

# Reputation Formation and the Evolution of Cooperation in Anonymous Online Markets

Andreas Diekmann,<sup>a</sup> Ben Jann,<sup>b</sup>  
Wojtek Przepiorka,<sup>c</sup> and Stefan Wehrli<sup>a</sup>

## Abstract

Theoretical propositions stressing the importance of trust, reciprocity, and reputation for cooperation in social exchange relations are deeply rooted in classical sociological thought. Today's online markets provide a unique opportunity to test these theories using unobtrusive data. Our study investigates the mechanisms promoting cooperation in an online-auction market where most transactions can be conceived as one-time-only exchanges. We first give a systematic account of the theoretical arguments explaining the process of cooperative transactions. Then, using a large dataset comprising 14,627 mobile phone auctions and 339,517 DVD auctions, we test key hypotheses about the effects of traders' reputations on auction outcomes and traders' motives for leaving feedback. Our statistical analyses show that sellers with better reputations have higher sales and obtain higher prices. Furthermore, we observe a high rate of participation in the feedback system, which is largely consistent with strong reciprocity—a predisposition to unconditionally reward (or punish) one's interaction partner's cooperation (or defection)—and altruism—a predisposition to increase one's own utility by elevating an interaction partner's utility. Our study demonstrates how strong reciprocity and altruism can mitigate the free-rider problem in the feedback system to create reputational incentives for mutually beneficial online trade.

## Keywords

cooperation, trust, reciprocity, reputation, online markets

Social exchange has always been an important part of human sociality and arguably a driving force in the evolution of the human mind (Cosmides and Tooby 1992). In archaic societies, social exchange spanned a network of multi-plex relations and was accompanied by elaborate ceremonies. A long tradition of research in anthropology and sociology on social exchange began with the work of Malinowski and Mauss. Ceremonial exchanges, such as the Kula (Malinowski 1922) or the potlatch (Mauss [1950] 1990), were formal rituals characterized by solemnity, decorum, and disinterested generosity; they occurred side by side

with economic exchanges, which, in contrast, were guided by utilitarian motives (Heath 1976). Social exchange was governed by norms of reciprocity (Gouldner 1960; Malinowski 1926; Mauss [1950] 1990), and a

---

<sup>a</sup>ETH Zurich

<sup>b</sup>University of Bern

<sup>c</sup>Nuffield College and University of Oxford

## Corresponding Author:

Wojtek Przepiorka, University of Oxford,  
Department of Sociology, Manor Road, Oxford,  
OX1 3UQ, United Kingdom  
E-mail: [wojtek.przepiorka@sociology.ox.ac.uk](mailto:wojtek.przepiorka@sociology.ox.ac.uk)

failure to reciprocate was punished by the loss of reputation and status (Blau 1964; Mauss [1950] 1990).

With trade taking place over ever longer distances, markets transcended the borders of small communities, and historical evidence shows that social and economic institutions evolved as a substitute for informal social control (Greif 1989; Zucker 1986). For example, trade in the early Middle Ages in Europe was characterized by geographic specialization, bookkeeping, and cashless payment. At that time, the Champagne fairs in France were a meeting point for traders from all over Europe. Milgrom, North, and Weingast (1990) discuss the emergence of a private adjudication system (the Law Merchant) that helped overcome trust problems among anonymous traders. Administered by private judges, it provided a platform for traders to settle disputes and document trading partners' dishonest behavior. Along with a system of notaries, such an institution could track and disseminate information about traders' past behavior so that dishonest traders could be denied access to the market.

More recently, similar information-collecting institutions have emerged to overcome trust problems between money lenders and borrowers. Credit bureaus, for example, function as information brokers that collect and collate information about borrowers' liabilities, credit histories (in particular, arrears and defaults), and other characteristics (Japelli and Pagano 2002). Credit bureaus are private institutions, often operated by a group of lenders, and reporting of borrower data is voluntary. To overcome the free-rider problem in data reporting, credit bureaus' services are based on reciprocity; only lenders who submit accurate information about their customers are granted access to a comprehensive customer database. This reciprocal information-sharing system creates financial incentives for lenders to contribute to the common good of customer information and for borrowers to maintain a good reputation by timely debt repayment. Japelli and Pagano (2002) show, moreover, that such information sharing is

indeed associated with higher lending and lower default rates across 46 countries.

In a similar vein, cooperation in peer-to-peer online markets is sustained by an electronic reputation system that collects and disseminates information about traders' past behavior (Kollock 1999; Resnick et al. 2000). In online markets, tens of thousands of anonymous buyers and sellers trade with each other every day. These traders are not connected through a social network. In most cases they interact only once (Resnick and Zeckhauser 2002), and—when cheating—they are unlikely to be caught and fined by a central authority because many of these transactions transcend national borders (Przepiorka 2013). Traders' willingness to submit feedback after transactions are finished is crucial for the functioning of the reputation system. However, unlike in credit markets, information about traders' reputations is a collective good freely available to anyone (Bolton, Katok, and Ockenfels 2004). Submitting feedback has a cost in terms of time and effort, so traders have no real incentive to submit feedback and a second-order free-rider problem may exist (Heckathorn 1989). Assuming strict rationality, actors would not rate their trading partners and, consequently, the whole market would degenerate or not evolve at all. How is it, then, that online markets are so successful and enjoy increasing popularity?

Here we argue that other-regarding preferences, such as altruism (Becker 1976) and strong reciprocity (Gintis 2000), play an important role in maintaining the provision of feedback in anonymous online markets. This feedback then generates the information necessary to create financial incentives for mutually beneficial trade. Based on a large sample of process-produced auction data obtained from eBay ( $N \approx 350,000$ ), we first corroborate previous studies' findings that reputation has a market value, thus making it rational for traders to build and maintain a good reputation. We investigate the effect of reputation on the probability of sale and the selling price in a high- and a low-cost product market (mobile phones versus DVDs, respectively)

and under high and low buyer uncertainty (used versus new mobile phones, respectively). Second, we investigate traders' motivation to leave feedback, without which reputational incentives would not accrue. Based on a large sample of timed rating events, we test partly competing hypotheses with regard to the possible motives behind traders' voluntary contributions to the feedback system. Our analyses take advantage of the repeated measures obtained for the same traders, to control for unobserved confounders (i.e., unobserved heterogeneity).

Our analysis brings together different strands of research and describes theoretically and empirically the mechanisms that guide traders' cooperative behavior in an online market. In particular, we show how the institutional set-up of an online market engages traders' moral sentiments and material interests to create opportunities for mutually beneficial trade.

## THEORY AND HYPOTHESES

An interaction between a buyer and a seller in a typical online-auction market can be conceived as a process comprising four stages. First (the *initiation stage*), a seller decides to initiate an auction and sets some relevant parameters. Second (the *bidding stage*), potential buyers place bids and the highest bidder gets the offered item. Third (the *transaction stage*), the buyer pays the highest bid (i.e., selling price) and the seller ships the item. Finally (the *feedback stage*), the buyer can rate the seller and, depending on implementation of the feedback mechanism, the seller can rate the buyer. Ratings can be positive or negative, are typically accompanied by a short written comment, and are immediately available to all market participants free of charge.

### *Transaction Stage*

The sequence of moves at the transaction stage is not predefined. However, because sellers cannot choose their buyers, most sellers demand advance payment and ship the item

only after payment is received (Diekmann, Jann, and Wyder 2009). In this case, buyers cannot inspect the product before payment and a trust problem arises: sellers could deliver poor quality goods or keep the money without delivering at all (Akerlof 1970; Dasgupta 1988; Güth and Ockenfels 2003).<sup>1</sup> How do traders overcome this trust problem?

Evidence from laboratory experiments suggests there may be a significant proportion of trustworthy sellers who care about the outcome of their trading partners due to other-regarding motives such as positive reciprocity, fairness considerations, and altruism (Camerer 2003). However, these trustworthy sellers will be driven out of the market because they cannot keep up with dishonest sellers' cheap offers. Bad experiences with dishonest sellers will decrease buyers' willingness to pay for a product, and thereby reduce trustworthy sellers' incentives to stay in the market (Akerlof 1970; Yamagishi et al. 2009). Alternatively, sellers may be trustworthy because they are embedded in dyadic relations or social networks (Buskens and Raub 2002; Granovetter 1992; Gulati and Gargiulo 1999; Jones, Hesterly, and Borgatti 1997) in which cooperation is enforced through direct and indirect reciprocity (Nowak and Sigmund 2005; see also Blau 1964), respectively.

