Socio-Economic Review, 2022, Vol. 20, No. 4, 1905–1932
doi: 10.1093/ser/mwab012
Advance Access Publication Date: 26 April 2021
Article

Downloaded from https://academic.oup.com/ser/article/20/4/1905/6252982 by Universiteitsbibliotheek Utrecht user on 14 December 2022

Article

The role of reputation systems in digital discrimination

Judith Kas [1,2,]* Rense Corten[1] and Arnout van de Rijt[1,3]

[1]Department of Sociology/ICS, Utrecht University, Padualaan 14, Utrecht 3584 CH, The Netherlands, [2]Research Unit Migration, Integration and Transnationalization, Berlin Social Science Center (WZB), Reichpietschufer 50, Berlin 10785, Germany and [3]Department of Political and Social Sciences, European University Institute, Via dei Roccettini 9, San Domenico de Fiesole 50014, Italy

**Correspondence:* judith.kas@wzb.eu

Abstract

Reputation systems are commonplace in online markets, such as on peer-to-peer sharing platforms. These systems have been argued to be a solution to (ethnic) discrimination on such platforms. This argument is based on empirical studies showing that ethnic disadvantages are smaller for users with ratings than for users without ratings. We argue that this conclusion may be premature, because minorities have a harder time accumulating ratings. The greater benefit of ratings to minorities may be offset by their troubles acquiring any, thereby diminishing the potential for reputation systems to reduce discrimination. We tested this counterargument using a unique data set that contains information on all interactions on a peer-to-peer motorcycle rental platform. We find that the reputation system does not reduce initial inequalities between otherwise comparable renters of different ethnicity. Platforms that wish to reduce discrimination should not only make their reputation system more effective, but also help users collect ratings.

Key words: discrimination, economic sociology, inequality, reputation, uncertainty

JEL classification: D13 household production and intra-household allocation, asymmetric and private information, D82 mechanisms design

1. Introduction

Discrimination of minorities is thought to be a widespread phenomenon in online marketplaces for goods (Edelman and Luca, 2014; Nunley *et al.*, 2011), for example, on second-hand selling platforms such as eBay, and platforms where individuals can rent and borrow goods (e.g. cars and tools). Internet platforms supporting such marketplaces enable direct

© The Author(s) 2021. Published by Oxford University Press and the Society for the Advancement of Socio-Economics. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs licence (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial reproduction and distribution of the work, in any medium, provided the original work is not altered or transformed in any way, and that the work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

communication between peers. They make personal profiles available that include names and photos of users to foster interpersonal trust (Bente *et al.*, 2012; Ert *et al.*, 2016), but this abundance of personal information also makes discrimination possible (Pope and Sydnor, 2011; Edelman and Luca, 2014; Ert *et al.*, 2016; Ge *et al.*, 2016; Cui *et al.*, 2019; Edelman *et al.*, 2017; Laouénan *et al.*, 2017; Mohammed, 2017; Wu *et al.*, 2017; Tjaden *et al.*, 2018; Wu and Jin, 2018; Jaeger *et al.*, 2019). Any discrimination in the platform economy results in fewer opportunities for users with certain demographic characteristics to buy, rent, sell and borrow goods. This outcome is disadvantageous both for the people who are discriminated against and for the platforms and possibly also for those individuals who discriminate, as it leads to suboptimal market outcomes where otherwise fruitful interactions are not realized.

Many have argued that reputation systems temper or even solve this problem of digital discrimination (Ert *et al.*, 2016; Abrahao *et al.*, 2017; Cui *et al.*, 2019; Mohammed, 2017; Tjaden *et al.*, 2018).[1] The argument goes that discrimination on platforms mostly originates from a lack of information about other users. This information is especially important for interactions in the (peer-to-peer) platform economy, as these platforms are characterized by lower levels of institutionalization and higher levels of interpersonal trust (Katz, 2015; ter Huurne *et al.*, 2017). By providing user-specific information on performance and trustworthiness through ratings from past transactions, reputation systems would supersede group stereotypes as a basis of partner choice and reduce or even eliminate unfounded inequalities. Whereas the use of photos and names always entails the risk of discrimination based on demographic characteristics, ratings are often considered a better and fairer way of reducing information asymmetry between users and providers.

Several studies have provided empirical support for the claim that reputation systems help overcome discrimination by showing that a difference between (ethnic) groups in the acceptance chances of a transaction request is reduced once users have received at least one positive rating. For example, in an online experiment with Airbnb users, reputation was found to partly offset the tendency of people to trust others who are similar to them more (Abrahao *et al.*, 2017). In another study it was even found that while guests on Airbnb with African-American-sounding names were 19.2 percentage points less likely to be accepted than those with white-sounding names, this difference completely disappeared when both guests had a positive rating (Cui *et al.*, 2019).

However, here we argue that earlier research may have overstated the extent to which reputation systems mitigate discrimination. It is critical to consider that not all users are equally likely to achieve that initial reputation necessary for acquiring trust on the platform. Most platforms only allow users to provide a rating *after* a transaction via the platform. By the very definition of discrimination, and as previous research shows, the probability that a first request from a user without any ratings is accepted, and thus the probability of receiving a first rating, strongly depends on a trustee's (ethnic) background (Abrahao *et al.*, 2017; Cui *et al.*, 2019). While a reputation system may decrease the importance of information retrieved from names and pictures once users have acquired ratings, it may altogether fail to reduce inequality in transaction volume as it gives majority members the ability to more quickly build a good reputation.

1 Reputation systems collect, aggregate and distribute feedback about trustees' past decisions to trustors (Resnick *et al.*, 2000).

Downloaded from https://academic.oup.com/ser/article/20/4/1905/6252982 by Universiteitsbibliotheek Utrecht user on 14 December 2022

In the current research, we first illustrate these inequality-sustaining effects of reputation systems with a simple model. To study the empirical relation between reputation systems and discrimination, we then analyse a unique data-set containing the complete historical records of user activity on a Dutch peer-to-peer motorcycle renting platform. This platform is similar in function and design to many other platforms, such as hospitality platform Airbnb and various carsharing platforms. However, whereas previous studies have analysed platforms in a static way, for example, comparing differences in the number of clicks an offer received (Tjaden *et al.*, 2018) or prices Airbnb hosts could charge (Edelman and Luca, 2014), the full historic record of all interactions on the online platform we study allows us to investigate the dynamics of inequality. Instead of comparing individuals with and without rating, we study the process of reputation-formation itself. In order to test whether reputation systems with time decrease or sustain discrimination on this platform, we analyse data on interactions at the platform at different time points as users accumulate reputation.

2. Theory

Modern-day exchange is increasingly mediated by online platforms. 'Platform economy' is an umbrella term encompassing several activities, such as selling, exchanging, borrowing and renting of goods and services. Leong and Belzer (2017) identified two distinctive features of platforms. First, platform economy businesses make money not by providing goods and services *per se*, but rather by connecting people who need particular goods and services with people who want to provide them. Second, to facilitate this connection efficiently, platform economy businesses rely on online platforms.

Trust and reputation formation are particularly important on sharing platforms, as interactions pose a high risk for the owners of the goods without an immediate reward. Owners may thus be strongly inclined to derive trustworthiness of others from past experiences of other owners as well as borrower demographics. When an owner decides to rent out their goods to another user, who is generally a stranger, he or she runs the risk of not getting back the good in a good state. Although most sharing platforms typically offer some support in solving problems between users, the legal structures owners can rely on are usually limited (ter Huurne *et al.*, 2017). Owners will therefore be motivated to carefully consider which renters can and which cannot be trusted before deciding to accept or reject a rental request. Our theoretical focus is the decision of the owner to accept or reject a request of a potential renter.

2.1 Discrimination in the platform economy

Unlike in traditional exchange, in online markets, people often exchange goods with perfect strangers from around the world (Frenken and Schor, 2017). There is limited opportunity to get to know a person before a transaction or to acquire information about the other person's unobserved qualities. The online nature of the platform economy implies several information asymmetries (ter Huurne *et al.*, 2017). First, consumers and providers are unsure about each other's intentions, leading to perceived personal safety risks, especially when the two actors meet offline after the online interaction. Second, consumers cannot check upfront whether the quality of the good they are buying, renting or borrowing is good. Akin to Akerlof's (1970) classical lemons problem, the risk of buying a low-quality good will result in market failure.

Downloaded from https://academic.oup.com/ser/article/20/4/1905/6252982 by Universiteitsbibliotheek Utrecht user on 14 December 2022

To mitigate this risk, platforms allow their users to create extensive user profiles, including names and profile pictures. Compared to traditional e-commerce companies, anonymity is lower in these platforms (Abrahao *et al.*, 2017). Names, photos and descriptions are used as a means of identity verification and are intended to foster an increased sense of personal contact (Dubois *et al.*, 2012; Guttentag, 2015). They also convey information about a person's ethnicity, gender and age and may thereby lead to discrimination on the basis of these demographic characteristics, with discrimination being defined as the unequal treatment of individuals or groups on the basis of their (demographic) characteristics (Pager and Shepherd, 2008).

Discrimination in online markets is believed to be partly caused by a lack of information about the potential transaction partner's characteristics that impact the value or merit of that person, such as trustworthiness (Ert *et al.*, 2016; Abrahao *et al.*, 2017; Cui *et al.*, 2019; Mohammed, 2017; Tjaden *et al.*, 2018). Individuals fill in those gaps on the basis of the perceived association between a person's demographic characteristics (such as ethnicity) and those relevant characteristics. This reliance on demographic characteristics due to a lack of information about relevant personal characteristics is called *statistical discrimination* (Becker, 1957). Applied to trust problems in the platform economy, statistical discrimination is the tendency to derive expectations about the trustworthiness of exchange partners from demographics on their personal profiles. When owners assume that renters from a specific demographic group are on average less trustworthy, they may place less trust in these individuals.

