The Effect of User Interactions on Shaping Online Trust: Evidence from a Large-scale Experiment

Research-in-Progress

Chenyang Zhu
Shanghai University of Finance and Economics
Shanghai, China
chenyang.zhu@163.sufe.edu.cn

Bruno Abrahao
NYU Shanghai
Shanghai, China
abrahao@nyu.edu

Anqi Zhao
National University of Singapore
Singapore
staza@nus.edu.sg

Paolo Parigi
Institute for Research in the Social Sciences, Stanford University
Stanford, CA
pparigi@stanford.edu

Abstract

Trust is critical to the healthy function and growth of organizations. In particular, the success of online platforms of resource exchange, which depends on enabling trust between strangers, hinges on understanding factors that contribute to the engineering of trust. While reputation systems have proved effective in fostering trust and in offsetting prevalent social biases, it has been challenging to measure the extent to which having a peer-to-peer experience shapes judgment of trustworthiness, both in other members and in the platform itself. We draw causal conclusions from a longitudinal experiment that tracked 3,374 Airbnb users over time. We found that the average causal effect of a good user interaction enhances trust in the platform while reducing the importance users place in social similarity in their decision-making process. The effect is homogenous across groups of different socio-demographic features, which shows evidence that almost all subgroups benefit from positive user interactions.

Keywords: sharing economy, natural experiment, trust, experience, causal inference

Introduction

The emergence of online platforms operating in the sharing economy space is redefining business practices and disrupting traditional markets. A large number of Internet users currently rely on sharing economy services, such as ridesharing, temporary rental accommodation in other people's homes, the hire of a "tasker", etc. These platforms offer the possibility to experience interactions and exchanges with perfect strangers, who often share little or no prior information about each other. As a result, trust becomes the core component to bridge the information gap between strangers. Accordingly, the success, healthy function, and growth potential of companies in the sharing economy rely heavily on the engineering of trustworthiness and on creating a sense of belonging to a community of strangers whose members have shared goals.

Previous research has shown that social biases based on demographic characteristics, such as race, region, and gender, pose difficult challenges to the development of trust on online platforms, thereby hindering the vision these companies share (Abrahao et al. 2017) and potentially limiting their growth.

The tendency to trust similar others, known in the social sciences literature as *homophily*, is a major source of preexisting social biases (Mcpherson et al. 2001) and results in the selection of partners as a byproduct of socio-demographic characteristics. Researchers have quantified the effects of reputation systems, an artificially engineered feature of these platforms, in offsetting these naturally rooted biases (Abrahao et al. 2017). Reputation systems, such as those based on reviews and ratings, build trust at scale by recording a community-generated measure of aggregated evaluations for the actions users take on the platform. The platform implements these systems with the hope that they serve as sources of signals that may redirect the decision-making process. By helping users assess trustworthiness, they remove hurdles to the establishment of mutual trust, which generate paralysis (Snijders 2009; Resnick et al. 2002).

In this work, we address a fundamental complementary question regarding the dynamic of trust in the context of mutual interactions on the platform. In particular, through a natural experiment, we investigate the degree to which different levels of interactions and experiences, affect the degree of trust users have in others and in the platform, as well as their inclination to operate under the influence of social biases. We found that positive experiences causally improve trust in both the platform and in other users with good reputations.

Background and Previous Work

Long before the sharing economy emerged, trust has proved critical in traditional virtual communities by facilitating information exchange (Ridings et al. 2002) and increasing commitments to an online community (Wu et al. 2009). As exchanges and products are introduced in online communities, the source and target of trust has become more crucial and multidimensional in e-commerce and sharing economy platforms. Specifically, trust can be decomposed into three main targets, namely, trust towards peers, towards platforms, and towards products (Hawlitschek 2016). The interaction of these three elements serves as a new framework to study the dynamics of trust in the sharing economy and many recent works address the relationship of either two or all of the elements.