Online consumer markets, however, merely coordinate the supply and demand of end-product sellers and ultimate buyers. These markets are not characterized by customized and complex exchanges, and, as we will show, repeated interactions between the same two traders are infrequent. Under these conditions, it is unlikely that network governance structures will emerge to solve potential exchange problems (Jones et al. 1997). Moreover, because online trade takes place across large geographic distances, offline commitment relations (Brown, Falk, and Fehr 2004; Kollock 1994; Yamagishi, Cook, and Watabe 1998) or trust networks (Cook, Rice, and Gerbasi 2004; DiMaggio and Louch 1998), which would allow traders to obtain and disseminate information about their interaction partners, are unlikely to form. Finally,

although online markets such as eBay try to establish norms of good conduct by providing discussion forums, chat rooms, and clubs where traders have the opportunity to exchange information on selected topics, to our knowledge, there is no evidence that eBay traders have developed shared social norms of cooperation. Even if they have, such norms will not prevent fraudulent traders from behaving opportunistically.<sup>2</sup>

In anonymous online markets, the lack of social embeddedness is compensated for through an electronic rating system that efficiently and systematically disseminates information about traders' reputations (Resnick et al. 2000). From an organizational behavior perspective, an electronic reputation system is an exogenously established, fully connected network that embeds anonymous online traders in weak ties, through which information about their reputations is transmitted, without structurally limiting their choice of potential interaction partners and the pursuit of their goals (Gulati and Gargiulo 1999; Uzzi 1997). In other words, in an anonymous online market with a reputation system, traders can be conceived as structurally homogeneous and their interactions as isolated dyads (Granovetter 1992).

The resulting social mechanism that resolves exchange problems is rather simple. With their reputations at stake, rational and self-regarding sellers who sufficiently care about their future business have a strong incentive to behave cooperatively. Even in interactions with buyers whom they are unlikely to meet again, these sellers will forgo the short-term temptation to abuse a buyer's trust to avoid a negative rating that would hamper future business. Market entrants, however, have no reputation and are therefore likely to be distrusted by buyers. Consequently, to enter the market, sellers with no feedback record have to lower prices to compensate buyers for the risk they take buying from a new seller. If new sellers sufficiently care about future business, their good reputations will reimburse them for this initial investment in the future (Przepiorka 2013; Resnick and Zeckhauser 2002; Shapiro

1983).<sup>3</sup> If buyers discriminate among sellers according to their reputations, selling prices will be positively (negatively) correlated with a seller's number of positive (negative) ratings. Moreover, if the market does not clear because supply exceeds demand, an analogous hypothesis can be derived with respect to sellers' positive (negative) ratings and the probability of sale.

Previous studies have found a stronger absolute effect of negative ratings than positive ratings on outcomes (Ba and Pavlou 2002), possibly because negative ratings carry more information—they express unmet expectations in relation to proper business practice—whereas positive ratings merely confirm expectations (Lucking-Reiley et al. 2007). Finally, with both new and used products, a potential buyer may fear that a seller will not deliver after payment or that the quality of the item will be lower than advertised. In the case of used products, uncertainty is higher for the buyer because description of the product condition is less standardized, leaving more room for dishonest behavior by sellers. For used products, a seller's reputation may thus be more important and exhibit a larger effect than in the case of new products (Cook et al. 2004; Kollock 1994).

In summary, our hypotheses for the transaction stage are as follows:

*Hypothesis 1:* The number of positive (negative) ratings increases (decreases) the probability of a sale (1a) and the selling price (1b).

*Hypothesis 2:* The effects of negative ratings on the probability of a sale and the selling price are stronger than the effects of positive ratings.

*Hypothesis 3:* The effects of positive and negative ratings are stronger for used products than for new products.

### *Feedback Stage*

We argued that the market value of a good reputation creates financial incentives for traders to behave cooperatively. However, rational and self-regarding traders will not

provide feedback after a transaction is finished and thus reputational incentives will not accrue. Participation in the feedback system can be compared to voting, in that, just as with democratic elections, there is a collective good problem (Downs 1957). If citizens eschew the costs of voting, as strict rationality principles predict, the democratic system collapses, as would the online market should no buyer provide feedback. Fortunately, however, a substantial proportion of citizens do vote, and a large proportion of traders do leave feedback. Similar explanations can be given for both behaviors.

Consider a one-sided reputation system in which only buyers can rate sellers after a transaction. There are several possible explanations for buyers' participation in the feedback system. First, as suggested by signaling theory (e.g., Gambetta 2009), a buyer might leave feedback to gain a reputation for being an active rater. Such a reputation signals to sellers that the buyer will publicly comment on how a transaction plays out, which might enforce cooperative behavior by the seller. Second, buyers might submit positive feedback because they are altruistic and the increase in the seller's utility elevates their own utility (Becker 1976).<sup>4</sup> Because we assume a decreasing marginal utility of positive ratings, the likelihood of an altruistic buyer submitting feedback will decrease with an increasing seller score. Third, some buyers might leave feedback because they gain satisfaction from contributing to the public good provided by the feedback system. This idea corresponds to the notion of "warm glow" altruism (Andreoni 1990; see also Dellarocas, Fan, and Wood 2004). Finally, there is good reason to believe that some buyers are willing to punish defection or reward cooperation without direct benefit. That is, giving feedback can be motivated by strong reciprocity (Fehr, Fischbacher, and Gächter 2002; Gintis 2000).<sup>5</sup> Returning a favor with a favor and responding to misdeeds with sanctions is observed in experiments and seems to be rooted in human nature (de Quervain et al. 2004).

Until May 2008, eBay.de had a feedback system in which both transaction partners

could leave feedback within 90 days. In such a two-sided rating system, in addition to the motives listed above, the feedback stage is open to strategic thinking. Although both buyers and sellers can be led by strategic (and nonstrategic) motives, it is primarily sellers who have an interest in acting strategically at the feedback stage (see also Bolton, Greiner, and Ockenfels 2013). For example, the best way for a seller to establish a good reputation is to engage in good conduct, but there is always a small probability that a buyer is dissatisfied for reasons beyond the seller's control. To safeguard against a buyer's (possibly unjustified) negative response, a seller will prefer to hold up the threat of retaliation. Because a negative rating is more detrimental for sellers who have not yet established a good reputation than for sellers with many positive ratings, we expect sellers with fewer positive ratings to be more likely to strive for the second-mover position at the feedback stage. On the other hand, a seller could try to elicit positive feedback from a buyer by giving positive feedback first. Such a strategy makes sense if the seller believes the buyer adheres to a norm of reciprocity (Gouldner 1960; Resnick and Zeckhauser 2002). If sellers make such an assertion, then the buyer's score will have a positive effect on the probability of the seller giving feedback first. A further variable likely to influence a seller's rating decision is whether the seller has already received a rating from the same buyer in a previous interaction. At the time of our data collection, repeated ratings from the same user did not affect the reputation score. Hence, a seller had no reason to elicit further positive feedback from a buyer by giving positive feedback first (see also Dellarocas et al. 2004).

For buyers, the nonstrategic motives mentioned earlier are still valid. In addition, if buyers are motivated by strong reciprocity, positive feedback from the seller will further increase their likelihood of submitting a positive rating. In fact, it is hard to think of reasons why a self-regarding buyer would reciprocate a seller's positive feedback. Such behavior would be an indication of strong reciprocity at the feedback stage. Furthermore, indirect

reciprocity (Nowak and Sigmund 2005), that is, rewarding sellers for their participation in the feedback system, might be another non-strategic motive in a two-sided system. According to indirect reciprocity, buyers will be more likely to give (positive) feedback to sellers with a good reputation.<sup>6</sup> Finally, due to eBay's reputation system ignoring repeated ratings, buyers should be less likely to submit positive feedback if they provided a rating to the same seller in a previous interaction. If no such effect exists, altruistic motives based on the buyer's belief that the seller gains utility from receiving feedback (Becker 1976) can be ruled out. For our empirical analysis, we confine ourselves to the following hypotheses (see also Dellarocas et al. 2004):

*Hypothesis 4: Reciprocity:*

If buyers and sellers are motivated by strong reciprocity at the feedback stage, their inclination to submit a positive (negative) rating increases after receiving positive (negative) feedback from the transaction partner.

*Hypothesis 5: Effect of seller's score on buyer's behavior:*<sup>7</sup>

- (a) If buyers are motivated by Becker's altruism, the seller's score should have a negative effect on the probability of feedback submission.
- (b) If buyers are motivated by indirect reciprocity, the seller's score should have a positive effect on the buyer's decision to provide feedback.

*Hypothesis 6: Sellers' strategic behavior:*

- (a) Sellers' inclination to make the first move at the feedback stage increases with their own score; and
- (b) Sellers' inclination to make the first move at the feedback stage increases with the buyer's score.

*Hypothesis 7: Previous interaction:*

- (a) Sellers are less likely to submit positive feedback if they have already received a rating from the same buyer in a previous interaction.

- (b) Buyers are less likely to submit positive feedback if they provided a rating to the same seller in a previous interaction.

## PREVIOUS FINDINGS

Most studies investigating reputation effects in online markets analyze data collected from eBay, and most of these studies find evidence of a positive relation between sellers' reputations and prices or sales. Due to space limitations, we provide a review of these studies in the online supplement (<http://asr.sagepub.com/supplemental>) (for earlier reviews, see Bajari and Hortacsu 2004; Resnick et al. 2006). Here, we instead focus on important aspects in the ongoing discussion about reputation formation in anonymous online markets.

In their review of 15 empirical eBay studies, Resnick and colleagues (2006) point out that these studies' quasi-experimental designs cannot account for unobserved heterogeneity in a satisfactory way. If unobserved factors, such as pre-purchase communication between buyers and sellers or the quality of the product description, are correlated with sellers' reputations, then estimates of the regression coefficients can be biased. In our dataset, we have repeated measures for a considerable proportion of both buyers and sellers. This allows us to deal with problems of unobserved heterogeneity by using fixed-effects regressions (Allison 2009).