In the current study, we focus on ethnicity as a sociodemographic dimension of discrimination, because we know from the labor market literature and earlier studies on the platform economy that this is a widespread and persistent form of discrimination (Rich, 2014; Zschirnt and Ruedin, 2016). We use data from a Dutch platform and follow the definition of Statistics Netherlands that classifies individuals' ethnicity by their biological parent's birth country. In 2019, 76.4% of the Dutch residents had parents who were born in the Netherlands (Statistics Netherlands, 2019). The second and third largest group consists of (children of) migrants from Turkey and Morocco. Labor market research shows that these migrants are disadvantaged in the Dutch labor market (Gracia *et al.*, 2016). On the platform that we study, discrimination then entails owners placing more trust in renters belonging to the ethnic majority (i.e. native Dutch renters) and renters belonging to the ethnic minority (i.e. non-native Dutch renters).

2.2 Reputation systems as a solution to statistical discrimination

Earlier research proposes that reputation systems may be the most promising solution to statistical discrimination, as they are believed to provide more specific and therefore more accurate information about a renter's behavior in a transaction through the platform than more diffuse sociodemographic information (Resnick et al., 2000; Robbins, 2017). This would then allow minority renters to acquire a personal reputation divorced from any initial statistical expectation about trustworthiness based on minority group membership. After an interaction, users are asked to leave a rating and a review about their interaction partner. Ratings are numerical evaluations, usually in the form of a rating between 1 and 5 stars. Reviews are written texts in which transactions and individuals are evaluated. In the current article, we empirically focus on numeric ratings. These ratings are displayed on user's profile pages for potential future interaction partners.

Reputation systems allow users to assess the expected quality of an interaction with a potential partner ('learning' according to Buskens and Raub (2002). Users with better

Downloaded from https://academic.oup.com/ser/article/20/4/1905/6252982 by Universiteitsbibliotheek Utrecht user on 14 December 2022

ratings are more likely to have pro-social interests and are therefore more likely to live up to expectations and honor agreements, while users with negative ratings likely care less about the interest of their partner, and are therefore more likely to abuse trust again in the future. In this way, ratings serve as a costly signal of trustworthiness and other unobservable characteristics of individuals, while they at the same time provide an incentive for pro-social behavior (Resnick and Zeckhauser, 2002; Bolton *et al.*, 2004; Fehrler and Przepiorka, 2013; Tadelis, 2016). For these reasons, requests from renters with a better reputation (i.e. more positive ratings) should be more likely to be accepted by the owners, and requests from renters with more negative ratings less likely.

Ratings on online platforms are typically highly left-skewed, with many very high ratings and only few ratings lower than 4.5 out of 5 stars (Zervas *et al.*, 2015; Teubner *et al.*, 2017; Teubner and Glaser, 2018). We, therefore, define a positive rating as having 5 stars and a negative rating as having fewer than 5 stars (Przepiorka *et al.*, 2017). This skewness may be due to underreporting of negative experiences (Fradkin *et al.*, 2018), or may be a direct effect of the reputation system preventing negative experiences. As users' present behavior on the platform may affect their possibilities for future behavior, they may be motivated to behave well to build and maintain their reputation (Buskens and Raub, 2002). The lack of variation in average rating reduces the informativeness of the average rating of a user, but may increase the importance of the number of ratings a user has.

Previous research has confirmed that trustees with a better reputation are trusted more often (Boero *et al.*, 2009; Gong and Yang, 2010; Charness *et al.*, 2011; Duffy *et al.*, 2013; Fehrler and Przepiorka, 2016). Especially, the first rating seems to matter (Duffy *et al.*, 2013; Frey and Van De Rijt, 2016). One study found that users with one positive rating received between 8.4% and 29.5% more trust than users without any ratings (Cui *et al.*, 2019).

When a renter has no ratings, the owner relies on the limited information that is available, such as names and photos. Yet demographic information is only a proxy for trustworthiness, while ratings contain direct information about a renter's trustworthiness, so when ratings are available this should trump demographics in trust assessment. This supremacy of reputation over demographic information is confirmed in empirical studies: The difference between individuals of different ethnicity in the probability to participate on platforms is much smaller when reputational information about the user is available (Ert *et al.*, 2016; Abrahao *et al.*, 2017; Cui *et al.*, 2019; Mohammed, 2017; Tjaden *et al.*, 2018). This line of research thus suggests the following hypothesis:

Hypothesis 1: The ethnic gap decreases with the number of positive ratings the renters have received.

Because of this 'compensation effect', reputation systems have been suggested to constitute a solution to digital discrimination by researchers (Abrahao *et al.*, 2017; Cui *et al.*, 2019; Ert *et al.*, 2016; Mohammed, 2017; Tjaden *et al.*, 2018) and practitioners (Murphy, 2016) alike. If the reputation system indeed reduces inequality between ethnic groups, we would expect that the difference between minority and majority renters gets smaller as they accumulate ratings. They can do so by submitting more requests, which is expected to result in more transactions and thus in a better reputation (under the assumption that most ratings are positive). We, therefore, again based on previous research, hypothesize that:

Hypothesis 2: The ethnic gap decreases with the number of requests renters have filed.

Downloaded from https://academic.oup.com/ser/article/20/4/1905/6252982 by Universiteitsbibliotheek Utrecht user on 14 December 2022

2.3 Dynamic interplay between reputation formation and discrimination

However, while reputation systems may indeed decrease discrimination for those users who have positive ratings, the hypothesis that reputation systems decrease overall discrimination on the platform is based on the implicit assumption that all users are equally likely to obtain a good reputation. Yet, as ratings can generally only be provided after an interaction on the platform, not all users are equally likely to obtain ratings. The results from previous research suggest that users who already have a good reputation are more often selected for new interactions, which in turn improves their reputation and thus the probability that their next request will be accepted. In a laboratory experiment, Frey and Van De Rijt (2016) show that these 'reputation cascades' lead to arbitrary inequality between equally trustworthy individuals. The chance to build a reputation thus depends on the probability that an initial request is accepted. If this probability differs between people of different ethnicity, members of some ethnic groups may have fewer opportunities for reputation building than others. We thus argue that minorities' requests for a transaction are less likely accepted, so that per filed request (either accepted, declined or not responded), ethnic minority renters receive fewer ratings than ethnic majority renters. As such, the greater benefit of positive ratings to minority members enabled by reputation systems is counteracted by minorities' greater struggles to obtain such ratings in the first place.

We use a simulation model to better understand this interplay between reputation and discrimination. We simulated a sequence of 10 rounds indexed t in each of which two renters, $i = 1, 2$, who have a different ethnicity ($E_i = 0$ for a renter with majority ethnicity, $E_i = 1$ for a renter with minority ethnicity), each submit a rental request to different owners.[2] These proposals are accepted and rejected based on a combination of minority group status and the number of positive ratings received. The outcome of a request in a given round, A_{it}, equals 1 if the request is accepted and 0 otherwise. We assume that all renters are equally trustworthy and that they receive a positive rating after every accepted request. The number of ratings a renter has accumulated thus equals the sum of previously accepted requests. The probability that the request is accepted in a given period is $0 \leq P_{it} \leq 1$. P_{it} is determined by a logistic function, $P(L_{it})_{it} = \frac{e^{L_{it}}}{1+e^{L_{it}}}$, where L_{it} depends on the ethnicity and reputation of the renter:

$$L_{it} = G - E_i * D * (1 + R * \sum_{1}^{t-1} A_{it})^{-C} + R * \sum_{1}^{t-1} A_{it}. \quad (1)$$

In Equation (1), G is the level of trust in ethnic majority renters. D denotes the level of discrimination. We assume that there is discrimination for actors without reputation: requests from renter 1 (belonging to the ethnic majority) are accepted more often than requests from renter 2 (who belongs to the ethnic minority): $D > 0$.R denotes the effect of

2 On many platforms, like eBay, Airbnb and Couchsurfing, most users but especially consumers (i.e. buyers, renters), transact only a handful of times (Resnick and Zeckhauser, 2002; Lauterbach *et al.*, 2009; Teubner, 2018). Renters on the motorcycle sharing platform that we study in the next sections on average submitted 2.5 requests since the start of the platform, so with 10 periods in our simulation model we study a time frame that includes the vast majority of user histories on our platform and most other platforms. Most renters will never submit 10 requests, and if the reputation system fails to reduce the inter-ethnic inequality in acceptance chances in a few rounds, most minorities will never experience equal chances.

Downloaded from https://academic.oup.com/ser/article/20/4/1905/6252982 by Universiteitsbibliotheek Utrecht user on 14 December 2022

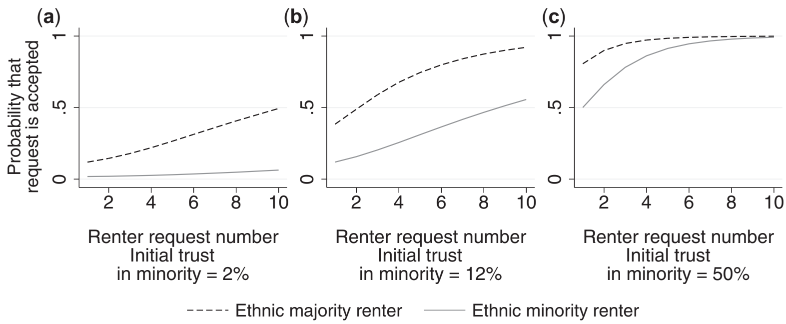

Figure 1 Simulation results: average acceptance rate of native and non-native Dutch renters over the rounds.

ratings. The more ratings a renter has, the higher the probability that the request is accepted: $R > 0$. The compensation effect identified in prior work that owners pay less attention to ethnicity when renters have more ratings is implemented as a non-negative parameter $C \geq 0$. The larger C is and the more ratings renter 2 has accumulated, the smaller the influence of discrimination. We systematically varied parameters G, D, R and C, assessing their effects on the size of the ethnic gap, defined as the absolute difference in the probability that majority and minority renters are trusted in a given round and parameter setting. We ran 1000 iterations per parameter combination.

