SOCIAL SCIENCES

Fighting bias with bias: How same-race endorsements reduce racial discrimination on Airbnb

Minsu Park[1]*†, Chao Yu[2]†*, Michael Macy[3,4]*

Recent studies have documented racial discrimination in online interactions, mirroring the historic bias observed offline. The sharing economy is especially vulnerable due to greater dependence on mutual trust in sharing a ride, residence, or date with a stranger. These services rely on user recommendations to build trust, but the effects of these peer evaluations on racial bias are only beginning to be explored. Using data from Airbnb, we examine in-group preference for same-race hosts as well as same-race recommendations. The unexpected result is that these two manifestations of racial bias are offsetting, not reinforcing. White guests largely overcame their racial bias in host selection when hosts were endorsed by previous white guests. Moreover, we found no evidence of racial bias in the affective enthusiasm of endorsements, which suggests that the preference for same-race endorsements is motivated by the race of the recommender, not the content of the recommendation.

INTRODUCTION

Over the past decade, the online sharing economy has flourished in several domains, including lodging, transportation, and even dating (1–3). This success is unexpected, given the risk to personal safety of sharing a home, accepting a ride, or going on a date with a complete stranger (4). Sharing economy platforms have addressed this challenge by designing user-driven reputation systems (e.g., reviews, badges, and ratings) that discourage misbehavior and encourage mutual trust (1, 4). In addition, most platforms also display users' names and photos, thereby reducing anonymity and conveying authenticity and accountability (1).

However, recent studies have documented an unintended consequence. Names and photos also reveal sociodemographic characteristics that can enable racial discrimination in users' choices of individuals with whom to interact (1–3, 5), thereby reinforcing racial barriers to participation in the sharing economy (5–9). These barriers are well documented. Black users wait longer for shared rides (10) and are more likely to have the driver cancel (3). Black hosts on Airbnb had 20% less demand than similar white-owned properties (11) and earn 12% less (12), while lodging requests from Black guests are 16% less likely to be accepted (2).

In response, Uber, a popular ridesharing platform, discontinued the display of racial identities of drivers and passengers when algorithmically optimizing matches (3). Although the design change was intended to combat racial bias, the strategy backfired. Although racial discrimination was avoided at the ride-request stage, it resurfaced in higher driver cancellation rates once the match was made and users' demographic information was revealed (3). Worse still, removing racial identities at the ride-request stage may have unintentionally strengthened the normative legitimacy of racial discrimination.

Previous research shows that reputation systems can reduce discrimination in the sharing economy (5, 13, 14). A study using 1 million Airbnb requests for accommodation found that the age and gender diversity of guests increased with the number of reviews a host received (13). Another study found that a single positive review of an Airbnb guest was sufficient to eliminate host reluctance to rent to guests with names that sounded African American (14).

Other studies show that users pay more attention to and place greater trust in third-party recommendations from in-group members. For example, Racherla et al. (15) found that "perceived similarity" based on the reviewer's photo, name, and bio-sketch increased the willingness to stay in a reviewer-recommended hotel, while inconsistent effects were reported in a similar study of digital camera recommendations (16). An experiment involving third-party racial similarity found that racially biased white participants were more responsive to criticism of racist language on Twitter when the criticism came from a user believed to be white (17). These findings are consistent with a half-century of social identity research on in-group bias that suggests greater responsiveness to same-race peer recommendations (18–20).

Growing evidence of in-group bias in third-party recommendations poses the central question that motivates our study: Does racial bias in the response to peer endorsements promote or attenuate racial bias in the selection of exchange partners in the sharing economy? The preference of whites for same-race recommendations, combined with racial bias in the strength of reviewers' endorsements, could reinforce racial discrimination against Black providers with few white customers and even fewer white endorsements. If so, previous studies may have overstated the effects of racial discrimination against Black providers by failing to control for the confounding effects of racial bias in third-party endorsements.