Sociologists have established that people tend to build connections with others that are similar to themselves in sociodemographic features, such as gender, age, marital status, etc. (Mcpherson 2001 et al.). However, the propensity to choose familiar others, or *homophily*, often adds inefficiencies to markets. This problem is particularly pronounced in online platforms. According to a recent study on a sharing platform, a host who rejects guests based on their race incurs a median loss of revenue around 65 to 100 dollars (Edelman et al. 2017), showcasing the extent to which homophily is causing moral concern and economic loss on online platforms. Platforms invariably deploy online reputation systems as a mechanism to signal trust, hoping to offset the friction that results from lack of information, homophily, and other biases. Researchers have found that reputation systems serve as an effective way to stimulate exchanges on online platforms and provide trustworthiness in uncertain environments (Snijders 2009; Dellarocas 2003), as well as to offset biases in judgments of trustworthiness (Abrahao et al. 2017). As the reputation system enables trust at scale, it reduces the cost by building trust at the community level as opposed to at the individual level (Ba 2001).

On a different aspect of trust dynamics, studies about the impact of products (or experiences) on platforms and peers often foreshadow the complexity of modeling user behavior when experiences are involved. A common ramification is the fallacy of composition, also known as substitution effect from social psychology, which often arises because platform users tend to evaluate the trustworthiness of an entire platform based on their few experiences (Bolton 2004; Kahneman 2011). If their experiences happen to be satisfying, the positive interactions would increase the member's trust in the platform and also help build up a sense of commitment to the community (Wu et al. 2009). When it comes to the platforms that involve frequent user-to-user interactions, only a few works have addressed the dynamics of the more diversified and more frequent experiences. A study on Airbnb users shows that guests are more likely to post a review (viewed as an act of engagement with the platform) when the real experience deviates from neutral (Bae 2017). However, how this significant effort to engage, i.e., post a review, shapes trust in others or in the platform remains unexplained.

The preceding question has remained open in studies of online trust. In traditional settings, the inability to isolate confounds due to the inflexibility and potential economic losses or ethical concerns of manipulating experience, e.g., by randomly assigning good or bad experiences to users, has refrained researchers from investigating this question with controlled experiments or A/B testing. Researchers frequently rely on limited information to study organizations, which is often insufficient for understanding the dynamics of trust between peers and between actors and platforms. Even with observational studies that utilizes a rich set of observational data, the analysis remains obscure for lack of causal arguments. In this work, we address these problems by using causal inference on a longitudinal natural experiment conducted on Airbnb users. In this setting, we can manipulate a rich set of socio-demographic features that underpin the properties of the interactions while investigating the effects of treatments on this population.

Study Design

Our work analyzes data from a longitudinal experiment, which tracked 3,374 Airbnb users over time (Abrahao et al. 2017). This allows us to to establish the causal effect of user interactions on trust on a sharing economy platform, using a diverse population with respect to socio-demographic backgrounds. We selected a random population of 100,000 users from Airbnb's database, while controlling for their roles as hosts or guests, the number of previous experiences, and whether or not they have future travel bookings in the platform. We built an independent website that implements our experiment, completely separated from the main Airbnb's platform.

Our design guidelines for the experiment aimed to mimic the process of selecting hospitality partners, i.e., host or guest, with limited information, while reflecting the risks users may expose themselves when selecting their partners. Moreover, a desirable feature of the experiment is to allow for comparability with previous literature. In light of this, instead of replicating the Airbnb experience, which is specific to that platform, we employed a more general framework from behavioral economics, namely an investment game (Berg et al. 1995; Houser et al. 2010). We believe the investment game reflects the same trade-offs users face in the platform when assessing risk via trust judgments while providing us with the ability to align the results with an extensive literature that spans over several decades (Berg et al. 1995; Houser et al. 2010).

The respondents are prompted to play an investment game with five other Airbnb profiles. Although the profiles are synthetic and engineered in advance, the experiment told the participants that they correspond to profiles of real users. The profiles are generated with different social distances from the participants, based on the information they entered when registering for the game. Here, social distance is a measure of the distance between two participants, computed with the Hamming distance between users' feature vectors, which include age, gender, region, and marital status. (Mcpherson 1983; Blau 1977). This design allows us to measure the similarity and identify possible bias, if any, between different users. For each participant, three profiles are generated with distance $d = 0,1,2$. In other words, the profiles are different in d features. We generate two other profiles with distance $d = 4$, which have all features different from the participant.