Some researchers argue that most eBay studies claim to estimate buyers' willingness to pay for a seller's reputation but, in fact, only estimate what buyers do pay, which is a poor estimate of their willingness to pay (Snijders and Weesie 2009). The argument is that the highest bids may depend on many other factors not observable in data obtained from eBay. For instance, bidding is a dynamic process and potential buyers' bidding decisions also depend on their beliefs about other buyers' reservation prices. Moreover, potential buyers could consider several offers simultaneously, but their initial choice sets are unobserved or difficult to define (but see Livingston 2003). In our study, we do not claim to be estimating buyers' willingness to pay for a

seller's reputation. For the sake of our argument, it is enough to show that buyers, on average, pay more for sellers with higher reputations and thus create financial incentives for sellers to behave cooperatively in the market.

The Achilles' heel of a reputation system remains market participants' willingness to rate each other. If participants do not provide truthful ratings, the market loses its capacity to identify fraudulent traders and reputation loses its value. For buyers and sellers on eBay, Resnick and Zeckhauser (2002) report a rating frequency of 52 and 61 percent, respectively; Dellarocas and colleagues (2004) find 68 and 78 percent; and Jian, MacKie-Mason, and Resnick (2010) report rates of 33 and 55 percent. Moreover, averaging across seven different countries' eBay sites, Bolton and colleagues (2013) report a rating frequency of 71 percent. They attribute this high rating frequency to eBay's two-sided rating system. On the one hand, a two-sided system provides incentives to leave feedback due to positive reciprocity expectations. On the other hand, the threat of retaliation against a negative rating leads to a positive evaluation bias. Based on a dataset of ratings obtained from eBay, Dellarocas and Wood (2008) also identify reciprocity as an important factor driving behavior at the feedback stage. Moreover, they suggest that not leaving feedback indicates dissatisfaction and, based on information obtained from the observed rating patterns, they estimate the proportion of good, bad, and mediocre transactions. They show that the 99 percent positive rating usually observed on eBay is a biased estimate of the proportion of good transactions. Accounting for the "sound of silence" (i.e., transactions where one or both traders did not leave feedback) gives an estimate of 79 and 86 percent for satisfied buyers and sellers, respectively. In a similar vein, Jian and colleagues (2010) developed a model to estimate the proportion of different rating strategies from eBay feedback data. They assume three rating strategies (*do not give feedback*, *give feedback unconditionally*, and *reciprocate feedback*) and estimate buyers' and sellers' propensity to choose one of

the strategies. They find that the median buyer and seller choose the reciprocal feedback strategy in 23 and 20 percent of cases, respectively, whereas they choose the unconditional feedback strategy in 38 and 47 percent of cases.

Our study adds to these results by analyzing time-specific feedback decisions. Whether trading partners provided feedback first is an explanatory variable in our analysis. We estimate discrete-time event-history models, an approach that does not rely on restrictive assumptions about the timing distribution of traders' feedback. We also account for unobserved heterogeneity by including buyer and seller fixed effects in our models. Note, however, that we observe behavior in our data; therefore, any inference we make about traders' underlying motives is based on the consistency of the behavioral data with our theoretical propositions.

## DATA

Our data, collected from the German eBay platform, contain information on auctions of mobile phones and DVDs that ended between December 1, 2004 and January 7, 2005 (for details on data collection, see the online supplement). We restrict our analysis to offers from sellers in Germany and exclude private auctions, where the buyer is unknown, as well as auctions in which multiple pieces are offered that can be purchased by multiple buyers. Some auctions ended early because the seller withdrew the auction before the first bid, the seller stopped the auction with the current bid as the winning bid, or the buyer bought the item at the "Buy It Now" (i.e., fixed) price. We exclude these cases from our analysis. Finally, minor inconsistencies in the regular expressions caused missing or ambiguous data on key variables for a small fraction of auctions, which we also exclude. In total, we excluded 14.5 percent of mobile phone auctions and 10.4 percent of DVD auctions for these reasons.

To estimate reputation effects on the highest bids in auctions, it is sensible to select a homogeneous good to avoid bias due to unobserved heterogeneity (Diekmann et al. 2009).

**Table 1.** Means and Standard Deviations of Key Variables

	New Mobile Phones	Used Mobile Phones	DVDs
Seller			
Seller's positive ratings (log)	3.77 (1.59)	3.74 (1.61)	7.00 (2.65)
Seller's negative ratings (log)	.42 (.70)	.47 (.71)	2.00 (2.01)
Buyer (if sold)			
Buyer's positive ratings (log)	3.06 (1.70)	3.16 (1.83)	3.92 (1.69)
Buyer's negative ratings (log)	.27 (.53)	.32 (.62)	.28 (.54)
Feedback (if sold)			
Time until seller feedback (censored)	32.6 (35.3)	36.5 (36.7)	24.9 (30.1)
Positive seller feedback	.736	.685	.843
Neutral seller feedback	.002	.005	.001
Negative seller feedback	.012	.016	.003
Seller has previous rating from buyer	.032	.031	.049
Time until buyer feedback (censored)	33.9 (35.7)	39.0 (36.8)	25.2 (30.2)
Positive buyer feedback	.706	.656	.835
Neutral buyer feedback	.007	.011	.004
Negative buyer feedback	.019	.020	.004
Buyer has previous rating from seller	.030	.037	.049
Auction			
Successful auction	.958	.956	.533
Selling price in Euros (if sold)	216.7 (37.4)	150.3 (47.1)	7.4 (10.2)
Number of observations	5,499	9,128	339,517
Number of sellers	4,341	7,687	33,166
Number of buyers	4,881	7,454	100,046

*Note:* Table S1 in the online supplement lists descriptive statistics of further control variables used in our model estimations.

To ensure homogeneity in our mobile phone data, we manually sorted out offers that contained complementary goods or multiple items (23.1 percent) and divided the sample into new (unused and in original packaging) and used products. The total sample then included 14,627 mobile phones of seven different models, 37.6 percent of which were new and 95.7 percent of which received at least one bid (i.e., their auctions were successful).<sup>8</sup> In our data, the mean selling price was 217 Euros for new mobile phones and 150 Euros for used mobile phones (see Table 1 for descriptive statistics of our key variables; additional variables are documented in Table S1 in the online supplement). To evaluate whether reputation effects can also be

observed for low-cost products, we analyze a sample of 339,517 DVD auctions, 53.3 percent of which were successful. We collected data on all DVD auctions in the mentioned time span, covering hundreds of different movies. Thus, product homogeneity cannot be assumed. Also, due to the lack of a reliable automatic categorization method, we do not distinguish between new and used DVDs. Moreover, unlike the mobile phone market, with 1.25 auctions per seller, on average, the DVD market, with an average of 10.24 auctions per seller, is dominated by a relatively small number of professional sellers.

We operationalize buyers' and sellers' reputations with their number of positive and negative ratings. Because distributions of

these variables are highly right-skewed, we take the logarithm of their values (after adding 1). Note that in terms of ratings, too, sellers in the DVD market are considerably larger than sellers in the mobile phone market.

At the feedback stage, we have the time (in days) until feedback for both the seller and the buyer of a successful auction. Recall that these variables are censored at 90 days. The mean duration seems to be lower in the DVD market, and positive feedback is very common, whereas neutral and negative ratings are rare. Negative ratings are more frequent in auctions for mobile phones than for DVDs, indicating that problematic transactions are more likely when they involve more complex and expensive products (although this could also be due to the lower degree of professionalism in the mobile phone market). Another important variable at the feedback stage is whether the seller or buyer has already received a rating from the same person in a previous interaction. Such repeated encounters are more likely in the DVD market (5 percent) than in the mobile phone market (3 percent).<sup>9</sup>

## RESULTS

### *Effect of Reputation on Auction Success*

We first turn to reputation effects on sales and selling price. Table 2 gives an overview of results for our key variables. We control for many additional variables in our model estimations, but due to space limitations, we report and discuss their effects in the online supplement only.

Models 1 through 3 in Table 2 show results from logistic regressions of auction success on sellers' positive and negative ratings. In line with Hypothesis 1, we find a significantly positive effect of positive ratings and a negative effect of negative ratings on the probability of sale for new mobile phones (Model 1) and for DVDs (Model 3). Contrary to our expectation, we find no significant effects for used mobile phones (Model 2). Recall, however, that over 95 percent of all mobile phones

were sold, which leaves relatively little information for the estimation of reputation effects on the probability of sale.