We find that what crucially determines whether the ethnic gap closes or widens over the course of rounds is G—D (see Appendix A). That is, whether the ethnic gap decreases with the number of requests renters have filed, is perpetuated or increases predominantly depends on the initial level of trust in ethnic minorities. Figure 1 shows the average rate with which requests from ethnic majority and ethnic minority renters are accepted over the rounds in the simulations for three different levels of initial trust in the minority. The converging lines in panel C of Figure 1 show that when the initial level of trust in ethnic minority renters is high, their initial disadvantage can be compensated through reputation, consistent with what past scholars have argued.

Panels A and B, however, show that when requests from ethnic minority renters are less frequently accepted, these renters do not get the opportunity to build a reputation and can, therefore, not benefit from the reputation system. As a result, the acceptance rates of the different renters diverge. Hence, initial differences between renters in the probability to be accepted for an interaction may accumulate over time, thereby diminishing the potential of reputation systems to decrease discrimination. In other words, a reputation system will reduce inequality only when minorities are commonly trusted also in the absence of any reputation, that is, when trust is not a major problem to begin with. These results continue to hold when the ethnic gap is measured as the ratio in odds of a request being accepted between majority and minority renters. For example, in panel A of Figure 1 this odds ratio increases between rounds 1 and 10 from 6.6 to 14.3 while in panel C it decreases from 1.5 to 1.0.

Based on these simulation results, we predict that given sufficiently low initial success chances of minorities, the ethnic gap will not diminish on a platform with a reputation system, in contrast to what earlier studies concluded. This is the main contribution of the article: we argue that, contrary to what is commonly believed, reputation systems do not necessarily

Downloaded from https://academic.oup.com/ser/article/20/4/1905/6252982 by Universiteitsbibliotheek Utrecht user on 14 December 2022

lead to a decrease in inequality in online markets. The purpose of the analyses is to test if the hypothesis from previous research that reputation system reduces discrimination still holds true when taking into account the fact that the reputation systems may reproduce existing biases (Hypothesis 2). We provided a theoretical explanation for why this hypothesis may be problematic, and we reevaluate this claim of previous research in a dynamic context.

3. Data and methods

We study a Dutch peer-to-peer motorcycle sharing platform that was founded in 2016. The platform operates within the so-called sharing economy and has a similar design and functionality as peer-to-peer carsharing platforms and hospitality platforms such as Airbnb. It is common on sharing platforms for users to advertise their goods and also this platform allows motorcycle owners to advertise their motorcycle. Renters can browse through the listed motorcycles and the personal profiles of the owners and send rental requests for specific time slots and for a predefined price. Before accepting or declining the request, the owner can view the personal profile of the renter, including first name, photo, personal description and ratings. These personal profiles are commonplace on online platforms. When the request is accepted, the renter pays the rental price to the platform and the owner and renter meet offline to hand over the motorcycle. Similar to the payment system on other platforms, the platform keeps the renter's payment until the transaction has been completed. After the rental period, the platform transfers the money paid by the renter to the owner. Both renter and owner are asked to provide a rating that is publicly displayed in their user profile. Many platforms automatically arrange full insurance options for rental goods, and also the present platform automatically insures motorcycles during the rental period and checks drivers' licences and fraud histories of all users.

The platform facilitates the sharing of a fragile good of high-value motorcycles with strangers. Risk is thus particularly high. Moreover, motorcycle owners attach certain non-material values to their motorcycles that cannot be reimbursed by insurances, as illustrated by the following quotes: 'How do you let some random person ride your motorcycle? I would never do that! It's quite hard to watch your baby go the first time. No matter how attached you are to your motorcycle, it's difficult watching someone else ride it off. Even when you sell it!' (Hooshmand, 2018). 'The first mental hurdle to get over is the fact that you'd be "cheating" on your bike if you decide to ride someone else's' (Klinger, 2018). In addition to the material and emotional risk of entrusting a stranger with a motorcycle, the transaction requires that the parties meet offline before and afterward which may pose personal safety concerns, especially in case of conflict.

We analyse the complete historical records of user activity on the platform. The data-set provided to us by the platform contains information on all (11 418) interactions that took place since the start of the platform, May 2016, to July 2017. We excluded unfinished requests that were not filed, requests that were cancelled by the platform,[3] and requests from renters who had rented from the same owner in the past (as we are interested in trust

3 The platform automatically cancels requests for two reasons: (a) the renter does not have the required driver's license; (b) renters can send multiple (similar) requests at the same time. When one of the requests is confirmed, the other requests are automatically cancelled.

Downloaded from https://academic.oup.com/ser/article/20/4/1905/6252982 by Universiteitsbibliotheek Utrecht user on 14 December 2022

between strangers). The remaining data include 7181 requests for 973 motorcycles sent by 2896 renters to 851 owners.

3.1 Dependent variable

Our dependent variable is the decision of the owner. They can either actively accept (2626 requests, 36.6%) or decline (2443 requests, 34.0%) a request, or not send a response at all, after which the request expires (2112 requests, 29.4%). An earlier study of discrimination on the basis of sexual orientation shows that discrimination on Airbnb is mostly driven by non-responses rather than outright rejection (Ahuja and Lyons, 2019). Because both result in an unfulfilled request, we combined the latter two categories into the category 'declined'. A robustness check shows that the results do not change when excluding expired requests from the analyses.

3.2 Independent variables

To operationalize the reputation of the renter, we created two continuous variables, indicating the number of positive ratings and the number of negative ratings. Since 92.1% of all ratings are 5-star ratings, we define a positive rating as having 5 stars and a negative rating as having fewer than 5 stars (Przepiorka *et al.*, 2017). The variable 'number of positive ratings' counts the number of 5-star ratings a renter received. The variable 'number of negative ratings' indicates the number of ratings with less than five stars a renter received. Renters without any ratings serve as the reference category.

We make use of data from the Dutch Civil Registration (DCR) to operationalize the users' ethnicity (Edelman *et al.*, 2017; Laouénan *et al.*, 2017; Hofstra and de Schipper, 2018). All names are verified by the insurance company with which the platform collaborates, so the names visible to the owners are the renters' real names. The DCR data are register data of those who have Dutch nationality and were alive and living in the Netherlands. We have aggregated DCR data that comprises 3800 unique first names (85.0% of the unique names in the platform data set, covering 93.0% of the users of the platform[4]). Per name we know the frequencies of combinations of the parents' birth countries. Based on the definition of Statistics Netherlands, country-combinations are classified into one of two ethnic origin groups: Native Dutch or non-native Dutch (Statistics Netherlands, 2018). In cases where both parents were born in the Netherlands, the combination is assigned to the native Dutch group. If only one or neither parent was born in the Netherlands, the combination is classified as non-native Dutch. Per name we then calculate the probability that a user with that name is native Dutch. We include this continuous variable in the analyses. For some of the tables and figures, we used dichotomous ethnicity variables rather than continuous ones. In such cases, we assigned users their most likely ethnicity. The final independent variable is

4 One hundred and twenty-three of the unique names in the platform data contained special symbols (√, ∏, etc.) and could therefore not be matched with the DCR-data. The appearance of these symbols is probably caused during the transition from the platform to Excel. There does not seem to be a relation between the ethnicity of the renter and these symbols (two examples: 'G√∂khan' and 'Mari√´lle'). Among the remaining 37 names that could not be matched non-native Dutch names seem to be overrepresented ('Dzahid' and 'Carone'), but there are also names that are more likely to belong to native-Dutch renters ('Juergen' and 'Paulpeter').

Downloaded from https://academic.oup.com/ser/article/20/4/1905/6252982 by Universiteitsbibliotheek Utrecht user on 14 December 2022

the cumulative request count of the renter, which is the total number of requests made by the renter at the time of the request, including the current request.

3.3 Control variables

We included all other information visible to motorcycle owners: the price and duration of the rental, the number of years a renter had been a member on the platform and the age and gender of the renter,[5] whether the renter had a profile picture and a linked Facebook account. We included the same information about owners and we also added the following control variables related to the motorcycles: weight, engine displacement (cc), power (hp), age of the motorcycle in years and the number of positive and negative ratings of the motorcycle. We also included the total number of requests submitted by a renter, to account for differences in the tendency to submit a new request.[6] The final control variable is time. As requests are not evenly distributed over time, with peaks in summer months and virtually no activity in winter, we operationalize time by the total number of completed requests at the time of the new request (divided by 1000).

3.4 Analytical strategy

The data have a cross-classified structure: requests are nested within renters and owners, but renters and owners are not nested within each other: the same renter interacts with multiple owners and vice versa. To test the hypotheses, we ran a cross-classified multilevel linear probability model with as dependent variable a dummy variable indicating whether the request was accepted or not. We used a Bayesian estimator with two MCMC (Markov Chain Monte Carlo) chains and non-informative priors in Mplus (Muthén and Muthén, 2017). We used the default convergence criterion of Mplus (Proportional Scale Reduction factor lower than 1.1). We used 50 000 iterations of which the first half is considered a burn-in phase.

We included random intercepts for renters and owners in all analyses. In the second and third models, we added random slopes for reputation. We used full information maximum likelihood (FIML) to deal with missing values. This method includes partial information in observations with missing values, which allowed us to use all available information (Collins *et al.*, 2001).

In the first model, we included the renter's ethnicity and reputation to check if there is indeed inequality between renters of different ethnicity, and to see if reputation increases trust. We control for the total number of requests submitted by the renter in all models. To the second model, we added the interaction between the renter's ethnicity and reputation to test if renters with a better reputation are discriminated less, testing Hypothesis 1. The third model

5 We apply the same methodology to estimate the gender of the users as for estimating their ethnicity.

6 We tested if past experiences at the platform affected renter's tendency to submit another request using a cross-classified multilevel logistic regression with the dependent variable indicating whether a renter submitted at least one more request after the current one. The renter's ethnicity, the fraction of previous successful requests and the interaction between the two are included as independent variables. We found that renters who have experienced more rejections are indeed more likely to file a new request ($b = -1.307$, 95% confidence interval $= -1.400, -1.215$). Moreover, we found that the interaction between the historical success rate and the renter's ethnicity is significant and negative ($b = -0.043$, 95% confidence interval $= -0.091, -0.003$). That means that even though renters who submit another request tend to have a lower success rate than renters who do not file another request, this is less so for non-native Dutch renters than for native Dutch renters.