Participants are randomly selected into two worlds, in order to take reputation levels of profiles into consideration. A profile's reputation is constructed by star ratings from 0 to 5 and the number of reviews it has. In both worlds, the reputation levels of the profiles with $d = 0,1,2$ and one profile with $d = 4$are set to be the same - the baseline reputation. The difference between the two worlds lies in the second profile with $d = 4$, which we now call $d = 5$ for convenience. In world 1, profiles with $d = 5$have a worse reputation than baseline. The purpose of this world is to test the extent to which social biases based on social distance interferes with trustworth judgement. Accordingly, in world 1, the most distant profile from the participant turns out to also be the one with the lowest reputation. In other words, in world 1, reputation does not alter the inclination to distrust a distant profile, and the main factor that distinguishes the profiles is social distance. In world 2, the profile at $d = 5$ has a better reputation than baseline. This setup allows us to learn our participant's trade-off in their decision-making process: a cognitive tension is introduced, as the most socially distant profile has the best reputation.

To play the game, we ask each subject to invest (or save) a fraction of a total amount of 100 credits in these profiles. The receivers may return a portion of the inflated (tripled) investment to the investor (or keep the whole investment, thereby defecting). In this class of games, the amount invested serves as a proxy of trust in each profile (Berg et al. 1995; Houser et al. 2010). In roughly six weeks after the first round of investment, we invite the participants to play the exact same game again, with the same profiles they had seen before, as we aim to analyze changes in behavior.

The treatment in our experiment is constructed by tracking users' traveling and hospitality histories between the two phases, as guest and host, from Airbnb's internal database. We take the 1,833 users who interacted using the platform between the two phases to form a treatment group, while users who did not interact make up a control group. In this paper, an interaction is defined by the history that users either traveled somewhere as guests or hosted someone as part of being members of Airbnb.

We also obtained the trip ratings of these interactions from Airbnb's database. The ratings show a skewed distribution where 75% of all trips are 5-star (out of 5-stars). Thus, we restrict the scope of our analysis to trips that were rated 5-star and define the treatment as having a satisfactory hospitality experience (5 stars) between the two phases. The control group is then defined as having no trips, thus no user-to-user interactions. Since units are not randomized to treatment or control, we matched each participant in the treatment group with a subject in the control group, based on socio-demographic features, reputation, and traveling history (Parigi et al. 2017). Thus, we are able to estimate causal effects and to explain the heteroskedasticity between groups. Different covariates are put together by weighting the inverse of their variances, and the matching is conducted separately in different worlds.

Results

We first estimate the causal effect of the treatment from the change of average investment from phase 1 to phase 2 in the investment game. The estimation is carried out across different social distances, thus accounting for the tradeoff between social distance and reputation level. We then proceed to estimate the causal effect with discrete choices, or the social distance of the profile that received the largest investment from a participant. Discrete choices focus on the participant's ordinal preference of profiles at different social distances. In the second part, we investigate whether the causal effect of positive user interactions would be different among socio-demographic groups.

1. The Average Causal Effect of Good Experiences on Shaping Trust

The figure above shows to what extent users with good experiences (treatment) changed their investment behavior in phase 2, compared with users with no-experience (control). Figure 1 shows for each social distance (x-axis) the average treatment effect of a good experience on the change of investment decisions (y-axis) after matching users into pairs. The results are shown separately for world 1 and world 2. In world 1, participants in the treatment group save 0.89 less than in control on average, with a 95% confidence interval [-1.315, -0.467]. Therefore, when users made decisions solely by social distance as in world 1, a good experience is very likely to boost the users' trust in all profiles, and thus in the platform. When there is a trade-off between reputation and social distance, as, in world 2, a significant change of trust in the platform is not always guaranteed. This could be explained by the variance of expectations users have when assessing from profiles exhibiting a dissonance between reputation and social distance.