Because we are dealing with auctions and not fixed-price offers, a large part of the reputation effect will be exerted on price. Table 2 also reports results from linear regressions of selling price on sellers' positive and negative ratings (Models 4 through 6). As expected, we observe a strong relation between a seller's reputation and the highest bid in our data. For both mobile phones and DVDs, the number of positive ratings has a significantly positive effect on price, and negative ratings reduce price.<sup>10</sup> Because both the number of ratings and the selling price are log-transformed, the estimated coefficients can be interpreted as a percent increase in the selling price due to a one percent increase in the number of ratings (the elasticity). Thus, doubling the number of positive ratings increases the selling price of new mobile phones (Model 4) by about  $.005 \times \ln(2) \times 100 = .35$  percent. For used mobile phones (Model 5) and DVDs (Model 6), the effect of a 100 percent increase in positive ratings is about .55 and 3.7 percent, respectively. Evaluating these relative effects at the mean prices for new and used mobile phones and DVDs (see Table 1) yields absolute effects of 61, 97, and 27 cents, respectively. Although these amounts may appear small at first sight, they are not, in fact, if one takes into account the range of sellers' reputations in our sample, where differences of 1,000 percent (e.g., 40 versus 440 positive ratings) and more are the rule rather than the exception. As Hypothesis 2 predicts, the absolute effect of negative ratings on price and sales is larger than the effect of positive ratings in all models (Models 1 through 6), although the difference is statistically significant at the 5 percent level only in Model 4 ( $p = .003$  for a two-sided test). In terms of absolute effects at mean prices, doubling the number of negative ratings decreases the price by 1.3 Euros, 2.2 Euros, and 52 cents in Models 4, 5, and 6, respectively.

In line with our expectation that used products will exhibit more uncertainty than new

Table 2. Effect of Reputation on Sales and Prices

	Product Sold (0/1)			Selling Price			Selling Price (with Fixed Effects)			
	1	2	3	4	5	6	7	8	9	10
	New Mobile Phones	Used Mobile Phones	DVDs	New Mobile Phones	Used Mobile Phones	DVDs	New Mobile Phones	Used Mobile Phones	DVDs	DVDs
Seller's positive ratings (log)	.344* (.136)	-.019 (.064)	.117** (.040)	.005*** (.001)	.008* (.003)	.053*** (.015)	.027** (.010)	-.002 (.045)	.016*** (.002)	.005 (.007)
Seller's negative ratings (log)	-.670* (.301)	-.207 (.132)	-.145* (.058)	-.011*** (.002)	-.018** (.006)	-.101** (.035)	-.055 (.034)	-.025 (.033)	-.012** (.005)	-.036** (.014)
McFadden <i>R</i> -squared	.858	.678	.133							
<i>R</i> -Squared				.844	.513	.111				
Number of observations	5,499	9,128	339,517	5,269	8,727	180,881	1,612	1,944	113,276	103,030
Number of sellers	4,341	7,687	33,166	4,242	7,474	30,018	585	691		6,901
Number of titles									18,054	15,964

Note: Models 1, 2, and 3: Coefficient estimates of logistic regressions are shown; the binary dependent variable is equal to one for successful auctions. Models 4, 5, and 6: Coefficient estimates of OLS regressions are shown; the dependent variable is the logarithm of the selling price (in Euros). Models 7, 8, 9, and 10: Coefficient estimates of fixed-effects regressions are shown (seller fixed effects in Models 7 and 8, title fixed effects in Model 9, title and seller fixed effects in Model 10); the dependent variable is the logarithm of the selling price (in Euros). Numbers in parentheses are robust standard errors (adjusted for seller-clusters). Models contain various control variables (starting price, length of product description, number of competing offers, and dummies for private profile, verified identity, Me-Page, PowerSeller, auction picture, thumbnail listing, bold listing, payment modes, auction duration, time and date of auction ending, and product subcategory); for detailed results see the online supplement.

\*  $p < .05$ ; \*\*  $p < .01$ ; \*\*\*  $p < .001$  (two-tailed tests).

products, and buyers will thus pay a higher reputation premium to sellers of used products (Hypothesis 3), the effects of positive and negative ratings on the selling price are stronger for used mobile phones (Model 5) than for new mobile phones (Model 6).<sup>11</sup> Recall, however, that reputation has a significant effect on sales only for new mobile phones. If we take into account both outcomes (the selling price and the probability of sale), the total effect of reputation is even larger for new than for used mobile phones, although the difference is statistically insignificant (not shown).<sup>12</sup>

As mentioned earlier, we cannot assume product homogeneity for our DVD dataset. We therefore replicate our analysis using a model with fixed effects for movie titles (Model 9).<sup>13</sup> Although smaller in absolute size than in the ordinary DVD model, the effects of positive and negative ratings are still highly significant.

One could object that our effects are spurious due to unobserved seller heterogeneity. For example, some sellers might be more successful for reasons unrelated to reputation and, therefore, more likely to have a good reputation, resulting in a spurious correlation between reputation and selling prices. We check this argument by analyzing the effects of changes in reputation for sellers who appear repeatedly in our dataset. We estimate models with seller fixed effects for new and used mobile phones (Models 7 and 8) and a model with crossed fixed effects for sellers and titles for DVDs (Model 10).<sup>14</sup> In the model for new mobile phones, effects of positive and negative ratings persist, as does the effect of negative ratings in the model for DVDs, but the effects disappear for used mobile phones. However, the fact that we still find reputation effects in most of our models provides strong support for the reputation hypothesis (Hypothesis 1). Note that accounting for seller fixed effects results in fairly conservative estimates of reputation effects. These estimates are based only on changes in sellers' reputations during the observation window of about one month, and all between-seller variability is discarded. Consequently,

model estimations are based on sellers with high trading frequency, who are more likely to have a good reputation, whereas newcomers, for whom an increase in reputation might be most valuable, tend to be excluded.

In summary, our analyses replicate previous findings of a positive (negative) effect of positive (negative) ratings on sales and prices (Hypothesis 1); we also find a larger absolute effect of negative ratings as compared to positive ratings (Hypothesis 2). Our additional model estimations also make it rather unlikely that the effects we find are spurious. Evidence for the uncertainty hypothesis (Hypothesis 3) is mixed, however; the reputation effect does not seem to be larger for used mobile phones than for new ones.

### *Participation in the Feedback System*

As argued earlier, if too few traders participate in the feedback forum, the reputation system will be ineffective in detecting cheaters, and reputational incentives for (first-order) cooperation at the transaction stage will not accrue. However, we do find high levels of (second-order) cooperation at the feedback stage. Buyers submitted positive feedback in over 65 percent of all transactions in our mobile phone sample, whereas only 3 percent of transactions were rated neutrally or negatively. In the DVD sample, the difference is even larger. Buyers rated 83 percent of transactions positively; neutral or negative feedback was submitted in less than 1 percent of cases (see Table 1). Furthermore, Table 3 shows that over 80 percent of transactions in the mobile phone market and 90 percent of DVD transactions were rated by at least one trader.

After an auction ends, traders have 90 days to submit their feedback; because traders can observe their partners giving them feedback, their feedback decisions are not independent from each other. Therefore, it is important to take the exact timing of traders' feedback decisions into account when analyzing their rating behavior. Figure 1 gives an overview of the distribution of feedback decisions over the 90 days. The survival functions (upper

**Table 3.** Feedback Patterns (%)

	Mobile Phones	DVDs
Auctions with mutual feedback		
Buyer rated first	32.3	48.6
Seller rated first	28.1	30.2
Simultaneous	.1	.3
Auctions with buyer feedback only	10.1	5.1
Auctions with seller feedback only	11.8	5.7
Auctions without feedback	17.7	10.1
Total	100.0	100.0
Observations	13,996	180,881

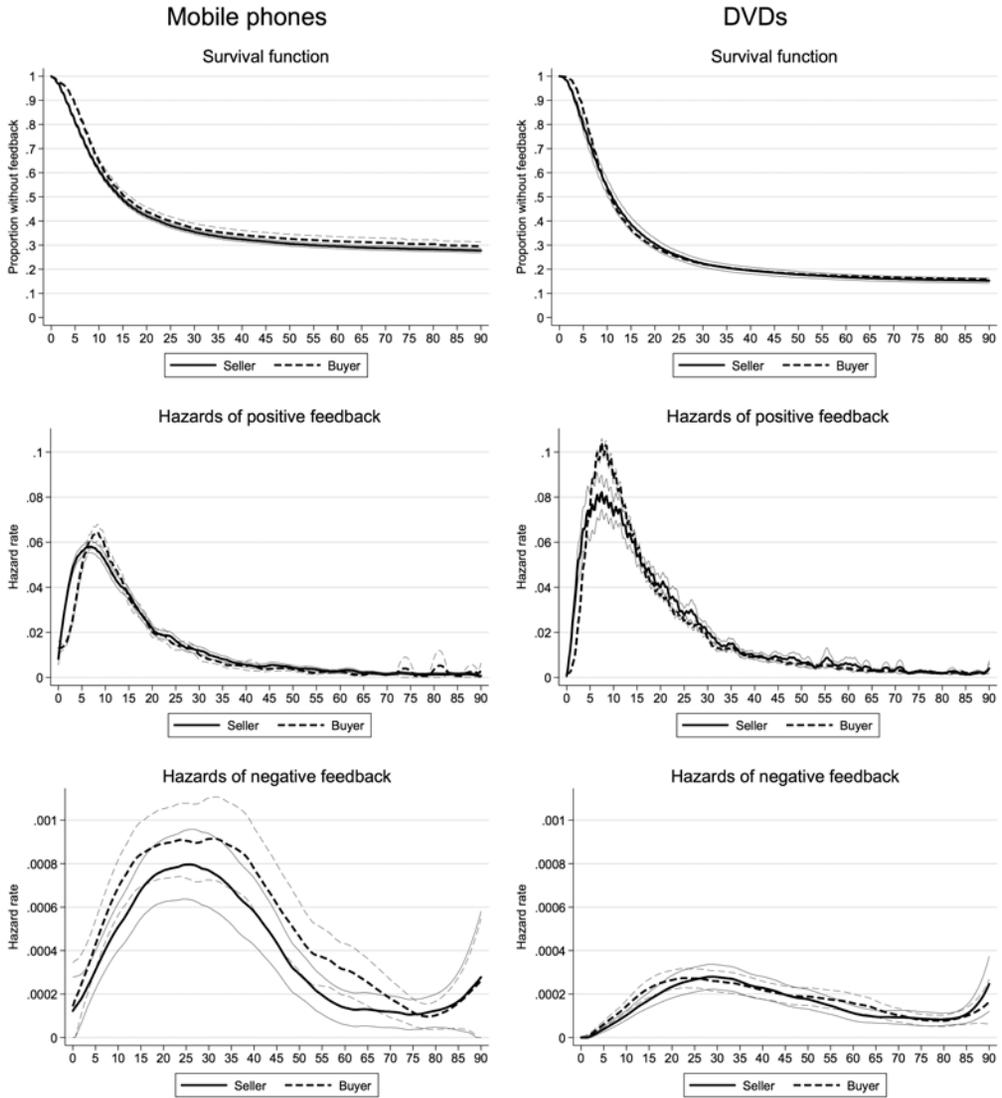
panel) show how the proportions of buyers and sellers who have not yet given feedback decrease over time; the hazard rates for positive and negative ratings (the middle and lower panel, respectively) show, roughly speaking, the feedback probability on a given day conditional on no feedback being given before; dotted lines indicate 95 percent confidence intervals (see the online supplement for computational details).