Downloaded from https://academic.oup.com/ser/article/20/4/1905/6252982 by Universiteitsbibliotheek Utrecht user on 14 December 2022

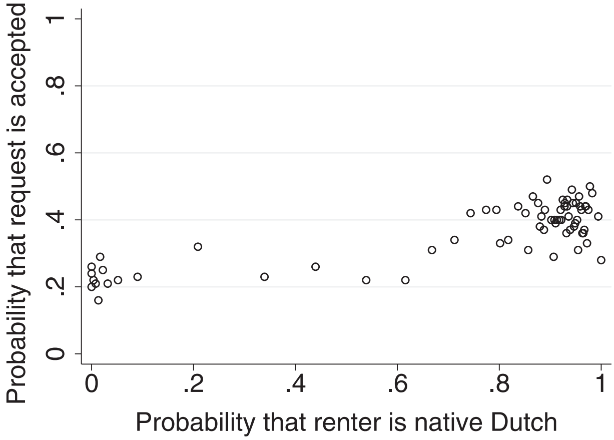

Figure 2 Average fraction of requests accepted, by renter's ethnicity. Each point represents 100 requests.

included the renter's reputation and ethnicity, which is interacted with the cumulative request count of the renter, testing Hypothesis 2.

We ran the same models with and without control variables. We ran the models in several steps. First, we only included variables at the level of the transaction (Level 1). We then removed insignificant variables before adding renter- and owner-level variables (Levels 2 and 3). We also removed insignificant variables before adding cross-level interactions, except when we were testing cross-level interactions between insignificant variables.

4. Results

4.1 Descriptive statistics

Figure 2 shows the relationship between the acceptance rate and the renter's ethnicity. Requests from renters with more Dutch-sounding names are accepted more often. Native Dutch requesters are about twice as likely to receive a positive response as non-native Dutch requesters.

Figure 3 shows the relationship between a renter's reputation and the probability that a request is accepted. Having more positive ratings seems to have a positive effect on the acceptance rate, both for native Dutch and non-native Dutch renters (panel A). Having negative ratings seems to be detrimental for the probability to receive trust, especially for non-native Dutch renters (panel B). A note here is that there are no non-native Dutch renters with more than five positive ratings, suggesting that indeed they have a harder time establishing a reputation.

Figure 4 shows how the accept rate changes with every request filed for renters of different ethnicity. Only renters who have submitted five or more requests are included to control for the changing composition of the sample, which may, for example, be caused because renters who were previously unsuccessful are more likely to submit another request.[7]

7 The figure looks similar when including only renters who submitted at least three or seven requests.

Downloaded from https://academic.oup.com/ser/article/20/4/1905/6252982 by Universiteitsbibliotheek Utrecht user on 14 December 2022

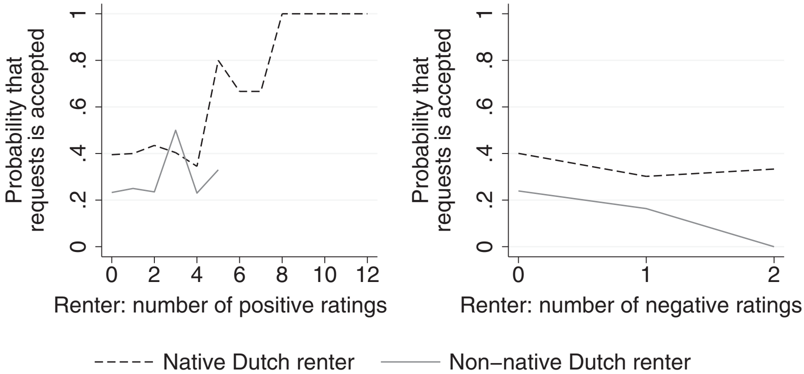

Figure 3 Average fraction of requests accepted, by renter's ethnicity and number of ratings.

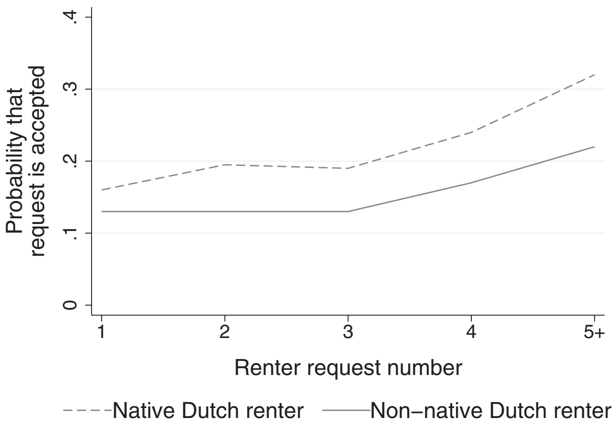

Figure 4 Fraction of requests accepted, by the cumulative number of requests made by the renter and the renter's ethnicity (i.e. '1' on the *x*-axis refers to the first request made by a renter, '2' to the second request made by the same renter). Only renters who submitted five or more requests are included. Data points representing the five or more requests are collapsed into one category.

Per renter, we randomly drew one request in order to account for individual differences in the probability that their request is accepted due to characteristics other than ethnicity. The figure is based on the average acceptance rate for 500 draws. The figure shows the acceptance chances for every request filed for native Dutch ($n = 255$) and non-native Dutch renters ($n = 87$). The ethnic gap is again visible: native Dutch renters seem to have a higher probability of getting a positive response than non-native Dutch renters. The figure shows an increase in the probability that a request was accepted with every additional request filed. Crucially, the acceptance rates of native and non-native Dutch renters do not converge. Apparently, the reputation system is not able to overcome persistently lower acceptance rates of minority requesters.

Downloaded from https://academic.oup.com/ser/article/20/4/1905/6252982 by Universiteitsbibliotheek Utrecht user on 14 December 2022

4.2 Regression results

Table 2 contains the results from the regressions without control variables and Table 3 the results with control variables. We find that requests from renters with more Dutch-sounding names have a higher probability of being accepted by the owners (Table 2, Model 1). This gap is persistent even when control variables are added (Table 3, Model 4).

We also find that having more positive ratings increases the probability of getting a positive response (Table 2, Model 1), and this effect is persistent when control variables are added (Table 3, Model 4). Having more negative ratings does not affect the acceptance probability, neither in the regressions without control variables (Table 2, Model 1), nor in the models with control variables ($b = -0.031$, 95% confidence interval: -0.089, 0.026, $P = 0.140$).

We expected that the ethnic gap would be smaller for renters with more and positive ratings (Hypothesis 1), but we do not find evidence for this in either the regression models with or without controls (Tables 2 and 3, Model 2 and 5). The findings suggest that having more positive ratings is equally beneficial for native Dutch and non-native Dutch renters. This implies that the ethnic gap does not become smaller once renters have ratings.

To test Hypothesis 2, we examine changes in the ethnic gap with every request filed by a renter. The results of this analysis are shown in Table 2, Model 3 (without controls) and Table 3, Model 6 (with controls). We find that the probability that a request is accepted increases with the number of requests a renter has already filed. However, we do not find that there is a difference between renters with Dutch-sounding names and non-Dutch sounding names in extent to which the probability that a request is accepted changes with the number of requests done by that renter. This means that the disadvantage non-native Dutch renters have *vis-a-vis* native Dutch renters does not decrease as they file more requests. This leads us to reject Hypothesis 2.

4.3 Robustness checks

When using a logistic regression instead of a linear probability model, there are only minor differences with the results in Tables 2 and 3. The coefficient of the number of negative ratings in Model 2 is insignificant when using a linear probability model, but is negative and significant when using a logistic regression. The interaction between having negative ratings and the ethnicity of the renter in Model 5 is significant and positive when using a logistic regression. This means that non-native Dutch renters suffer more from negative ratings. When using a logistic regression, the positive effect of having positive ratings disappears in Model 6.

In the main analyses, we operationalized reputation as the number of positive and negative ratings. When reputation is operationalized as having no ratings versus one or more ratings the results do not change. When taking the natural logarithm of the number of ratings, the coefficient of the interaction between the ethnicity of the renter and the number of negative ratings a renter has is positive and significant in the model with control variables, which means that non-native Dutch renters experience more negative effects of having negative ratings. The other results remain the same.

Excluding non-responses (i.e. requests that are neither accepted nor rejected) from the analyses does not affect the results.

Downloaded from https://academic.oup.com/ser/article/20/4/1905/6252982 by Universiteitsbibliotheek Utrecht user on 14 December 2022

Table 1 Descriptive statistics

Variable	Level	N_{miss}	Mean	Standard deviation	Minimum	Maximum
Dependent variables						
Accepted	I	0	0.37	0.48	0	1
Declined	I	0	0.34	0.47	0	1
No response	I	0	0.29	0.46	0	1
Independent variables (renter characteristics)						
Dutch ethnicity	R	223	0.79	0.30	0	1
Number of positive ratings	I	0	0.24	0.70	0	12
Number of negative ratings	I	0	0.03	0.19	0	2
Request number	I	0	2.84	2.75	0	27
Control variables						
Number of rental days	I	0	1.81	1.49	1	33
Motorcycle characteristics						
Day price	I	2	66.06	25.53	20.9	184.59
Number of positive ratings	I	0	3.60	6.60	0	53
Number of negative ratings	I	0	0.78	1.54	0	9
Age	I	4	12.83	7.47	0	61
Weight	I	189	217.97	40.53	93	585
Engine displacement (CC)	I	12	865.24	269.13	113	2294
Horsepower (HP)	I	0	87.74	29.99	10	173
Renter characteristics						
Age	I	2	35.56	10.81	20	80
Member number of years	I	0	0.20	0.36	0	2.12
Female	R	191	0.12	0.31	0	1
Profile picture (%)	R	0	0.61	0.49	0	1
Facebook verification	R	0	0.26	0.44	0	1
Cumulative total number of requests (per 1000)	I	0	3.59	2.07	0.00	7.18
Owner characteristics						
Number of positive ratings	I	0	2.31	3.67	0	33
Number of negative ratings	I	0	0.16	0.50	0	6
Age	I	2	35.74	10.21	18	73
Member number of years	I	0	0.64	0.53	0	2.29
Dutch ethnicity	O	142	0.85	0.22	0	1
Female	O	117	0.08	0.25	0	1
Profile picture (%)	O	0	0.81	0.39	0	1
Facebook verification	O	0	0.37	0.48	0	1

Note: Variables are measured at the level of the interaction/request (I), renter (R) or owner (O).
$R_{requests} = 7181$.
$N_{renters} = 2896$.
$N_{owners} = 851$.