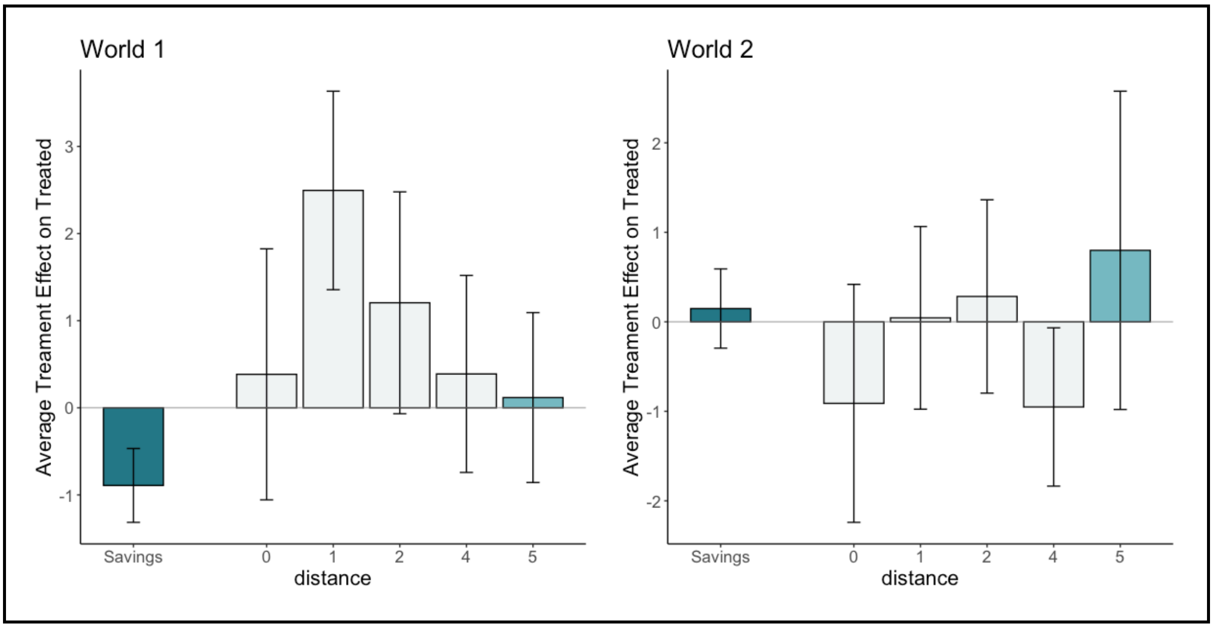

Figure 1. The average treatment effect of good experience on the treated group. Recall that profiles of distance from 0 to 4 have baseline reputations, and profiles of distance 5 have lower reputation in world 1 and higher reputation in world 2.

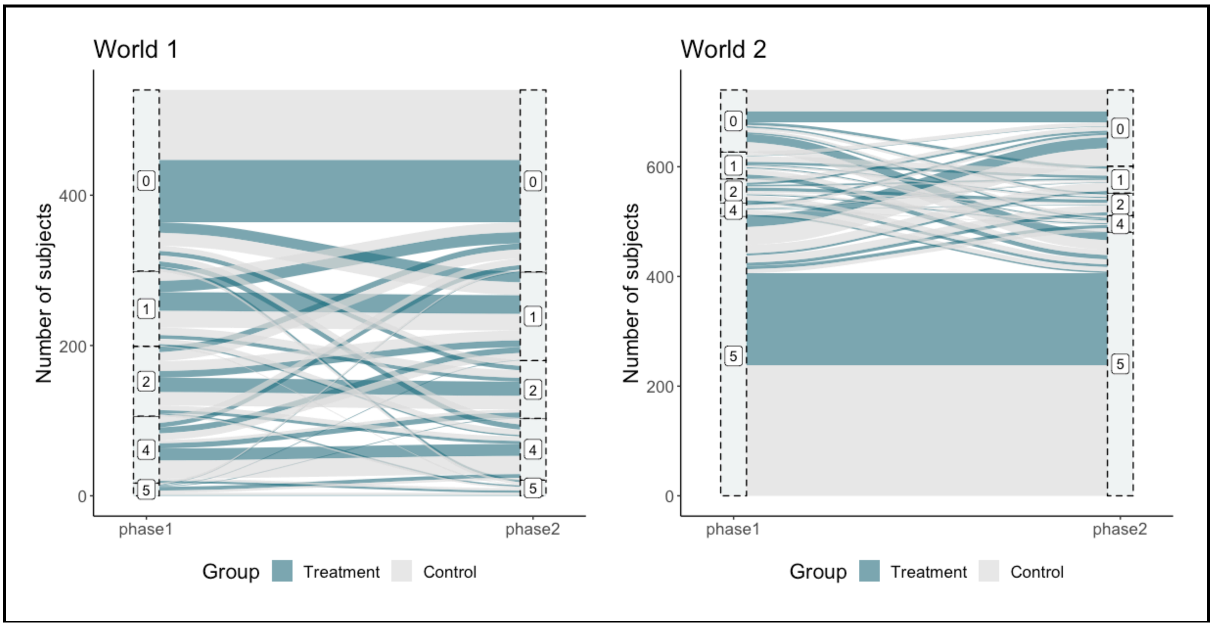

Figure 2. Change of the bulk from phase 1 to phase 2.