To find out, we analyzed data on host selection and peer recommendations in New York City using "instant bookings" collected from Airbnb, the leading online marketplace for shared accommodations. New York City has the booking density needed to control for neighborhood racial composition, and instant bookings are

[1]Division of Social Science, New York University Abu Dhabi, Abu Dhabi, UAE. [2]Department of Communication, Cornell University, Ithaca, NY, USA. [3]Department of Sociology, Cornell University, Ithaca, NY, USA. [4]Department of Information Science, Cornell University, Ithaca, NY, USA.
*Corresponding author. Email: minsu.park@nyu.edu (M.P.); cy395@cornell.edu (C.Y.); mwm14@cornell.edu (M.M.)
†These authors contributed equally to this work.

needed to isolate guests' racial preferences since they allow guests to select a host but do not permit hosts to choose a preferred guest. We used the Face++ facial recognition algorithm to classify users' race based on their profile photos (www.faceplusplus.com/). Visually revealed racial identities may differ from the user's self-identification, but Airbnb does not include race in user profiles and it is the racial identity perceived by guests that influences their behavior (see Materials and Methods for details on the data collection, facial recognition, and measurement of racial bias).

Our analysis focused on two patterns of possible racial bias among white users: guest preference for same-race hosts and guest responsiveness to same-race recommendations. We then tested whether racial bias in peer recommendations reinforces or attenuates racial bias in the selection of a host. We measured racial bias in host selection by comparing the probability that a white guest will select an Airbnb host, broken down by the race of the host. We measured host preferences among Black and Asian guests as a benchmark for the assessment of racial bias among white users. As an additional benchmark, we report the probability to choose a host whose race is unidentified (i.e., the host does not have a profile photo or their photo could not be algorithmically labeled by Face++).

Our focus on racial discrimination among white guests is not because they exhibit more pronounced same-race preferences compared to Black or Asian guests. On the contrary, all three racial groups share a similar preference for same-race hosts. We focus on white discrimination because of racial inequality in access to the sharing economy in general and the underrepresentation of disadvantaged minorities on Airbnb in particular. Same-race preferences are important for all three groups but for very different reasons and with very different policy implications. In-group bias among white guests reinforces racial inequality, while in-group bias among Black and Asian guests promotes more equal opportunity for hosts from underrepresented groups.

RESULTS

Preference for same-race hosts

In Fig. 1, the x axis represents the host's race and the y axis reports the probability of choosing the host relative to the expected probability if the choice were random. The expected probability corrects for the overrepresentation of white users on Airbnb (see Materials and Methods for additional details). The x axis also includes hosts that are racially unidentified, as a comparative baseline. The probability of choosing a racially unidentified host is close to the expected probability for each of the three guest groups, which confirms the randomization procedure used to account for the effects of the racial composition of hosts on Airbnb.

The results show that white guests are nearly three percentage points more likely to choose a white host ($M = 0.027$, $P < 0.001$) than would be expected by chance (given the racial distribution of hosts in the Airbnb population). White guests are also below chance in their likelihood to stay with Black ($M = -0.046$, $P < 0.001$) and Asian hosts ($M = -0.016$, $P < 0.001$).

The same-race preference is not limited to white guests. Asian ($M = 0.030$, $P < 0.001$) and Black ($M = 0.052$, $P < 0.001$) guests are also more likely to choose same-race hosts. Although white and Asian guests have roughly equal same-group preferences,

white guests are more reluctant to stay with Black hosts than are Asian guests ($D = -0.041$, $P < 0.001$).

Same-race preference for a host does not necessarily reflect racial bias among white guests. An alternative explanation is racial inequality in the distribution of wealth such that white hosts can afford more premium amenities and better locations. An obvious problem with this explanation is that a racial disparity in property attributes cannot explain same-race preferences among Black and Asian guests. Nevertheless, we confirmed the robustness of the results reported in Fig. 1 controlling for a set of measures for the intrinsic appeal of the property, including price, amenities, privacy, and neighborhood racial composition (see tables S3, S5, and S7). We found little between-race difference in listing attributes and little change in the preference for same-race hosts when controlling for intrinsic appeal.

Preference for same-race endorsements

The results in Fig. 1 are consistent with previous studies on racial bias in the selection of exchange partners in the sharing economy. Same-race preferences in host selection could reflect racial bias in how guests write and respond to endorsements, but this possibility has received comparatively little attention in previous research. It is the problem to which we now turn.