4.4 Control variables

Table 3 includes the coefficients of the control variables that were included in the main analyses. We find that requests from older renters and from renters with a profile picture have a

Downloaded from https://academic.oup.com/ser/article/20/4/1905/6252982 by Universiteitsbibliotheek Utrecht user on 14 December 2022

Table 2 Results of the multilevel cross-classified linear probability model of outcome of the request: accepted (0/1) without control variables

Independent variables	Model 1	Model 2	Model 3
Intercept	0.321***	0.331***	0.385***
	(0.288, 0.352)	(0.285, 0.364)	(0.337, 0.435)
Main effects			
Renter Dutch ethnicity	0.144***	0.137***	0.161***
	(0.117, 0.175)	(0.107, 0.185)	(0.113, 0.205)
Renter number of positive ratings (ref. cat. = no reviews)	0.058***	0.070***	–
	(0.042, 0.073)	(0.041, 0.113)	
Renter number of negative ratings	0.002	−0.032	–
	(−0.053, 0.057)	(−0.149, 0.052)	
Renter request number	–	–	0.041***
			(−0.045, −0.037)
Interactions			
Renter Dutch × Renter number of positive ratings	–	−0.007	–
		(−0.062, 0.025)	
Renter Dutch × Renter number of negative ratings	–	0.049	–
		(−0.044, 0.207)	
Renter Dutch × Renter request number	–	–	−0.003
			(−0.014, 0.006)
Control variables			
Renter total number of requests	−0.024***	−0.025***	−0.041***
	(−0.026, −0.021)	(−0.028, −0.022)	(−0.045, −0.037)
Residual variance renter-level	0.001***	0.002***	0.007***
	(0.000, 0.003)	(0.000, 0.005)	(0.002, 0.015)
Residual variance owner-level	0.039***	0.039***	0.042***
	(0.033, 0.046)	(0.033, 0.046)	(0.035, 0.049)
n	7181	7181	7181

Note: 95% confidence interval in parentheses.
*Indicates significance at $P = 0.05$ (one-tailed tests).
**Indicates significance at $P = 0.01$ (one-tailed tests).
***Indicates significance at $P = 0.001$ (one-tailed tests).

higher probability of being accepted, while requests from female renters and renters who have verified their Facebook account are less likely to be accepted. Native Dutch, female and older owners, as well as owners with a profile picture accept more requests. Owners who have been a member for a longer time and owners who have verified their Facebook account are less likely to accept a request. Requests for longer rental periods and for motorcycles with a higher number of negative ratings are accepted more often.

Requests for older motorcycles have a lower probability of being accepted. The probability that a request is accepted decreases over time. The probability that a request is accepted does not vary with the number of rental days.

Table 3 does not include the coefficients for insignificant control variables at the request level, which were removed before adding higher level control variables. Table B1 in Appendix B shows the model with all request-level control variables. We found that the day

Downloaded from https://academic.oup.com/ser/article/20/4/1905/6252982 by Universiteitsbibliotheek Utrecht user on 14 December 2022

Table 3 Results of the multilevel cross-classified linear probability model of the outcome of the request: accepted (0/1) with control variables

Independent variables	Model 4	Model 5	Model 6
Constant	−0.104*	0.078	0.068
	(0.000, 0.206)	(−0.019, 0.174)	(−0.033, 0.173)
Main effects			
Renter Dutch ethnicity	0.138***	0.135***	0.144***
	(0.108, 0.169)	(0.095, 0.170)	(0.100, 0.191)
Renter number of positive ratings (ref. cat. = no ratings)	0.028**	0.034	0.028**
	(0.011, 0.045)	(−0.015, 0.072)	(0.010, 0.046)
Renter number of negative ratings	–	−0.111*	–
		(−0.201, 0.002)	
Renter request, number	0.034***	0.036***	0.036***
	(0.028, 0.040)	(0.030, 0.043)	(0.027, 0.044)
Interactions			
Renter Dutch × Renter, number of positive ratings	–	−0.003	–
		(−0.051, 0.047)	
Renter Dutch × Renter, number of negative ratings	–	0.119	–
		(−0.023, 0.207)	
Renter Dutch × Renter request, number	–	–	−0.001
			(−0.011, 0.007)
Control variables			
Rental characteristics			
Number of rental days	0.002	0.002	0.002
	(−0.005, 0.009)	(−0.005, 0.009)	(−0.005, 0.009)
Cumulative total number of interactions (per 1000)	−0.017***	−0.017***	−0.017***
	(−0.024, −0.010)	(−0.024, −0.010)	(−0.024, −0.009)
Renter characteristics			
Renter total number of requests	−0.039***	−0.040***	−0.039***
	(−0.043, −0.035)	(−0.044, −0.036)	(−0.043, −0.035)
Renter female	−0.037*	–	–
	(−0.072, −0.001)		
Renter has profile picture	0.056***	0.056***	0.057***
	(0.032, 0.082)	(0.033, 0.082)	(0.032, 0.081)
Renter has Facebook verification	−0.035**	−0.036**	−0.037**
	(−0.061, −0.010)	(−0.060, −0.012)	(−0.062, −0.011)
Renter age	0.001**	0.001**	0.001**
	(0.000, 0.002)	(0.000, 0.002)	(0.000, 0.003)
Owner characteristics			
Owner Dutch ethnicity	0.138***	–	–
	(0.108, 0.169)		
Owner age	0.005***	0.005***	0.005***
	(0.004, 0.007)	(0.004, 0.007)	(0.004, 0.007)
Owner member, number of years	−0.071***	−0.073***	−0.073***
	(−0.102, −0.040)	(−0.104, −0.042)	(−0.103, −0.042)
Owner female	0.025	–	–
	(−0.039, 0.091)		

continued

Downloaded from https://academic.oup.com/ser/article/20/4/1905/6252982 by Universiteitsbibliotheek Utrecht user on 14 December 2022

Table 3 *Continued*

Independent variables	Model 4	Model 5	Model 6
Owner has profile picture	0.140***	0.126***	0.126***
	(0.086, 0.195)	(0.073, 0.178)	(0.075, 0.177)
Owner has Facebook verification	−0.030	–	–
	(−0.068, 0.009)		
Motorcycle characteristics			
Motorcycle age	−0.003**	−0.003**	−0.003**
	(−0.006, −0.001)	(−0.006, −0.001)	(−0.006, −0.001)
Motorcycle, number of negative ratings	0.013*	0.013*	0.013*
	(0.001, 0.025)	(0.001, 0.025)	(0.001, 0.025)
Residual variance renter-level	0.002***	0.004***	0.004***
	(0.000, 0.005)	(0.001, 0.009)	(0.000, 0.012)
Residual variance owner-level	0.035***	0.035***	0.035***
	(0.029, 0.042)	(0.029, 0.042)	(0.029, 0.042)
Observations	7181	7181	7181

Note: 95% confidence interval in parentheses.
***$P < 0.001$,
**$P < 0.01$,
*$P < 0.05$ (one-tailed test).

price, weight, engine displacement (cc) and power of the motorcycle, duration of the membership of the renter, reputation of the owner and the number of positive ratings of the motorcycles do not affect the probability that a request is accepted.

4.5 Exploratory analysis

To better understand the mechanisms leading up to those the findings, we explored a number of potential explanations. First, we explored whether owners on the platform indeed place more trust in renters who have the same ethnicity as they have (homophily). Homophily is treated as a behavioral outcome measure here, since we do not have information about the mechanisms leading to this outcome. In another set of cross-classified linear probability models with the acceptance of requests as dependent variable, we included the difference between the renter's ethnicity and the owner's ethnicity (i.e. the difference in the extent to which their names sound native Dutch) as the independent variable. We also included the renter's ethnicity to see whether the distance between the ethnicity of the renter and the owner explains all of the discrimination.

The results of these regressions are given in Table 4, Model 7. When the renter and owner are of the same ethnicity, the probability that the request is accepted is higher. However, even when controlled for the difference between the owner's and renter's ethnicity, requests from renters with non-Dutch sounding names are accepted less often. This means that homophily is not the only explanation for the ethnic gap. Besides a preference for renters who have the same ethnicity as they have, non-native Dutch owners also have a preference for native Dutch renters. Table 5 shows the acceptance rates of different combinations of renters and owners. It seems that the homophily effect is mostly present for native Dutch owners and owners of Turkish ethnicity.

Downloaded from https://academic.oup.com/ser/article/20/4/1905/6252982 by Universiteitsbibliotheek Utrecht user on 14 December 2022

Table 4 Results of the exploratory analyses

Independent variables	Model 7: DV = accepted	Model 8: DV = accepted
Intercept/threshold	0.363***	0.344***
	(0.311, 0.417)	(0.305, 0.386)
Main effects		
Distance ethnicity	−0.063**	–
	(−0.115, −0.012)	
Renter Dutch	0.109***	0.143***
	(0.059, 0.158)	(0.103, 0.182)
Experience owner	–	−0.019***
		(−0.028, −0.011)
Interaction effect		
Renter Dutch × experience owner	–	0.002
		(−0.005, 0.009)
Variance intercept renter-level	0.002***	0.003***
	(0.001, 0.007)	(0.000, 0.007)
Variance intercept owner-level	0.040***	0.048***
	(0.034, 0.048)	(0.040, 0.058)
n	7181	7181

Note: 95% confidence interval in parentheses.
*Indicates significance at $P = 0.05$ (two-tailed tests).
**Indicates significance at $P = 0.01$(two-tailed tests).
***Indicates significance at $P = 0.001$ (two-tailed tests).