We further investigate the trade-off between reputation and social distance by looking at the change of the ordinal preference of each social distance between the two waves. We are most interested in the ordinal change of the distance of the profile that received the largest investment for each participant, in each phase, or the *bulk* of the investment in short. The Sankey plot of change in the bulk is shown in Figure 2, where for each world, the left column shows the bulk in phase 1 and the right column shows the bulk in phase 2. The dark color refers to the treatment group and the light one to the control group. Subjects that do not invest in any profiles are removed, resulting in a total of 1,479 subjects (624 in World 1 and 855 in World 2). Participants in world 1 tend to choose the highest reputation and more familiar distance 0 (275 in phase 1 and 273 in phase 2), while world 2 sees a larger proportion of choices in distance 5 (583 in phase 1, and 550 in phase 2). Comparing the two phases, more than half of the participants (352, or 56% in World 1 and 568, or 66% in World 2) stick with the same profile distance when assigning the bulk of the investments, supported by Cohen's Kappa coefficient shown in Table 1.A (Fleiss 1981). The Low. and Up. columns in the table refers to lower and upper bound.

Table 1. Model Results

A. Cohen's Kappa coefficient

World	Group	Low.	$\hat{\kappa}$ Est.	Up.
world 1	Treatment	0.31	0.38	0.45
	Control	0.32	0.40	0.47
world 2	Treatment	0.27	0.34	0.41
	Control	0.30	0.37	0.44

B. McFadden Discrete Choice Model

Param.	Value	Std.	|t| value
β_0	-0.598	0.029	20.717
β_1	0.546	0.127	4.307
β_2	-0.351	0.132	2.663

To figure out whether a good experience is a cause for the change in ordinal preference, we adopted the McFadden discrete choice model in panel data settings (McFadden 1973). This method is also known as the logistic regression model to an ordered factor response. For each participant, we denote $d_1, d_2 \in \{0, 1, 2, 4, 5\}$ as the social distance of profile that received the bulk of investment in phase 1 and phase 2. The model can be written as

$$logit\ P(d_2 \leq d) = f(d) = \eta_d + \beta_0 d_1 + \beta_1 1_{w1} 1_{Tr} + \beta_2 1_{w2} 1_{Tr}, \forall d \in \{0,1,2,4,5\}$$

where $\beta_0, \beta_1, \beta_2$ denote parameters to be estimated, $1_{w1}, 1_{w2}, 1_{Tr}$ denote indicators for world 1, world 2 and treatment, and η_k is the intercept. We run the model on the matched dataset. The results in Table 1.B reveal several key observations. The parameter of the bulk in phase 1, or β_0, shows that the choices in the two phases have a certain degree of agreement, which is consistent with results from Cohen's Kappa.

Secondly, the significant parameters β_1 and β_2 of two worlds show seemingly opposite treatment effects but actually point to the same origin. Take world 1 as an example, a treatment brings a 0.546 (± 0.254) increase into the cumulative logit function. Thus, $P(d_2 \leq 4)$(possibility to choose a better reputation) increases, while $P(d_2 = 5)$ (choosing a worse reputation) decreases. On the other hand, in world 2, $P(d_2 \leq 4)$ decreases, and $P(d_2 = 5)$increases. In other words, the probability of choosing profiles with better reputations increases in both worlds. Notice how the change of probability to invest the bulk to each distance $P(d_2 = d) = \sigma(f(d)) - \sigma(f(d-1))$, where $\sigma(x) = logit^{-1}(x)$ is the sigmoid function and $d \geq 1$, depends on which profile the subject chooses in phase one. However, the absolute change to distance 1, 2, 4 tends to be larger than other distances on average, because of the shape of the derivative of the sigmoid function. Thus, positive user interactions increase the probability to invest in less similar profiles in world 1 and profiles with a larger reputation in world 2.

These results strongly suggest the argument that good experiences causally improve the trustworthiness of the reputation system and the trustworthiness of less similar profiles, thus reducing the level of homophily in the participant's decision-making process.