Survival curves for sellers and buyers are synchronized in both markets, but ratings are submitted faster in the DVD market. After 10 days, almost 50 percent of sellers and buyers had submitted a rating in the DVD market, but only about 40 percent had done so in the mobile phone market. Most notably, sellers are somewhat more likely to quickly submit a positive rating. That is, the hazard rate of a positive rating is higher for sellers than for buyers over the first couple of days in both markets. This indicates that some sellers might act strategically when trying to elicit a positive rating from the buyer by taking the first step. For negative ratings, the picture is different (note that the scale of the hazard rate is magnitudes smaller than for positive ratings). The curves are very similar for sellers and buyers, although in the mobile phone market the hazard rate for buyers is higher than for sellers. Furthermore, negative ratings are submitted more slowly than positive ratings. For positive ratings, hazard rates peak between days 5 and 10; for negative ratings, they are highest between days 15 and 35. Interestingly, unlike for positive ratings, we observe an increase in the hazard rates for negative rat-

ings between days 80 and 90 (although confidence intervals tend to be large as data mass gets small in this region). This indicates that some actors may delay their negative rating to escape retaliation.

The survival curves and hazard rates in Figure 1 only provide descriptive information on the distribution of ratings over time. To analyze the effects of covariates, we divide the process time into days and approximate the hazard rates using discrete-time linear probability models (Long 1997; Mood 2010) containing indicator variables for different segments of the process time (i.e., we model the baseline hazard rates shown in Figure 1 as step functions).<sup>15</sup> We also include two-way buyer and seller fixed effects in the models to control for unobserved buyer and seller characteristics that might otherwise bias the results. For instance, a positive relation between traders' reputations and their rating probability could be due to the fact that frequent raters accumulate more ratings because their ratings are reciprocated. Likewise, transaction quality or speed might be higher among traders with better reputations, thus increasing the likelihood of receiving positive feedback or decreasing the time-to-feedback reaction. Such potential biases are kept at a minimum by including seller and buyer fixed effects in the models.

Table 4 provides estimates from the discrete-time linear probability models for positive and negative feedback from buyers and sellers in the DVD market (we do not present results for mobile phones here because the



**Figure 1.** Survival Functions and Hazard Rates of Sellers' and Buyers' Rating Decisions

sample is too small for the two-way fixed-effects estimation). The displayed coefficients are scaled by a factor of 100; they can be interpreted as percentage-point effects on the conditional probability of submitting a rating on a specific day given that no rating had yet been submitted.

Our results are consistent with the hypothesis that ratings are driven by strong reciprocity (Hypothesis 4). All models in Table 4

include time-dependent variables indicating whether the trading partner submitted a positive, neutral, or negative rating first. Once a seller submits a positive rating, the buyer's conditional probability of submitting a positive rating sharply increases by about five percentage points (Model 11); for sellers, this effect is even stronger (Model 12). Likewise, the probability that a negative rating will be reciprocated with a negative rating is very

**Table 4.** Hazards of Positive and Negative Feedback in the DVD Market

	Positive Feedback		Negative Feedback	
	11	12	13	14
	Buyer	Seller	Buyer	Seller
Positive first move by partner (time-dependent)	5.106*** (.112)	13.754*** (1.013)	.003 (.004)	-.006** (.002)
Neutral first move by partner (time-dependent)	.410 (.732)	2.176*** (.461)	-.011** (.004)	.440*** (.109)
Negative first move by partner (time-dependent)	-.544 (.327)	1.859*** (.461)	.535** (.166)	1.871*** (.311)
Positive ratings (log)	-.880*** (.183)	-.421 (.234)	-.025* (.012)	-.012** (.004)
Negative ratings (log)	-.616 (.508)	-.225 (.258)	-.102* (.051)	-.008 (.006)
Partner's positive ratings (log)	-.840*** (.137)	-.259 (.269)	.017*** (.005)	.021*** (.006)
Partner's negative ratings (log)	1.023*** (.259)	-.894 (.755)	-.043** (.014)	-.059 (.035)
Previous interaction rating				
Received only	-.067 (.459)	.943 (.596)		
Provided only	-2.051*** (.428)	-1.448* (.676)		
Received and provided	-1.147*** (.219)	-.158 (.286)		
Received or provided			.002 (.005)	.006* (.003)
Number of observations	96,055	96,055	96,055	96,055
Number of events	80,601	80,343	310	223
Number of sellers	9,309	9,309	9,309	9,309
Number of buyers	26,188	26,188	26,188	26,188

*Note:* The table shows coefficient estimates for effects on the conditional probability of submitting a rating on a specific day given that no rating had been submitted yet (scaled by a factor of 100) for discrete-time linear probability models (LPMs) with seller and buyer fixed effects (standard errors in parentheses, adjusted for buyer-clusters in Models 11 and 13 and for seller-clusters in Models 12 and 14). Time indicators parameterizing the baseline hazard are not displayed.

\* $p < .05$ ; \*\* $p < .01$ ; \*\*\* $p < .001$  (two-tailed tests).

high (Models 13 and 14). Note that these effects are much smaller because of the low baseline probability.<sup>16</sup> A somewhat inconclusive finding is the positive effect of a buyer's negative rating on the seller's probability of giving a positive rating.

Moreover, our results support Hypothesis 5a, not 5b. The more positive ratings a seller has, the less likely the buyer is to submit feedback. That is, in addition to strong reciprocity, buyers' rating behavior is consistent with

Becker's altruism rather than with warm glow altruism, signaling, or indirect reciprocity. The remaining evidence in Table 4 is consistent with the view that buyers, in particular, take their interaction partners' utility into account when leaving feedback. First, in line with Hypothesis 7b, buyers (and sellers) seem to acknowledge that traders to whom they provided a rating before do not benefit from another rating. Second, the more negative ratings a seller has, the more likely a buyer is to

leave a positive rating. This might reflect buyers' willingness to support sellers who want to redeem their good reputation. Finally, we find that buyers (and sellers) are reluctant to give a negative rating if a trading partner already has many negative ratings, but they are more likely to give a negative rating to a trading partner with many positive ratings. This might reflect the fact that, after a failed transaction, traders feel greater disappointment toward trading partners with good reputations than toward those with bad reputations.

We do not find support for our hypotheses regarding sellers' strategic behavior at the feedback stage. Neither sellers' own number of positive ratings (Hypothesis 6a) nor buyers' number of positive ratings (Hypothesis 6b) exhibit a positive effect on sellers' probability of leaving positive feedback. Also, sellers are not less likely to give feedback to a buyer from whom they have received a rating in a previous interaction, which contradicts Hypothesis 7a.

In summary, we observe high levels of participation in the feedback system, which seem to be driven mainly by strong reciprocity and altruism. Both sellers and buyers reciprocate received ratings from their transaction partners (Hypothesis 4) and appear to take their interaction partners' utility into account in their feedback decisions (Hypotheses 5a and 7b). We do not find evidence for sellers' strategic considerations, as our corresponding hypotheses cannot be confirmed (Hypotheses 6a, 6b, and 7a). However, this does not mean that strategic motives are unimportant or do not exist. For example, according to our descriptive results, sellers are on average quicker than buyers to submit positive feedback, which can be seen as a strategic attempt to elicit reciprocation.

## DISCUSSION AND CONCLUSIONS

In commerce, trust problems have been ubiquitous ever since humans started trading commodities. Cultural evolution, together with the creativity of social engineers, led to a

great variety of institutions intended to help mitigate these trust problems. These institutions include the reputation system of the Maghribi traders in the Middle Ages described by Greif (1989), the "law merchant" system analyzed by Milgrom and colleagues (1990), the formation of formal social structures in nineteenth-century North America explored by Zucker (1986), and informal network governance structures as described by Jones and colleagues (1997). Factors such as the degree of traders' social embeddedness (Buskens and Raub 2002; DiMaggio and Louch 1998; Granovetter 1992), the type of commodity (Kollock 1994), the complexity of the interaction (Jones et al. 1997), and, most importantly, the lack of timely and inexpensive information about a trading partner's reputation (Milgrom et al. 1990) have been important determinants in the evolution of these institutions.