Table 5 Number of accepted requests and acceptance rate (in parentheses) per renter-owner ethnicity combination

		Owner, *n* (%)				
		Dutch	Moroccan	Turkish	Other	Total
Renter	Dutch	2123	61	50	134	2296
		(39.9)	(33.7)	(32.7)	(46.2)	(39.7)
	Moroccan	174	16	16	13	203
		(24.5)	(21.1)	(30.8)	(26.5)	(24.3)
	Turkish	114	11	7	9	137
		(28.6)	(33.3)	(36.84)	(33.3)	(29.3)
	Other	107	5	1	16	134
		(24.1)	(16.7)	(7.1)	(38.1)	(24.7)
	Total	2395	88	73	176	2626
		(36.8)	(29.2)	(30.5)	(41.8)	(36.6)

Second, we explored if renters without Dutch-sounding names indeed receive fewer ratings. Using a linear regression of the number of ratings per filed request on the renter's ethnicity shows that this is indeed the case. Per request, renters with completely Dutch-sounding names receive 0.18 ratings more than renters with names that do not sound Dutch

Downloaded from https://academic.oup.com/ser/article/20/4/1905/6252982 by Universiteitsbibliotheek Utrecht user on 14 December 2022

at all ($b = 0.181$, $z = 7.34$, $P < 0.001$). This difference is entirely caused by the difference in acceptance rate between renters with the Dutch ethnicity and renters with other ethnicity, as renters with less Dutch-sounding names receive slightly more ratings per accepted request ($b = -0.055$, $z = -2.05$, $P = 0.041$). There is a significant difference in the number of stars owners rate renters of different ethnicity with: compared to names that are completely native Dutch, renters with non-native Dutch names on average get ratings that are 0.13 points lower [$b = 0.134$, 95% confidence interval = (0.058, 0.207), $P < 0.001$].

Lastly, we explored whether more experienced owners behave differently than less experienced owners. Since there are very few cases in which the renter turns out to be untrustworthy (as reflected in the extremely high reputation scores and the low frequency of insurance claims), we would expect that more experienced owners are better able to select the relevant information for assessing the trustworthiness of the renter. Hence, we expect that more experienced owners trust more and discriminate less. To test this hypothesis, we ran a multilevel linear probability model with the outcome of the request (accepted or not) as the dependent variable. We included the renter's ethnicity and the number of completed transactions of the owner, as well as the interaction between the two. The results of this regression are in Table 4, Model 8. We find that more experienced owners are less likely to accept a request, and that they are not less likely to discriminate. This suggests that more experienced owners are more selective, possibly because they receive more requests in general.

5. Discussion

Reputation systems are often proposed as the most promising solution to (ethnic) discrimination in online markets (Abrahao *et al.*, 2017; Cui *et al.*, 2019; Ert *et al.*, 2016; Mohammed, 2017; Tjaden *et al.*, 2018). In the current article, we argue that reputation systems fail to solve discrimination except in cases where minorities have little problem finding a transaction partner to begin with. Using simulations, we develop a new argument about the interplay between discrimination and reputation systems. We argue that reputation systems may sustain inequality between renters with different ethnic backgrounds, even when discriminated renters can compensate their initial disadvantage with reputation. According to our theory, reputation systems may reproduce the biases of its users: only individuals who are given the opportunity to participate in interactions can accumulate ratings.

Previous research concluded that reputation systems decrease inequality. We tested if this hypothesis is still supported when taking into account that the reputation formation process may reproduce existing biases. To this end, we drew on data from a Dutch motorcycle-sharing platform. We find instead that the reputation system fails to reduce the ethnic gap. The difference between majority and minority members in the likelihood of having a request granted persists even for renters with positive ratings. Regardless of the reputation of the renter, requests from renters with an ethnic minority background are less likely to be accepted. We show that this decreases their probability of getting a (positive) rating, which in turn further decreases their chances to participate in future interactions. This implies that, contrary to what has been assumed in previous research, reputation systems may reproduce the biases of its users and therefore that the extent to which they may reduce inequality is limited.

We argue that our research setting allows for a particularly clean test of the hypothesis that reputation systems may increase discrimination, given that it involves a clear trust

Downloaded from https://academic.oup.com/ser/article/20/4/1905/6252982 by Universiteitsbibliotheek Utrecht user on 14 December 2022

problem. The platform we studied is an exemplary case of a sharing platform in which consumers grant each other temporary access to underutilized physical assets (Frenken and Schor, 2017). The platform is similar to many other sharing platforms in its procedures to match providers to consumers, the institutions in place that stimulate trust formation (e.g. secure payment system, insurances) and the risks that owners and renters encounter. Moreover, rather than investigating the effect of ethnicity and reputation on the number of clicks an offer receives or on a proxy of the number of bookings, we had access to a complete data set of all interactions that ever occurred on the platform. Our data set contains data on real and complete user profiles, which allowed us to study a more natural setting than in lab experiments where user profiles are constructed by the researchers.

A limitation of our test of the theory is that there are only very few ethnic minority renters with many positive ratings on the platform we studied. This limits the possibility to test to what extent reputation may compensate for the initial disadvantage of these renters. At the same time, this limitation in itself serves as a proof of our argument: we predicted that ethnic minority renters are less likely to obtain ratings, which is reflected in the low number of ethnic minority renters with many positive ratings.

Our contribution to the literature on discrimination and reputation is two-fold. First, we reevaluated the claim that if objective information derived from third parties (positive ratings) is available, in-group preferences no longer matter for acceptance rates or are at least significantly reduced (Abrahao *et al.*, 2017; Cui *et al.*, 2019; Laouénan *et al.*, 2017; Tjaden *et al.*, 2018; Ert and Fleischer, 2019). Although reputation is beneficial both for native ethnic majority and ethnic minority renters, we do not find evidence for this 'compensation hypothesis' (Hypothesis 1). Motorcycle owners may generally attach more non-material value to their motorcycles than users of other platforms to their shared possessions, as the quotes in the introduction suggest. This may strengthen the preference for renters with specific characteristics, as owners may simply prefer not to have certain renter with certain characteristics ride their bike (their 'baby'), regardless of how trustworthy that renter has proven to be. Another explanation for why we do not find this 'compensation effect' may be that reputation information is not the type of information that owners are looking for. While reputation information did not decrease the ethnic gap, other types of information may succeed in doing so. A final explanation for the lack of support for Hypothesis 1 ('compensation effect') is that the discrimination on this platform may not be statistical (i.e. caused by a lack of information), but based on other motives. As opposed to statistical discrimination, taste-based discrimination refers to a preference for certain characteristics over other characteristics, without an underlying expectation of qualities related to these characteristics (Becker, 1957). If discrimination is not caused by a lack of relevant information, providing more or better information is not expected to decrease discrimination. This reasoning could also explain why more experienced owners do not discriminate less.

Second, and more importantly, we argued that rating systems may fail to overcome inequalities caused by discriminatory tendencies, even in the presence of a compensation effect. While past field experiments convincingly show that ethnic background is less of a determinant of success on online platforms among users with profiles that contain (artificially created) positive ratings than among profiles lacking such ratings, we emphasize that groups that are discriminated against are less likely to obtain positive ratings in the first place, precisely because they are less likely to be accepted. Our simulation model shows that, as a result, the ethnic gap in acceptance chances may actually grow rather than recede within

Downloaded from https://academic.oup.com/ser/article/20/4/1905/6252982 by Universiteitsbibliotheek Utrecht user on 14 December 2022

empirically reasonable time spans. Even if ethnic minority users can compensate their initial disadvantage with a good reputation, initial differences between (ethnic) groups may be sustained due to a process of cumulative advantage. Evidence of similar cumulative advantage processes has been found in a wide variety of fields, such as education and poverty (Diprete and Eirich, 2006). Our data-set provides a unique opportunity to assess the dynamic process in which initial inequalities may be reproduced, because we observe complete trajectories of acceptance rates of individual users, across all their interactions on the platform. While we find that reputation benefits renters of all ethnicities, we show that with time the reputation system maintains the disadvantage of ethnic minority renters.

This is important knowledge in the light of efforts to mitigate discrimination in the rapidly growing platform economy. The emphasis on an experience of 'personal contact' through revealing personal details by peer-to-peer market platforms has the unintended consequence of opening the door to discrimination based on precisely this personal information, and our results indicate that the effectiveness of 'institutional' solutions as implemented by platforms to mitigate such discrimination may be limited. In particular, practitioners should be aware of the possibility that reputation systems may exacerbate rather than reduce inequalities between users with different (ethnic) backgrounds. While reputation systems may improve the chances for some individuals to participate on the platform, they may, as an unintended consequence, reduce the chances for others. Platforms that aim to reduce discrimination should not only make their reputation system more effective, they should also reduce initial differences between users with different backgrounds, and help first-time users to acquire their first rating, especially if a user belongs to a group that is discriminated against on the platform.

Funding

This work was supported by the Dutch Research Council, grant 452-16-002.

Acknowledgements

The authors thank Jeroen Weesie for his advice on the statistical analysis.

References

Abrahao, B, Parigi, P., Gupta, A. and Cook, K. S. (2017) 'Reputation Offsets Trust Judgments Based on Social Biases among Airbnb Users', Proceedings of the National Academy of Sciences, **114**, 9848–9853.

Ahuja, R. and Lyons, R. C. (2019) 'The Silent Treatment: LGBT Discrimination in the Sharing Economy', *Oxford Economic Papers*, **71**, 564–576.

Akerlof, G. (1970) 'The Market for Lemons: Qualitative Uncertainty and the Market Mechanism', *Quarterly Journal of Economics*, **84**, 235–251.

Becker, G. S. (1957) *The Economics of Discrimination*, Chicago, IL, University of Chicago Press.

Bente, G., Baptist, O. and Leuschner, H. (2012) 'To Buy or Not to Buy: Influence of Seller Photos and Reputation on Buyer Trust and Purchase Behavior', *International Journal of Human-Computer Studies*, **70**, 1–13.

Boero, R. et al. (2009) 'Reputational Cues in Repeated Trust Games', *The Journal of Socio-Economics*, **38**, 871–877.