2. Homogeneous Effects of Good Experiences across Sub-population

The second question concerns whether good experiences have the same effect across socio-demographic groups and for different numbers of times using the platform. We mainly consider the subject's socio-demographic features, age, gender, and marital status, and the number of trips as guests and hosts that the participants have experienced before phase 1, i.e., the pre-game users' history of Airbnb usage.

We fit a multi-level difference-in-difference regression model on the change of investment from phase 1 to phase 2. To study the interaction of treatment with these covariates, we put an indicator of treatment before each feature. The multilevel structure, which takes the change of investment on five social distances as layer one and each subject as layer two, is adopted to adjust for the covariance structure in the investment of each user. The parameter estimation is visualized in Figure 3. Each dot shows the treatment effect of each feature and its 95% confidence interval. The intercept, shown as the third feature in the graph, captures the average change of investment that is unrelated to treatments, showing that subjects were getting more familiar with the game and tended to increase their investment in general.

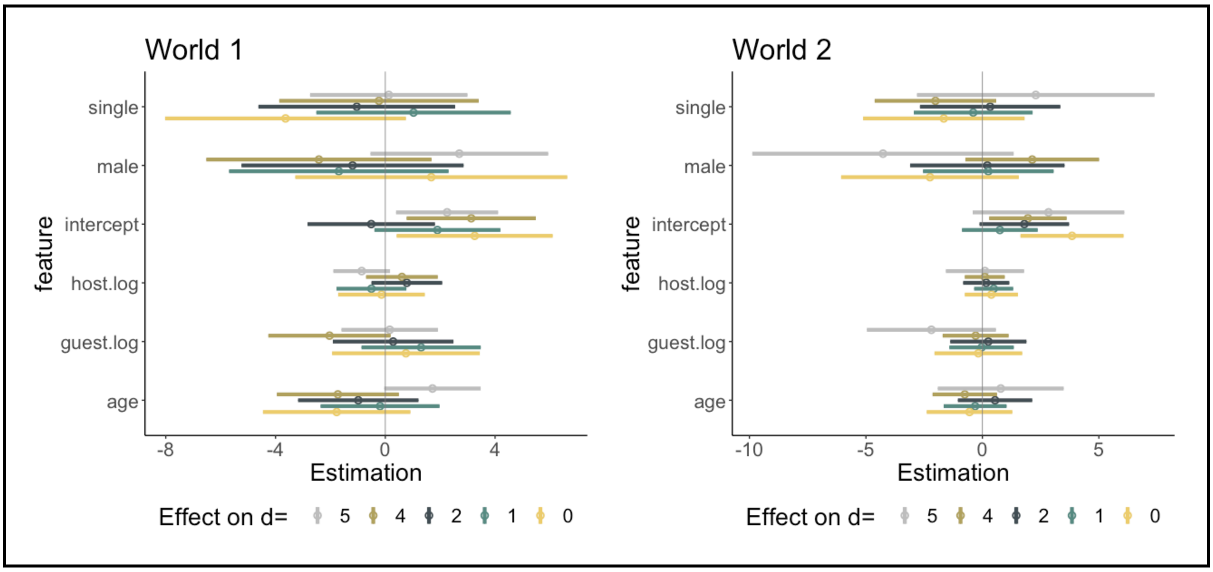

Figure 3. The effect of a good experience on different socio-demographic groups. Parameter estimation from the multi-level difference in difference model.

The above figure shows how a good experience affects the development of trust in groups of different socio-demographic features and different previous experiences on the platform. In both worlds, there is limited evidence that socio-demographic features determine the effect of good experiences, indicating that belonging to any of the subgroups or having different numbers of pre-experiment experiences on the platform are not indicators of how good experience could successfully change the trust towards peers and the platform. The effect of a good experience is close to homogeneous across subgroups in our experiment.

Conclusion

In this paper, we established through causal inference the effect of good experiences on shaping trust towards the platform and other peers in a platform of sharing economy. Several key finds are of theoretical interest and offer insights to answer fundamental questions in organizational function and evolution. Good experiences with other peers can causally increase users' trust in sharing platforms, and such experiences can mitigate (but not offset) the effect of preexisting homophily in the decision-making process. Users who had positive experiences tend to extend their trust to more distant others. There is also limited evidence that the effect of good experiences is indifferent across socio-demographic groups, thus resulting in a close-to homogenous uplift of trust in almost all subpopulations on the platform.