Recent technological progress has drastically diminished the cost of fast information transmission, and auction markets were among the first to establish online reputation systems for economic transactions. Based on a seller's reputation, buyers can trade off prices against the possibility of being cheated, and the price premium for a good reputation provides a financial incentive for sellers' cooperative behavior. Moreover, traders do not have to be embedded in offline social networks to access information about other traders as long as they have access to the market platform online. In the past decade, online reputation systems have spread all over the web, and the Internet's technical possibilities have brought their functioning close to perfection. However, the existence and stability of online reputation systems cannot be taken for granted. A reputation system's effectiveness in promoting (first-order) cooperation crucially depends on traders' voluntary provision of feedback. Because a trader gains no direct benefits from leaving feedback, participation in the feedback system is subject to a (second-order) free-rider problem.

Our study investigates the role of reputation in promoting cooperation and the process

of reputation formation in the field. That is, we explore the mechanisms underlying cooperation by observing transactions and ratings among traders in anonymous online markets in an unobtrusive way. We consider online markets as integrated systems of institutional rules that govern traders' decisions, and our analysis covers the entire chain of interactions between a buyer and a seller, starting with the seller's offer and ending with either or both leaving feedback. Our theoretical considerations are rooted in classical sociology and anthropology, with their insights into the roles of trust, reciprocity, and reputation in social exchange, but we also draw on recent findings from experimental social sciences. Based on a large sample of data from the mobile phone and DVD markets, we test various and partly competing hypotheses derived from our theoretical argument.

In both markets, we find that sellers' number of positive ratings has a positive effect on sales and selling prices, and sellers' negative ratings have the opposite effect (Hypothesis 1). Moreover, in line with Hypothesis 2, negative ratings have a greater effect on sales and prices than do positive ratings. This corroborates previous findings that reputation has a market value and that buyers take into account information about sellers' reputations when they place their bids (see the online supplement for a review of previous findings). However, the empirical evidence for the uncertainty hypothesis (Hypothesis 3) is mixed. As hypothesized, reputation has a larger effect on price for used mobile phones than for new ones, but this is not the case for the probability of sale. The gross effect of reputation, taking into account both the price and the probability of sale, seems even larger for new mobile phones than for used ones. A possible explanation for this finding is that used products attract buyers who are willing to take a risk and therefore do not take information about sellers' reputations into account as carefully as buyers of new products do.

At the feedback stage, cooperation is very high. Sixty percent of transactions in the mobile phone market and 80 percent in the

DVD market were rated by both sellers and buyers, and only a small fraction of transactions had no rating at all. This is comparable to what Dellarocas and colleagues (2004) and Bolton and colleagues (2013) found in their eBay data. Our study corroborates, moreover, that buyers' and sellers' behavior at the feedback stage is largely consistent with strong reciprocity (Hypothesis 4). Although one receives no apparent benefits from reciprocating a rating received from one's trading partner, the likelihood of providing similar feedback sharply increases for both buyers and sellers after receiving a rating. The effects are strong and highly significant for both positive and negative ratings. This is in line with statistics reported in Bolton and colleagues (2013) and with what Dellarocas and Wood (2008) and Jian and colleagues (2010) found based on their model estimations.

Besides reciprocity, altruistic preferences seem to motivate traders' (in particular, buyers') decisions. Buyers are more likely to give a positive rating to traders with few positive or many negative ratings (Hypothesis 5a). This contradicts Jian and colleagues (2010), who found that both buyers and sellers are more likely to give feedback to experienced traders than to inexperienced traders. We argue that this could be an artifact resulting from Jian and colleagues (2010) disregarding transaction quality in their model estimations. Transaction quality is likely to be higher for experienced traders and could thus lead to a positive correlation between traders' experience and the probability of giving feedback. We circumvent this problem by including partner fixed effects in our model estimations. If we do not include them, our results are similar to those of Jian and colleagues (2010). Moreover, we find that buyers (and sellers) are less likely to give positive feedback to the same trading partner repeatedly (Hypothesis 7b). Hypotheses 5a and 7b are in line with Becker's (1976) notion of altruism; our results are consistent with the idea that traders take their interaction partners' utility into account when leaving feedback and are more likely to give feedback the more beneficial it is for their partner.

Contrary to our expectations and unlike the results by Dellarocas and colleagues (2004) would suggest, we find no support for our hypothesis regarding sellers' strategic considerations at the feedback stage. In particular, we find no support for Hypotheses 6a and 6b, that sellers strategically take into account their own or the buyers' score at the feedback stage. Dellarocas and colleagues' (2004) conclusion was based on the finding that a buyer's first move at the feedback stage had a negative effect on the seller's overall probability of leaving feedback. We argue that this result could be an artifact due to their use of a probit model; if sellers have a lower a priori probability of providing a rating, buyers will more likely be the first movers at the feedback stage. The event-history models used in our analysis do not suffer from this problem because the buyer's first move is treated as a time-varying variable. Yet, traders seem to anticipate their trading partners' reciprocal behavior, especially in the case of negative feedback, as some try to avoid retaliation by submitting their negative rating close to the end of the rating period. Thus, we cannot conclude that strategic motives are unimportant in traders' decisions to leave feedback; they are difficult to identify in our data.

In Gouldner's (1960:176) words, the norm of reciprocity is a "starting mechanism" [that] helps to initiate social interaction." Reciprocity is also a cornerstone in Blau's (1964) social exchange theory. Both authors elaborate on the principle of reciprocity and outline its implications for the emergence and stability of exchange relations. Their notion of reciprocity and the reciprocity we observe in our data is similar to the notion of "strong reciprocity" pioneered by Gintis (2000). According to strong reciprocity, actors will return a favor with a favor and retaliate against an unfriendly act without expecting to be compensated for the costs they incur in doing so. In laboratory experiments, Fehr and colleagues (2002) demonstrated that strong reciprocity is a key mechanism in promoting cooperation in voluntary contribution games,

and they sparked a cross-disciplinary debate concerning the determinants of human social cooperation (Guala 2012). In our study, we show how the interplay of simple institutional rules and human motivation establishes a high level of cooperation among anonymous traders in online markets.

Based on our findings we argue that, in these markets, altruism, strong reciprocity, and, most likely, also strategic motives play an important role in maintaining (second-order) cooperation at the feedback stage, which generates the information necessary to create financial incentives for (first-order) cooperation at the transaction stage. In other words, giving positive or negative feedback—a second-order issue—can be conceived as rewarding or punishing a trading partner's first-order cooperation or defection, respectively (Resnick and Zeckhauser 2002). Given that the constrained set-up of online markets mirrors laboratory conditions relatively well, our findings bolster the external validity of experimental evidence of social cooperation and peer punishment. At the same time, our study makes apparent that unobtrusive behavioral data from online platforms need to be complemented with laboratory and field experiments, which are better suited to disentangling the various motives behind human behavior. Although strong reciprocity and altruism appear as the most likely motives behind the behavioral patterns we observe at the feedback stage, carefully designed experiments should be used to verify our results and rule out alternative explanations.

Electronic reputation systems have been constantly refined to better meet the requirements of a steadily growing online community. In fact, the change in eBay.de's reputation system in spring 2008 was guided by a thorough theoretical and empirical analysis that combined field data from online markets with evidence from laboratory experiments (Bolton et al. 2013). The main concern with the old two-sided feedback system was that it might foster a positive evaluation bias. On the one hand, it encouraged positive evaluations due to positive reciprocity expectations; on

the other hand, it deterred truthful negative evaluations due to negative reciprocity expectations. Without abandoning the old system, buyers were newly given the opportunity to leave a more detailed and anonymous seller rating. That is, after a transaction, buyers could rate sellers on several attributes—such as item description, communication, transaction speed, and costs—and sellers were not able to identify the buyer who rated them because they were only shown average feedback scores on each attribute. Surprisingly, although the costs for leaving feedback increased for buyers, both buyers' and sellers' conventional rating behavior hardly changed 10 weeks into the system change, and 70 percent of buyers who gave conventional feedback also gave detailed seller feedback (Bolton et al. 2013). This provides additional support for our assertion that cooperation at the feedback stage is largely motivated by strong reciprocity and altruism, rather than mere strategic considerations encouraged by the two-sided feedback system.

## Funding

This research was partly supported by the Swiss National Science Foundation (grant number 100017\_124877).

## Acknowledgments

We thank Stephanie Eckman, Debra Hevenstone, Claudia Jenny, and six anonymous reviewers for their perceptive and helpful comments.

## Data Note

An online supplement to this article with additional information on previous research, data collection, and results is available from *ASR*'s website (<http://asr.sagepub.com/supplemental>). The data and analysis files used for this article can be downloaded at <http://www.socio.ethz.ch/research/datafiles>.