Downloaded from https://academic.oup.com/ser/article/20/4/1905/6252982 by Universiteitsbibliotheek Utrecht user on 14 December 2022

Bolton, G. E., Katok, E. and Ockenfels, A. (2004) 'Trust among Internet Traders', *Analyse & Kritik*, **26**, 185–202.

Buskens, V. and Raub, W. (2002) 'Embedded Trust: Control and Learning', *Group Cohesion, Trust and Solidarity*, **19**, 167–202.

Charness, G., Du, N. and Yang, C. L. (2011) 'Trust and Trustworthiness Reputations in an Investment Game', *Games and Economic Behavior*, **72**, 361–375.

Collins, L. M., Schafer, J. L. and Kam, C. M. (2001) 'A Comparison of Inclusive and Restrictive Strategies in Modern Missing Data Procedures', *Psychological Methods*, **6**, 330–351.

Cui, R., Li, J. and Zhang, D. J. (2019) 'Reducing Discrimination with Reviews in the Sharing Economy: Evidence from Field Experiments on Airbnb', Management Science, **66**, 1071–1094.

Diprete, T. A. and Eirich, G. M. (2006) 'Cumulative Advantage as a Mechanims of Inequality: A Review of Theoretical and Empirical Developments', *Annual Review of Sociology*, **32**, 271–297.

Dubois, D., Willinger, M. and Blayac, T. (2012) 'Does Players' Identification Affect Trust and Reciprocity in the Lab?', *Journal of Economic Psychology*, **33**, 303–317.

Duffy, J., Xie, H. and Lee, Y. J. (2013) 'Social Norms, Information, and Trust among Strangers: Theory and Evidence', *Economic Theory*, **52**, 669–708.

Edelman, B. and Luca, M. (2014) *Digital Discrimination: The Case of Airbnb.com*, Harvard Business School NOM Unit Working paper No. 14–054, Boston, MA, Harvard Business School.

Edelman, B., Luca, M. and Svirsky, D. (2017) 'Racial Discrimination in the Sharing Economy: Evidence from a Field Experiment', *American Economic Journal*, **9**, 1–22.

Ert, E. and Fleischer, A. (2019) 'The Evolution of Trust in Airbnb: A Case of Home Rental', *Annals of Tourism Research*, **75**, 279–287.

Ert, E., Fleischer, A. and Magen, N. (2016) 'Trust and Reputation in the Sharing Economy: The Role of Personal Photos on Airbnb', *Tourism Management*, **55**, 62–73.

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**, 139–145.

Fehrler, S. and Przepiorka, W. (2016) 'Choosing a Partner for Social Exchange: Charitable Giving as a Signal of Trustworthiness', *Journal of Economic Behavior & Organization*, **129**, 157–171.

Fradkin, A., Grewal, E. and Holtz, D. (2018) 'The Determinants of Online Review Informativeness: Evidence from Field Experiments on Airbnb (No. 2939064)', *SSRN Electronic Journal*, accessed at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2939064 on May 15, 2017.

Frenken, K. and Schor, J. (2017) 'Putting the Sharing Economy into Perspective', *Environmental Innovation and Societal Transitions*, **23**, 3–10.

Frey, V. and Van De Rijt, A. (2016) 'Arbitrary Inequality in Reputation Systems', *Scientific Reports*, **6**, 1–5.

Ge, Y. et al. (2016) *Racial and Gender Discrimination in Transportation Network Companies*, NBER Working Paper Series No. w22776, Cambridge, MA, National Bureau of Economic Research, accessed at https://doi.org/10.3386/w22776 on November 1, 2016.

Gong, B. and Yang, C. L. (2010) 'Reputation and Cooperation: An Experiment on Prisoner's Dilemma with Second-Order Information No. 1549605', *SSRN Electronic Journal*, accessed at https://doi.org/10.2139/ssrn.1549605 on December 5, 2016.

Gracia, P., Vázquez-Quesada, L. and Van de Werfhorst, H. G. (2016) 'Ethnic Penalties? The Role of Human Capital and Social Origins in Labour Market Outcomes of Second-Generation Moroccans and Turks in The Netherlands', *Journal of Ethnic and Migration Studies*, **42**, 69–87.

Guttentag, D. (2015) 'Airbnb: Disruptive Innovation and the Rise of an Informal Tourism Accommodation Sector', *Current Issues in Tourism*, **18**, 1192–1217.

Downloaded from https://academic.oup.com/ser/article/20/4/1905/6252982 by Universiteitsbibliotheek Utrecht user on 14 December 2022

Hofstra, B. and de Schipper, N. C. (2018) 'Predicting Ethnicity with First Names in Online Social Media Networks', *Big Data & Society*, **5**, 1–14.

Hooshmand, D. (2018) 'Everything You Wanted to Know about Riders Share: How to Make Passive Income with a Motorcycle', accessed at https://motofomo.com/riders-share-review-passive-income/ on November 15, 2018.

ter Huurne, M. et al. (2017) 'Antecedents of Trust in the Sharing Economy: A Systematic Review', *Journal of Consumer Behaviour*, **16**, 485–498.

Jaeger, B. et al. (2019) 'The Effects of Facial Attractiveness and Trustworthiness in the Sharing Economy', *Journal of Economic Psychology*, **75**, 102125.

Katz, V. (2015) 'Regulating the Sharing Economy', *Berkeley Technology Law Journal*, **30**, 11–29.

Klinger, S. (2018) 'Riders Share: New Motorcycle Sharing Service Launched—ADV Pulse', accessed at https://www.advpulse.com/adv-news/motorcycle-sharing-service-riders-share/ on November 15, 2018.

Laouénan, M. et al. (2017) 'Ethnic Discrimination on an Online Marketplace of Vacation Rental Ethnic Discrimination on an Online Marketplace of Vacation Rentals', accessed at https://hal.archives-ouvertes.fr/hal-01514713/ on July 12, 2018.

Lauterbach, D. et al. (2009) 'Surfing a web of trust: reputation and reciprocity on Couchsurfing.com', *2009 International Conference on Computational Science and Engineering*, Baton Rouge, LA, USA.

Leong, N. and Belzer, A. (2017) 'The New Public Accommodations: Race Discrimination in the Platform Economy', *Georgetown Law Journal*, **105**, 1271–1322.

Mohammed, A. R. S. (2017) 'Designing for Racial Impartiality: The Impact of Relocating Host Photos on the Airbnb Website', accessed at https://www.aeaweb.org/conference/2019/preliminary/paper/6D5dQa8h at July 13, 2018.

Murphy, L. (2016) *Airbnb's Work to Fight Discrimination and Build Inclusion'*, A Report Submitted to Airbnb, Washington, DC, Laura Murphy & Associates.

Muthén, L. K. and Muthén, B. O. (2017) *Mplus User's Guide*, 8th edn, Los Angeles, CA, Muthén & Muthén.

Pager, D. and Shepherd, H. (2008) 'The Sociology of Discrimination: Racial Discrimination in Employment, Housing, Credit, and Consumer Markets', *Annual Review of Sociology*, **34**, 181–209.

Pope, D. G. and Sydnor, J. R. (2011) 'What's in a Picture? Evidence of Discrimination from Prosper.com', *Journal of Human Resources*, **46**, 53–92.

Przepiorka, W., Norbutas, L. and Corten, R. (2017) 'Order without Law: Reputation Promotes Cooperation in a Cryptomarket for Illegal Drugs', *European Sociological Review*, **33**, 752–764.

Resnick, P. et al. (2000) 'Reputation Systems', *Communications of the ACM*, **43**, 45–48.

Resnick, P. and Zeckhauser, R. (2002) 'Trust among Strangers in Internet Transactions: Empirical Analysis of eBay' s Reputation System', *Advances in Applied Microeconomics*, **11**, 127–157.

Rich, J. (2014) *What Do Field Experiments of Discrimination in Markets Tell Us? A Meta Analysis of Studies Conducted since 2000*, IZA Discussion Papers No. 8584, Bonn, IZA.

Robbins, B. G. (2017) 'Status, Identity, and Ability in the Formation of Trust', *Rationality and Society*, **29**, 408–448.

Statistics Netherlands (2018) 'Definitions: Native Dutch', accessed at https://jeugdmonitor.cbs.nl/en/definitions/Native on December 12, 2018.

Statistics Netherlands (2019) 'Hoeveel mensen met een migratieachtergrond wonen in Nederland?', accessed at https://www.cbs.nl/nl-nl/dossier/dossier-asiel-migratie-en-integratie/hoeveel-mensen-met-een-migratieachtergrond-wonen-in-nederland- on October 29, 2019.

Tadelis, S. (2016) 'Reputation and Feedback Systems in Online Platform Markets', *Annual Review of Economics*, **8**, 321–340.

Downloaded from https://academic.oup.com/ser/article/20/4/1905/6252982 by Universiteitsbibliotheek Utrecht user on 14 December 2022

Teubner, T. (2018) 'The Web of Host-Guest Connections on Airbnb—A Social Network Perspective', *Journal of Systems and Information Technology*, **20**, 262–277.

Teubner, T. and Glaser, F. (2018) 'Up or Out: The Dynamics of Star Rating Scores on Airbnb', *ECIS 2018 Proceedings*, London, Educational Collaborative for International Schools.

Teubner, T., Hawlitschek, F. and Dann, D. (2017) 'Price Determinants on Airbnb: How Repuation Pays off in the Sharing Economy', *Journal of Self-Governance and Management Economics*, **5**, 53–80.

Tjaden, J. D., Schwemmer, C. and Khadjavi, M. (2018) 'Ride with Me–Ethnic Discrimination in Social Markets', *European Sociological Review*, **34**, 418–432.

Wu, J., and Jin, M. (2018) 'Signaling Peer Trust in Accommodation-sharing Services: Effects of Similarity and Reviews on Listing Sales', *Wuhan International Conference on e-Business*, Wuhan, WHICEB.

Wu, J., Ma, P. and Xie, K. L. (2017) 'In Sharing Economy We Trust: The Effects of Host Attributes on Short-Term Rental Purchases', *International Journal of Contemporary Hospitality Management*, **29**, 2962–2976.