Acknowledgements

This work is supported by the National Science Foundation (NSF) grant #1257138. Abrahao was supported by a National Natural Science Foundation of China (NSFC) grant #61850410536.

Reference

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 (114:37), pp. 9848–9853 (doi: 10.1073/pnas.1604234114).

Ba, S. 2001. "Establishing online trust through a community responsibility system," Decision Support Systems (31:3), pp. 323–336 (doi: 10.1016/s0167-9236(00)00144-5).

Bae, S. J., Lee, H., Suh, E.-K., and Suh, K.-S. 2017. “Shared experience in pretrip and experience sharing in posttrip: A survey of Airbnb users,” Information & Management (54:6), pp. 714–727 (doi: 10.1016/j.im.2016.12.008).

Berg, J., Dickhaut, J., and McCabe, K. 1995. "Trust, Reciprocity, and Social History", Games and Economic Behavior (10:1), pp. 122-142(doi: 10.1006/game.1995.1027).

Blau, P. M. 1977. “A Macrosociological Theory of Social Structure,” American Journal of Sociology (83:1), pp. 26–54 (doi: 10.1086/226505).

Bolton, G. E., Katok, E., and Ockenfels, A. 2004. “How Effective Are Electronic Reputation Mechanisms? An Experimental Investigation,” Management Science (50:11), pp. 1587–1602 (doi: 10.1287/mnsc.1030.0199).

Dellarocas, C. N. 2003. “The Digitization of Word-of-Mouth: Promise and Challenges of Online Feedback Mechanisms,” SSRN Electronic Journal (doi: 10.2139/ssrn.393042).

Edelman, B., Luca, M., and Svirsky, D. 2017. “Racial Discrimination in the Sharing Economy: Evidence from a Field Experiment,” American Economic Journal: Applied Economics (9:2), pp. 1–22 (doi: 10.1257/app.20160213).

Hawlitschek, F., Teubner, T., and Weinhardt, C. 2016. “Trust in the Sharing Economy,” Die Unternehmung (70:1), pp. 26–44 (doi: 10.5771/0042-059x-2016-1-26).

Houser, D., Schunk, D., and Winter, J. 2010. “Distinguishing trust from risk: An anatomy of the investment game,” Journal of Economic Behavior & Organization (74:1-2), pp. 72–81 (doi: 10.1016/j.jebo.2010.01.002).

Kahneman, D. (2011). Thinking, fast and slow, New York: Farrar, Straus and Giroux.

McFadden, D. 1973. “Conditional logit analysis of qualitative choice behaviour,” Frontiers in econometrics, pp. 105-142

Mcpherson, M. 1983. “An Ecology of Affiliation,” American Sociological Review (48:4), p. 519 (doi: 10.2307/2117719).

Mcpherson, M., Smith-Lovin, L., and Cook, J. M. 2001. “Birds of a Feather: Homophily in Social Networks,” Annual Review of Sociology (27:1), pp. 415–444 (doi: 10.1146/annurev.soc.27.1.415).

Parigi, P., Santana, J., and Cook, K. 2017. "Online Field Experiments", Social Psychology Quarterly (80:1), pp. 1-19(doi: 10.1177/0190272516680842).

Resnick, P., and Zeckhauser, R. 2002. “Trust among strangers in internet transactions: Empirical analysis of eBay s reputation system,” The Economics of the Internet and E-commerce Advances in Applied Microeconomics, pp. 127–157 (doi: 10.1016/s0278-0984(02)11030-3).

Ridings, C. M., Gefen, D., and Arinze, B. 2002. “Some antecedents and effects of trust in virtual communities,” The Journal of Strategic Information Systems (11:3-4), pp. 271–295 (doi: 10.1016/s0963-8687(02)00021-5).

Snijders C, W. J. 2009. “Online programming markets”, in Cook, K., Snijders, C., Cheshire, C., and Buskens, V. 2009. ETrust: Forming Relationships in the Online World, New York: Russell sage Foundation.

Wu, J.-J., Chen, Y.-H., and Chung, Y.-S. 2010. “Trust factors influencing virtual community members: A study of transaction communities,” Journal of Business Research (63:9-10), pp. 1025–1032 (doi: 10.1016/j.jbusres.2009.03.022).