Racial bias could affect how endorsements are written as well as how readers respond. We used VADER (Valence Aware Dictionary for Sentiment Reasoning) (21) to compare the level of enthusiasm between endorsements from same- and other-race guests (see the "Sentiment analysis" section in Materials and Methods for additional details). The sentiment score can range from −1 (most negative) to 1 (most positive), with scores between −0.05 and 0.05 considered to be neutral. The mean sentiment scores for endorsements written by same-race (0.968) and other-race guests (0.970) were close to 1.0 (the upper limit) and nearly identical ($P = 0.305$), indicating that there were almost no negative reviews and other-race endorsements were no less positive than endorsements for hosts of the same race.

Using the Linguistic Inquiry and Word Count (LIWC), we also tested six additional affective measures: certitude, moralization, politeness, prosocial, risk, and dissatisfaction (see fig. S7). For Black hosts, all three groups of guests used polite words more frequently than they did for other hosts, Black and Asian guests used prosocial language more often, Black guests referred to risk more often, and white guests expressed lower certitude. For white hosts, Asian guests expressed higher certitude. For Asian hosts, there were no statistically significant differences in content compared to other hosts. The magnitude of these differences was very small and only marginally significant, and there were no significant differences for any of the other three LIWC measures that we tested. In short, using eight affective measures in two lexical sentiment analysis tools, we found no evidence of higher enthusiasm in same-race endorsements. If there is racial bias in peer recommendations, it is mainly in the response to endorsements, not in their content.

We isolated guest responses to endorsements from responses to the host's race by using bookings with racially unidentified hosts. Airbnb displays up to six endorsements on the listing's front page, without the need for the guest to click through to see earlier reviews (see Materials and Methods for details on the measurement of guest responses). In Fig. 2A, the x axis reports the number of front-page endorsements from previous guests of the same race as the guest, among hosts with six front-page endorsements. Note that,
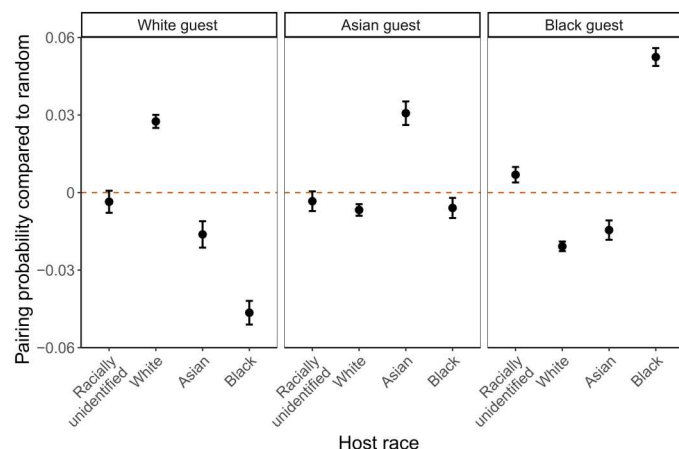
Fig. 1. Racial bias in host selection broken down by race. The y axis reports the probability to choose a host depending on the race of the host (x axis). Left to right: Each panel represents white, Asian, and Black guests' choices of hosts. As expected, racially unidentified hosts were close to chance in guest preferences. In contrast, guests clearly preferred same-race hosts. Black guests had stronger same-race preference than Asian ($D = 0.022$, $P < 0.001$) and white ($D = 0.025$, $P < 0.001$) guests. While Black and Asian guests had similar levels of reluctance to stay with other-race hosts, white guests preferred Asian over Black hosts ($D = -0.030$, $P < 0.001$).

in all cases, the hosts had six endorsements on their front pages; all that varies is the number of endorsements written by a previous guest of the same race as the current guest. We pooled five or more same-race endorsements together due to the smaller number of same-race reviews by Black guests (given the underrepresentation of Black users on Airbnb). The y axis is the probability of choosing a racially unidentified host relative to the expected probability if the choice were random, broken down by the race of the guest (identical to the measure used in Fig. 1 except that the host is racially unidentified).

The results show that the impact of same-race endorsement is greater than the impact of same-race host. The probability that a white guest chooses the host increases markedly with the proportion of endorsements that were from previous guests who were also white, from -0.240 to 0.056 points ($P < 0.001$) as the number of same-race endorsements increases from zero to five or more.