## Notes

1. Most economic and social exchange is sequential. Actor A transfers resources to actor B and expects B to reciprocate at some point in the future so that both A and B earn the gains from trade. A trust problem arises when A cannot expect B to reciprocate with certainty (Coleman 1990). We call A's expectation

regarding B's propensity to reciprocate "trust." A trusting A is more likely to transfer resources to B. We call B's propensity to reciprocate "trustworthiness." B can be trustworthy because she is committed to act in A's interest out of self-interest (Hardin 2002), or due to other-regarding preferences and internalized norms of fairness and reciprocity (for other notions of trust, see Cook, Hardin, and Levi 2007).

2. eBay Germany has 14.5 million active users. We counted 142 eBay clubs with a total of 2,686 members representing .019 percent of active users (eBay.de, March 28, 2013). A somewhat higher number of participants are in discussion forums, but relative to the millions of users and transactions, their proportion is negligible. Moreover, repeated transactions with the same trader are rare. The density of the trader network will thus be close to zero, even for submarkets. Finally, online relations may be "too thin" (Hardin 2004) to establish norms of good conduct to the same extent as offline relations would (Granovetter 1992).
3. On eBay, traders can leave feedback only after a transaction, and for every transaction the market platform charges a small fee. Building a good reputation, even with sham transactions, thus requires time and money. This makes a good reputation costly to fake and the reputation system fairly robust against deliberate attempts to influence it. On platforms where the possibility to leave feedback is not restricted to traders of a particular transaction, reputations can be faked more easily. eBay's reputation system is further stabilized because the platform bans traders with too many negative evaluations and may prevent them from opening a new account under a different alias.
4. Note that Becker's altruism model is agnostic regarding the emotional triggers of altruistic behavior. Regret and guilt, for instance, can be subsumed under the model.
5. Fehr and colleagues (2002:3) point out that the "essential feature of strong reciprocity is a willingness to sacrifice resources for rewarding fair and punishing unfair behavior *even if this is costly and provides neither present nor future material rewards for the reciprocator*" (emphasis in original). Behavior as defined by strong reciprocity cannot therefore be explained by reciprocal altruism (Trivers 1971) or by reputation and indirect reciprocity (Nowak and Sigmund 2005).
6. A high number of ratings can be interpreted as an indication of being an active rater due to the high correlation between incoming and outgoing ratings.
7. The two motives described in Hypotheses 5a and 5b are not mutually exclusive and may be present simultaneously in a heterogeneous population. Empirical estimates, however, will provide information about which motive is predominant. Other motives, such as the "warm glow" or signaling,

- might also be present, but we see no way to disentangle them with our data. These motives would be consistent with a null effect of sellers' scores on buyers' behavior. Furthermore, effects described in Hypothesis 5 could also be caused by fear of disagreeable e-mails from the seller if one does not leave feedback. However, we cannot offer any clear ideas about how such fears would relate to a seller's score.
8. We collected data for five more models, but we excluded these cases from our analysis because the models were released prior to 2004. The excluded cases are composed almost entirely of used products and thus are more heterogeneous than the offers of newer models. Including these older models in our analysis does not change our main findings.
  9. We calculated these numbers based on traders' entire rating histories. They are lower bounds for the actual proportions of repeated encounters because transactions appear in the feedback history only if a rating occurred.
  10. Censoring at the starting price might bias effects of positive and negative ratings on price because our estimation is based on successful auctions only. To evaluate the robustness of our results, we applied various models that take censoring into account (e.g., censored normal regression and Heckman's sample selection model). Results are very similar in all models (not shown).
  11. A bootstrap test of the joint null-hypothesis that the absolute coefficients of positive and negative ratings are not larger in the model for used mobile phones than in the model for new mobile phones yields a  $p$ -value of .04. This  $p$ -value is an *empirical strength probability* computed as the fraction of bootstrap samples in which the point estimates for the coefficients are in accordance with the null-hypothesis (see Davison, Hinkley, and Young 2003). The  $p$ -values for the separate hypotheses are .13 for positive and .10 for negative ratings.
  12. We determined the total effect by computing average marginal effects on the logarithm of the expected price, where the expected price was computed as the predicted selling price weighted by the estimated probability of sale.
  13. Title matches could not be found for all auctions, even after removing special characters. The sample size in this analysis is thus reduced to 113,276 sold items. These items were sold by 15,545 different sellers and comprised 18,054 different titles with 6.3 auctions per title, on average (the largest group being *Harry Potter and the Prisoner of Azkaban*, with 383 auctions).
  14. For the two-way fixed-effects model, we use Schmieder's (2009) implementation of an algorithm proposed by Guimarães and Portugal (2010).
  15. We chose the segments such that a sufficient number of rating events are available in each interval. For positive feedback we use one-day intervals up

to day 30, followed by two-day intervals up to day 50, followed by five-day intervals. For negative feedback, the intervals in days are: 1–10, 11–15, 16–20, 21–25, 26–30, 31–40, 41–50, 51–70, and 71–90. The exact choice of intervals does not seem to matter much for our findings.

16. In the subsample used for the two-way fixed-effects estimation, the average conditional probability of feedback being given on a specific day is about 3 percent for positive feedback and .01 percent for negative feedback.

## References

- Akerlof, George A. 1970. "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism." *Quarterly Journal of Economics* 84:488–500.
- Allison, Paul D. 2009. *Fixed Effects Regression Models*. Thousand Oaks, CA: Sage.
- Andreoni, James. 1990. "Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving." *Economic Journal* 100:464–77.
- Ba, Sulin and Paul A. Pavlou. 2002. "Evidence of the Effect of Trust Building Technology in Electronic Markets: Price Premiums and Buyer Behavior." *MIS Quarterly* 26:243–68.
- Bajari, Patrick and Ali Hortacsu. 2004. "Economic Insights from Internet Auctions." *Journal of Economic Literature* 42:457–86.
- Becker, Gary S. 1976. "Altruism, Egoism and Genetic Fitness: Economics and Sociobiology." *Journal of Economic Literature* 14:817–26.
- Blau, Peter M. 1964. *Exchange and Power in Social Life*. New York: Wiley.
- Bolton, Gary E., Ben Greiner, and Axel Ockenfels. 2013. "Engineering Trust: Reciprocity in the Production of Reputation Information." *Management Science* 59:265–85.
- Bolton, Gary E., Elena Katok, and Axel Ockenfels. 2004. "How Effective Are Electronic Reputation Mechanisms? An Experimental Investigation." *Management Science* 50:1587–1602.
- Brown, Martin, Armin Falk, and Ernst Fehr. 2004. "Relational Contracts and the Nature of Market Interactions." *Econometrica* 72:747–80.
- Buskens, Vincent and Werner Raub. 2002. "Embedded Trust: Control and Learning." Pp. 167–202 in *Group Cohesion, Trust and Solidarity*, edited by E. J. Lawler and S. R. Thye. Amsterdam: JAI Press.
- Camerer, Colin F. 2003. *Behavioral Game Theory*. Princeton, NJ: Princeton University Press.
- Coleman, James S. 1990. *Foundations of Social Theory*. Cambridge, MA: The Belknap Press of Harvard University Press.
- Cook, Karen S., Russell Hardin, and Margaret Levi. 2007. *Cooperation Without Trust?* New York: Russell Sage Foundation.
- Cook, Karen S., Eric R. W. Rice, and Alexandra Gerbasi. 2004. "The Emergence of Trust Networks Under