Zervas, G., Proserpio, D. and Byers, J. (2015) 'A First Look at Online Reputation on Airbnb, Where Every Stay is above Average (No. 2554500), *SSRN Electronic Journal*, accessed at https://doi.org/10.2139/ssrn.2554500 on January 23, 2020.

Zschirnt, E. and Ruedin, D. (2016) 'Ethnic Discrimination in Hiring Decisions: A Meta-Analysis of Correspondence Tests 1990–2015', *Journal of Ethnic and Migration Studies*, **42**, 1115–1134.

Appendix A: Simulations

Table A1 shows the parameter settings used in the simulations.

Table A1 Parameter settings used in the simulation

Parameter	Values tested
Number of rounds	$t = 10$
Number of renters	$i = 2$
Initial level of trust in majority	$G \in \{-2, -1, 0, 1, 2\}$
Discrimination	$D \in \{1, 2\}$
Importance of reputation	$R \in \{0.5, 1, 1.5, 2\}$
Compensation effect	$C\left\{\begin{array}{c}0.0, 0.1, 0.2, 0.3, 0.4, 0.5, \\ 0.6, 0.7, 0.8, 0.9, 1.0\end{array}\right\}$

Simulation results

Figure A1 shows the relation between the ethnic gap and the input parameters over the rounds, averaged over all levels of the other parameters. Especially the lines in the graph of the initial level of trust in the ethnic minority are interesting. At low levels of trust, inequality increases over the rounds, while at higher levels of trust, the ethnic gap becomes smaller over the rounds. Discrimination has a positive effect on the ethnic gap, but does not affect the change in the ethnic gap over the rounds. The lines of the importance of reputation and the compensation effect are more or less horizontal and at the same level, which means that they do not affect the (change in) the ethnic gap.

Downloaded from https://academic.oup.com/ser/article/20/4/1905/6252982 by Universiteitsbibliotheek Utrecht user on 14 December 2022

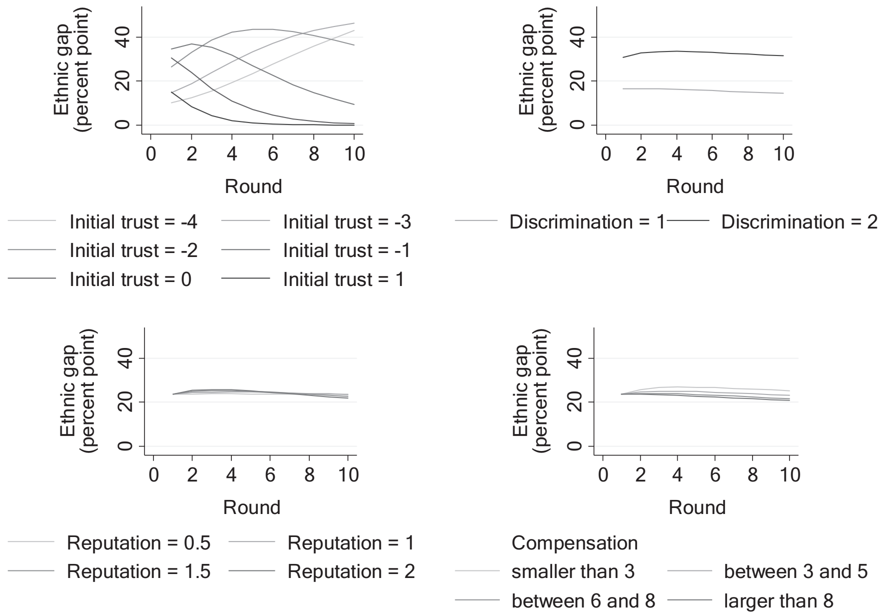

Figure A1 Simulation results: change in the ethnic gap (difference in the probability that ethnic majority and minority renters are trusted) over the rounds.

The results of a regression analyses with the ethnic gap as the dependent variable and the parameters and round number as independent variables are in Table A2. The ethnic gap is defined as the absolute difference in the probability that minority and majority renters are accepted. The average acceptance probability across all 1000 iterations is calculated using a logistic regression for every combination of the input parameters, and is used as input in the regression models in Table A2.

We added the interaction between round number and the four parameters one-by-one to investigate which parameters have the strongest influence on the change in the ethnic gap. The results of these regressions are in Table A2. As we would expect, the coefficients of the initial level of trust in minority renters and compensation effect are negative, which means that the ethnic gap is smaller when these variables have higher values.

The results do not change when operationalizing the ethnic gap as the odds ratio of acceptance between minority and majority renters. While a stronger compensation effect closes the ethnic gap somewhat, the magnitude of the effect is limited. More discrimination leads to a wider ethnic cap. The importance of reputation does not affect the absolute ethnic gap, but turns negative when operationalizing the ethnic gap as the relative difference between minority and majority renters.

The coefficient of the round number is negative, which indicates that on average the ethnic gap decreases over time. However, this change in the ethnic gap between round 1 and round 10 is negligible because the size of the change is very small. In the first round, the average probability that ethnic majority renters are trusted is 22.3 percentage points higher than the

Downloaded from https://academic.oup.com/ser/article/20/4/1905/6252982 by Universiteitsbibliotheek Utrecht user on 14 December 2022

Table A2 Results of a linear regression with the ethnic gap (difference in probability that majority and minority renters are trusted) as the dependent variable

	Model 1	Model 2	Model 3	Model 4	Model 5
Main effects					
Round number	−0.002	−0.002	−0.002	−0.002	−0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Initial level of trust in minority renters (G—D)	−0.074	0.064	−0.074	−0.074	−0.074
	(0.002)	(0.004)	(0.002)	(0.002)	(0.002)
Discrimination (D)	0.060	0.060	0.054	0.060	0.060
	(0.002)	(0.002)	(0.005)	(0.002)	(0.002)
Importance of reputation (R)	0.001	0.001	0.001	0.009	0.001
	(0.002)	(0.002)	(0.002)	(0.005)	(0.002)
Compensation (C)	−0.013	−0.013	−0.013	−0.013	−0.005
	(0.002)	(0.002)	(0.002)	(0.002)	(0.005)
Interactions					
Round × initial level of trust in minority renters	–	−0.025	–	–	–
		(0.001)			
Round × discrimination	–	–	0.001	–	–
			(0.001)		
Round × importance reputation	–	–	–	−0.001	–
				(0.001)	
Round × compensation	–	–	–	–	−0.001
					(0.001)
Constant	0.250	0.250	0.250	0.250	0.250
	(0.005)	(0.004)	(0.005)	(0.005)	(0.005)
Observations	4400	4400	4400	4400	4400
R^2	0.370	0.528	0.371	0.371	0.371

Notes: The following variables are standardized: initial level of trust, discrimination, reputation, compensation. Standard errors in parentheses. A positive number indicates a larger ethnic gap. Based on simulated data

probability that ethnic minority renters are trusted. In the last round this gap is 22.2 percentage points. When operationalizing the ethnic gap as odds ratio, the effect of round number is also minimal.

The negative coefficient of the interaction between the round number and the initial level of trust in ethnic minority renters indicates that for low levels of trust the ethnic gap increases over the rounds. For higher initial trust levels this effect is reversed: the ethnic gap decreases over time. This can be explained by a combination of factors. When initial trust increases, it becomes easier for ethnic minority renters to acquire ratings and to profit from their reputation. However, a higher initial trust rate also implies that there is less room for an increase in inequality as the maximum probability that a request accepted is fixed. The coefficients of the interactions between the round number and the other three parameters are very small, so they have a minimal effect on the change in the ethnic gap over time.

Downloaded from https://academic.oup.com/ser/article/20/4/1905/6252982 by Universiteitsbibliotheek Utrecht user on 14 December 2022

Appendix B: Additional regression results

Table B1 Results of the multilevel cross-classified linear probability model of outcome of the request on the motorcycle rental platform: accepted (0/1) with request level control variables

Independent variables	Model 1
Threshold	0.133**
	(0.034, 0.302)
Main effects	
Renter, number of positive ratings (ref. cat. = no ratings)	0.037***
	(0.020, 0.055)
Renter, number of negative ratings	−0.031
	(−0.089, 0.026)
Renter request, number	−0.009***
	(−0.014, −0.004)
Control variables	
Rental characteristics	
Number of rental days	0.008*
	(0.001, 0.015)
Cumulative total number of interactions (per 1000)	−0.007*
	(−0.015, 0.001)
Renter characteristics	
Renter age	0.003***
	(0.002, 0.004)
Renter member, number of years	0.008
	(−0.025, 0.041)
Owner characteristics	
Owner, number of positive ratings	−0.002
	(−0.007, 0.004)
Owner, number of negative ratings	−0.027
	(−0.060, 0.007)
Owner age	0.006***
	(0.004, 0.008)
Owner member, number of years	−0.069***
	(−0.103, −0.035)
Motorcycle characteristics	
Day price	0.000
	(−0.002, 0.001)
Motorcycle, number of positive ratings	0.002
	(−0.001, 0.005)
Motorcycle, number of negative ratings	0.016*
	(0.001, 0.031)
Motorcycle age	−0.003*
	(−0.006, 0.000)
Motorcycle weight	0.000
	(−0.001, 0.000)

continued

Downloaded from https://academic.oup.com/ser/article/20/4/1905/6252982 by Universiteitsbibliotheek Utrecht user on 14 December 2022

Table B1 *Continued*

Independent variables	Model 1
Motorcycle CC	0.000
	(0.000, 0.000)
Motorcycle HP	0.000
	(−0.001, 0.001)
Variance intercept renter-level	0.002^{***}
	(0.001, 0.005)
Variance intercept owner-level	0.042^{***}
	(0.035, 0.050)
n	7181

Note: 95% confidence interval in parentheses.
$^{***}P < 0.001$,
$^{**}P < 0.01$,
$^{*}P < 0.05$ (one-tailed tests).

Downloaded from https://academic.oup.com/ser/article/20/4/1905/6252982 by Universiteitsbibliotheek Utrecht user on 14 December 2022