As with the same-race preference for hosts, the effect of same-race endorsements is not limited to white guests. The booking probability increases from -0.033 to 0.296 points ($P < 0.001$) for Black guests and from -0.025 to 0.189 points ($P < 0.001$) for Asian guests. Figure 2A also shows that the overall booking probabilities for white guests are much lower, compared to Black and Asian guests. This indicates that white guests are more reluctant than other guests to book with hosts whose race is unknown.

Given that the race of the host is unknown, a plausible interpretation is that guests use the racial composition of endorsements as a proxy for the unknown race of the host. We therefore replicated Fig. 2A using bookings where the race of the host is known; hence, the effect of same-race endorsement is not needed to infer the race of the host. The results in Fig. 2B are nearly identical to Fig. 2A except that the race of the host is known to the guest. The y axis measures the probability of booking with a randomly chosen host of known race, with equal weighting for hosts regardless of race. The probability that a white guest chooses a host of known

but random race increases from -0.195 to 0.045 ($P < 0.001$) as the number of same-race endorsements increases from zero to five or more. Same-race endorsements also increase the probability of a booking for Black (from -0.025 to 0.248, $P < 0.001$) and Asian guests (from -0.024 to 0.157, $P < 0.001$).

Racial bias in host selection (Fig. 1) and in the response to endorsements (Fig. 2) is consistent with decades of social identity research on in-group demographic preferences (18–20). Figure 3 addresses the central question that motivates this investigation: Are these two manifestations of racial bias mutually reinforcing? The results show the opposite. Figure 3 is identical to Fig. 2, except that the y axis now measures the booking probability when the race of the host is known, broken down by race. The figure reports results for hosts of the same race (orange) and different race (navy). The choices of other-race hosts were sufficiently similar that we simplified the analysis by combining both other-race groups (see fig. S1 for results with all nine combinations of the race of guest and host).

The results show that same-race endorsements markedly increase not only the probability that white guests will choose a host of the same race but also the probability that they will choose a host of a different race. Among white guests (upper panel), the probability of choosing a same-race host relative to chance increases from -0.155 to 0.067 ($P < 0.001$) and the probability for a host of a different race increases from -0.213 to 0.041 ($P < 0.001$) as the number of same-race endorsements increases from zero (all endorsements were other race) to five or more (all endorsements were the same race). The gap between the orange and navy lines diminishes as the number of same-race endorsements increases. This gap measures the difference in the effect of white endorsements, depending on the race of the host. The reduction in the size of the gap with the number of same-race endorsements indicates that the effect of white endorsements on white guests is stronger for Black and Asian hosts.

The middle (Asian guests) and lower (Black guests) panels show similar main effects of same-race endorsement but the opposite interaction with the race of the host. Among Asian guests, the booking probability with a host of the same race increases from -0.024 to 0.229 ($P < 0.001$) as the number of same-race endorsements increases from zero to five or more. The booking probability for a host of another race also increases from -0.027 to 0.131 ($P < 0.001$). Among Black guests, the booking probability increases from -0.011 to 0.350 ($P < 0.001$) and from -0.036 to 0.233 ($P < 0.001$) for a host of the same race and of another race, respectively.

Although the effect of same-race endorsements is similar for all three groups of guests, the interaction with the race of the host is very different. As the number of same-race endorsements increases, the gap between the orange and navy lines gets larger for Black and Asian guests, while it gets smaller for white guests (see fig. S2 for additional analysis). This suggests that, compared to Black and Asian guests, white guests are less affected by the race of the host as the number of same-race endorsements increases.

