*. Berlin: De Gruyter Oldenbourg. 373–392.
- Przepiorka W and Diekmann A (2013) Temporal embeddedness and signals of trustworthiness: experimental tests of a game theoretic model in the United

- Kingdom, Russia, and Switzerland. *European Sociological Review* 29(5): 1010–1023.
- Przepiorka W, Norbutas L and Corten R (2017) Order without law: reputation promotes cooperation in a cryptomarket for illegal drugs. *European Sociological Review* 33(6): 752–764.
- Raub W (2004) Hostage posting as a mechanism of trust: binding, compensation, and signaling. *Rationality and Society* 16(3): 319–365.
- Raub W and Weesie J (1990) Reputation and efficiency in social interactions: an example of network effects. *American Journal of Sociology* 96(3): 626–654.
- Rawlins C and Johnson P (2007) Selling on ebay: persuasive communication advice based on analysis of auction item descriptions. *Journal of Strategic E-Commerce* 5(2): 75–82.
- Resnick P and Zeckhauser R (2002) Trust among strangers in internet transactions: empirical analysis of ebay's reputation system. *Advances in Applied Microeconomics* 11(2): 127–157.
- Resnick P, Kuwabara K, Zeckhauser R, et al. (2000) Reputation systems. *Communications of the ACM* 43(12): 45–48.
- Sabin N and Reed-Tsochas F (2020) Able but unwilling to enforce: cooperative dilemmas in group lending. *American Journal of Sociology* 125(6): 1602–1667.
- Sena MP, Heath CE and Webb M (2005) The impact of ebay ratings and item descriptions on auction prices: a comparison of designer watches and DVDs. *Contemporary Research in E-Marketing* 1: 149–165.
- Simpson B and Willer R (2015) Beyond altruism: sociological foundations of cooperation and prosocial behavior. *Annual Review of Sociology* 41(1): 43–63.
- Snijders C and Zijdemans R (2004) Reputation and internet auctions: ebay and beyond. *Analyse & Kritik* 26(1): 158–184.
- Sorenson O (2014) Status and reputation: synonyms or separate concepts? *Strategic Organization* 12(1): 62–69.
- Tadelis S (2016) Reputation and feedback systems in online platform markets. *Annual Review of Economics* 8(1): 321–340.
- ter Huurne M, Ronteltap A, Corten R, et al. (2021a) Antecedents of Trust in the Sharing Economy: A Systematic Review. *Journal of Consumer Behaviour* 16(6) Inderscience Publishers (IEL): 485–498
- Teubner T and Glaser F (2018) Up or out: the dynamics of star rating scores on airbnb. In: ECIS 2018 Proceedings, 2018, June 23–28 2018, Portsmouth, UK.
- Teubner T, Hawlitschek F and Dann D (2017) Price determinants on airbnb: how reputation pays off in the sharing economy. *Journal of Self-Governance and Management Economics* 5(4): 53–80.
- Thiemann M and Lepoutre J (2017) Stitched on the edge: rule evasion, embedded regulators, and the evolution of markets. *American Journal of Sociology* 122(6): 1775–1821.
- Tjaden JD, Schwemmer C and Khadjavi M (2018) Ride with me-ethnic discrimination, social markets. *European Sociological Review* 34(4): 418–432.



- Uzzi B (1996) The sources and consequences of embeddedness for the economic performance of organizations: the network effect. *American Sociological Review* 61(4): 674–986.
- Weigelt K and Camerer C (1988) Reputation and corporate strategy: a review of recent theory and applications. *Strategic Management Journal* 9(5): 443–454.
- Wilson RK and Sell J (1997) “Liar, Liar...”: cheap talk and reputation in repeated public goods settings. *Journal of Conflict Resolution* 41(5): 695–717.
- Wrong DH (1961) The oversocialized conception of man in modern sociology. *American Sociological Review* 26(2): 183–193.
- Yamagishi T and Yamagishi M (1994) Trust and commitment in the United States and Japan. *Motivation and Emotion* 18(2): 129–166.
- Yamagishi T, Cook KS and Watabe M (1998) Uncertainty, trust, and commitment formation in the United States and Japan. *American Journal of Sociology* 104(1): 165–194.
- Yang K (2008) Review: cooperation without trust? *Public Administration Review* 68(6): 1164–1166.
- Zelizer VA (2012) How I became a relational economic sociologist and what does that mean? *Politics & Society* 40(2): 145–174.
- Zucker LG (1986) Production of trust: institutional sources of economic structure, 1840–1920. *Research in Organizational Behavior* 8: 53–111.