- Uncertainty: The Case of Transitional Economies—Insights from Social Psychological Research.” Pp. 193–212 in *Creating Social Trust in Post-Socialist Transition*, edited by J. Kornai, B. Rothstein, and S. Rose-Ackerman. New York: Palgrave MacMillan.
- Cosmides, Leda and John Tooby. 1992. “Cognitive Adaptations for Social Exchange.” Pp. 163–228 in *The Adapted Mind*, edited by J. H. Barkow, L. Cosmides, and J. Tooby. New York: Oxford University Press.
- Dasgupta, Partha. 1988. “Trust as a Commodity.” Pp. 49–72 in *Trust: Making and Breaking Cooperative Relations*, edited by D. Gambetta. Oxford: Basil Blackwell.
- Davison, A. C., D. V. Hinkley, and G. A. Young. 2003. “Recent Developments in Bootstrap Methodology.” *Statistical Science* 18:141–57.
- Dellarocas, Chrysanthos, Ming Fan, and Charles A. Wood. 2004. “Self-Interest, Reciprocity, and Participation in Online Reputation Systems.” Working Paper 205, MIT Center for Digital Business.
- 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* 54:460–76.
- Diekmann, Andreas, Ben Jann, and David Wyder. 2009. “Trust and Reputation in Internet Auctions.” Pp. 139–65 in *eTrust: Forming Relationships in the Online World*, edited by K. S. Cook, C. Snijders, V. Buskens, and C. Cheshire. New York: Russell Sage.
- DiMaggio, Paul and Hugh Louch. 1998. “Socially Embedded Consumer Transactions: For What Kinds of Purchases Do People Most Often Use Networks?” *American Sociological Review* 63:619–37.
- Downs, Anthony. 1957. *An Economic Theory of Democracy*. New York: Harper & Row.
- Fehr, Ernst, Urs Fischbacher, and Simon Gächter. 2002. “Strong Reciprocity, Human Cooperation, and the Enforcement of Social Norms.” *Human Nature* 13:1–25.
- Gambetta, Diego. 2009. “Signaling.” Pp. 168–94 in *The Oxford Handbook of Analytical Sociology*, edited by P. Hedström and P. Bearman. Oxford: Oxford University Press.
- Gintis, Herbert. 2000. “Strong Reciprocity and Human Sociality.” *Journal of Theoretical Biology* 206:169–79.
- Gouldner, Alvin W. 1960. “The Norm of Reciprocity: A Preliminary Statement.” *American Journal of Sociology* 25:161–78.
- Granovetter, Mark. 1992. “Problems of Explanation in Economic Sociology.” Pp. 25–56 in *Networks and Organizations: Structure, Form, and Action*, edited by N. Nohira and R. G. Eccles. Boston: Harvard Business School Press.
- Greif, Avner. 1989. “Reputation and Coalitions in Medieval Trade: Evidence on the Maghribi Traders.” *Journal of Economic History* 49:857–82.
- Guala, Francesco. 2012. “Reciprocity: Weak or Strong? What Punishment Experiments Do (and Do Not) Demonstrate.” *Behavioral and Brain Sciences* 35:1–15.
- Guimarães, Paulo and Pedro Portugal. 2010. “A Simple Feasible Alternative Procedure to Estimate Models with High-Dimensional Fixed Effects.” *Stata Journal* 10:628–49.
- Gulati, Ranjay and Martin Gargiulo. 1999. “Where Do Interorganizational Networks Come From?” *American Journal of Sociology* 104:1439–93.
- Güth, Werner and Axel Ockenfels. 2003. “The Coevolution of Trust and Institutions in Anonymous and Non-anonymous Communities.” Pp. 157–74 in *Jahrbuch für Neue Politische Ökonomie*, edited by M. J. Holler, H. Kliemt, D. Schmidtchen, and M. Streit. Tübingen: Mohr Siebeck.
- Hardin, Russell. 2002. *Trust and Trustworthiness*. New York: Russell Sage Foundation.
- Hardin, Russell. 2004. “Internet Capital.” *Analyse & Kritik* 26:122–38.
- Heath, Anthony. 1976. *Rational Choice and Social Exchange: A Critique of Exchange Theory*. Cambridge: Cambridge University Press.
- Heckathorn, Douglas D. 1989. “Collective Action and the Second-Order Free-Rider Problem.” *Rationality and Society* 1:78–100.
- Jappelli, Tullio and Marco Pagano. 2002. “Information Sharing, Lending and Defaults: Cross-Country Evidence.” *Journal of Banking & Finance* 26:2017–45.
- Jian, Lian, Jeffrey K. MacKie-Mason, and Paul Resnick. 2010. “I Scratched Yours: The Prevalence of Reciprocation in Feedback Provision on eBay.” *The B.E. Journal of Economic Analysis & Policy* 10:Article 92.
- Jones, Candace, William S. Hesterly, and Stephen P. Borgatti. 1997. “A General Theory of Network Governance: Exchange Conditions and Social Mechanisms.” *Academy of Management Review* 22:911–45.
- Kollock, Peter. 1994. “The Emergence of Exchange Structures: An Experimental Study of Uncertainty, Commitment, and Trust.” *American Journal of Sociology* 100:313–45.
- Kollock, Peter. 1999. “The Production of Trust in Online Markets.” Pp. 99–123 in *Advances in Group Processes*, Vol. 16, edited by E. J. Lawler and M. W. Macy. Amsterdam: JAI Press.
- Livingston, Jeffrey A. 2003. “What Attracts a Bidder to a Particular Internet Auction?” Pp. 165–87 in *Organizing the New Industrial Economy*, edited by M. R. Baye. Amsterdam: Elsevier.
- Long, J. Scott. 1997. *Regression Models for Categorical and Limited Dependent Variables*. Thousand Oaks, CA: Sage.
- Lucking-Reiley, David, Doug Bryan, Naghi Prasad, and Daniel Reeves. 2007. “Pennies from eBay: The Determinants of Price in Online Auctions.” *Journal of Industrial Economics* 55:223–33.
- Malinowski, Bronislaw. 1922. *Argonauts of the Western Pacific: An Account of Native Enterprise and Adventure in the Archipelagoes of Melanesian New Guinea*. London: Routledge & Kegan Paul.

- Malinowski, Bronislaw. 1926. *Crime and Custom in Savage Society*. London: Routledge & Kegan Paul.
- Mauss, Marcel. [1950] 1990. *The Gift: The Form and Reason for Exchange in Archaic Society*. London: Routledge.
- Milgrom, Paul R., Douglass C. North, and Barry R. Weingast. 1990. "The Role of Institutions in the Revival of Trade: The Law Merchant, Private Judges, and the Champagne Fairs." *Economics and Politics* 2:1–23.
- Mood, Carina. 2010. "Logistic Regression: Why We Cannot Do What We Think We Can Do, and What We Can Do About It." *European Sociological Review* 26:67–82.
- Nowak, Martin A. and Karl Sigmund. 2005. "Evolution of Indirect Reciprocity." *Nature* 437:1291–98.
- Przepiorka, Wojtek. 2013. "Buyers Pay For and Sellers Invest in a Good Reputation: More Evidence from eBay." *Journal of Socio-Economics* 42:31–42.
- de Quervain, Dominique J.-F., Urs Fischbacher, Valerie Treyer, Melanie Schellhammer, Ulrich Schnyder, Alfred Buck, and Ernst Fehr. 2004. "The Neural Basis of Altruistic Punishment." *Science* 305:1254–58.
- Resnick, Paul and Richard Zeckhauser. 2002. "Trust among Strangers in Internet Transactions: Empirical Analysis of eBay's Reputation System." Pp. 127–57 in *The Economics of the Internet and E-Commerce*, edited by M. R. Baye. Amsterdam: Elsevier.
- Resnick, Paul, Richard Zeckhauser, Eric Friedman, and Ko Kuwabara. 2000. "Reputation Systems." *Communications of the ACM* 43:45–48.
- Resnick, Paul, Richard Zeckhauser, John Swanson, and Kate Lockwood. 2006. "The Value of Reputation on eBay: A Controlled Experiment." *Experimental Economics* 9:79–101.
- Schmieder, Johannes F. 2009. "gprep: Stata Module to Estimate Regressions with Two Dimensional Fixed Effects." Boston College, Department of Economics, Statistical Software Components, S457048.
- Shapiro, Carl. 1983. "Premiums for High Quality Products as Return to Reputation." *Quarterly Journal of Economics* 98:659–80.
- Snijders, Chris and Jeroen Weesie. 2009. "Online Programming Markets." Pp. 166–85 in *eTrust: Forming Relationships in the Online World*, edited by K. S. Cook, C. Snijders, V. Buskens, and C. Cheshire. New York: Russell Sage Foundation.
- Trivers, Robert L. 1971. "The Evolution of Reciprocal Altruism." *Quarterly Review of Biology* 46:35–57.
- Uzzi, Brian. 1997. "Social Structure and Competition in Interfirm Networks: The Paradox of Embeddedness." *Administrative Science Quarterly* 42:35–67.
- Yamagishi, Toshio, Karen S. Cook, and Motoki Watabe. 1998. "Uncertainty, Trust and Commitment Formation in the United States and Japan." *American Journal of Sociology* 104:165–94.
- Yamagishi, Toshio, Masafumi Matsuda, Noriaki Yoshikai, Hiroyuki Takahashi, and Yukihiro Usui. 2009. "Solving the Lemons Problem with Reputation." Pp. 73–108 in *eTrust: Forming Relationships in the Online World*, edited by K. S. Cook, C. Snijders, V. Buskens, and C. Cheshire. New York: Russell Sage Foundation.
- Zucker, Lynne G. 1986. "Production of Trust: Institutional Sources of Economic Structure, 1840–1920." *Research in Organizational Behavior* 8:53–111.

**Andreas Diekmann** is Professor of Sociology at the Swiss Federal Institute of Technology, ETH-Zurich. His research interests focus on theories of social cooperation, experimental game theory, research methods and statistics, and environmental and population sociology. His publications have appeared in *Proceedings of the Royal Society B*, the *Journal of Conflict Resolution*, *Sociological Methods & Research*, and the *American Sociological Review*. He is currently conducting research on social norms and institutions, and on environmental protests in China.

**Ben Jann** is Associate Professor of Sociology at the University of Bern, Switzerland. His research interests include social-science methodology, statistics, social stratification, and labor market sociology. Recent publications include papers in *Sociological Methodology*, *Sociological Methods & Research*, and *Public Opinion Quarterly*.

**Wojtek Przepiorka** is a research fellow at Nuffield College and the Department of Sociology at University of Oxford. His research interests are in quantitative methodology, in particular experimental methods, analytical sociology, behavioral game theory, and the evolution of social cooperation. His work has been published in *European Sociological Review*, *Evolution and Human Behavior*, and *Proceedings of the Royal Society B*. In his current research, he studies the signaling benefits of altruistic acts and mechanisms of social norm enforcement.

**Stefan Wehrli** is a PhD candidate in sociology at the Swiss Federal Institute of Technology, ETH-Zurich, and manager of the ETH Decision Science Laboratory (DeSciL). His research interests include decision science and social network analysis.