As with Fig. 1, we tested the robustness of the results reported in Fig. 3 controlling for a set of measures for the intrinsic appeal of the property (see tables S4, S6, and S8). Including the controls had no qualitative effect. We also replicated Fig. 3 using the total number of same-race endorsements out of all reviews, not just those on the host's front page (see fig. S3).
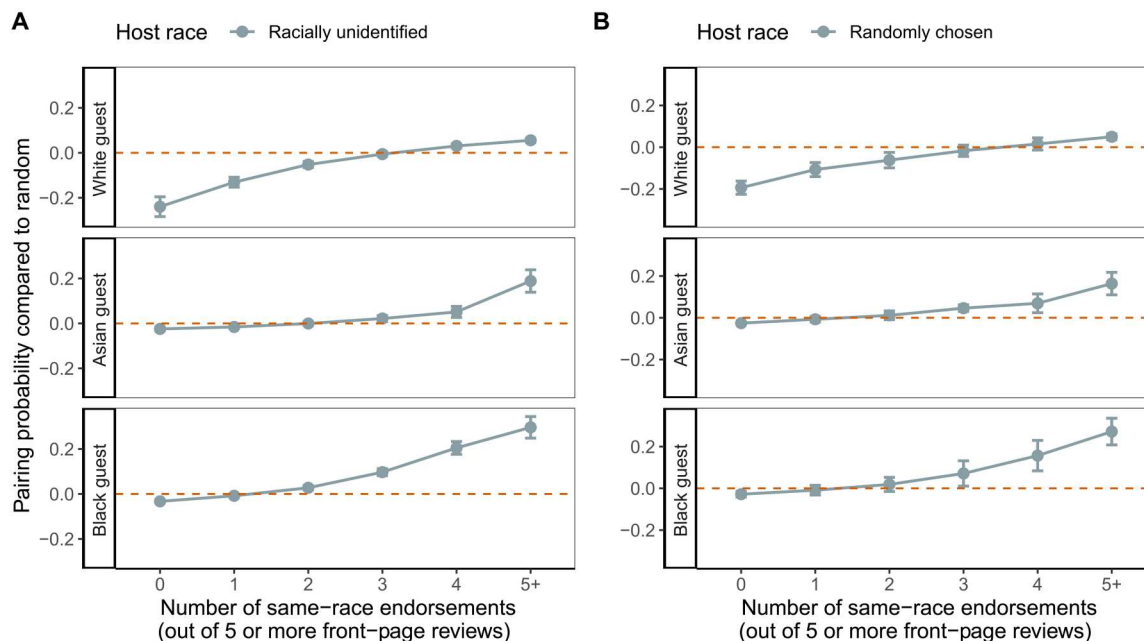
Fig. 2. Response to same-race endorsements of a racially unidentified or randomly chosen host. The *x* axis is the number of same-race endorsements for a racially unidentified host (**A**) and a randomly chosen host (**B**), out of five or more front-page reviews. The *y* axis is the probability that a white (top), an Asian (middle), or a Black guest (bottom) chooses a host, compared to chance. In (A), as same-race endorsement increases from zero to five or more, the normed probability of booking with a racially unidentified host (relative to chance) increases from −0.240 to 0.056 for white guests, from −0.025 to 0.189 for Asian guests, and from −0.033 to 0.296 for Black guests ($P < 0.001$ for all three groups). In (B), the probability of booking with a randomly chosen host of known race, with equal weighting for hosts regardless of race, shows almost identical results. As the number of same-race endorsements increases from zero to five or more, the normed probability increases from −0.195 to 0.045 points for white guests, from −0.024 to 0.157 points for Asian guests, and from −0.025 to 0.248 points for Black guests ($P < 0.001$ for all three groups).

DISCUSSION

Although racial discrimination is against the law in the United States (*22*), in-group preferences still affect many daily behaviors, such as choosing friends and neighborhoods (*23–26*). These in-group preferences are also evident in the sharing economy, from rides to dating (*2*, *3*). Using data from Airbnb in New York City, we confirmed previous studies showing that racial bias exists in online lodging markets as well.

We also tested for racial bias in peer recommendations. We found no difference in the affective strength of endorsements for same- and other-race hosts. Instead, racial bias in endorsements occurs in the response that they elicit. The race of endorsers had even more impact than the race of hosts. Yet, unexpectedly, the higher responsiveness to same-race endorsements had a net positive effect on guests' willingness to choose a host from a race other than their own.

The effect of white endorsements on the willingness of white guests to book other-race hosts is not because white guests write more positive reviews. On the contrary, nearly all reviews are positive, regardless of the race of the host or the reviewer (*27*, *28*). Moreover, guests of all races are more responsive to same-race endorsements, which suggests that it is the race of the reviewer, and not racial differences in review content, that accounts for the same-race responsiveness.

Nevertheless, there is an important difference in the effect of same-race endorsements on the willingness to book with other-race hosts. Because the market is white dominated, a Black or Asian host is more likely to be endorsed by white guests than is a

white host to be endorsed by Black or Asian guests. Thus, the net effect of same-race endorsement is to reduce racial discrimination on Airbnb.

These results have policy implications for reducing racial inequality in access to the sharing economy. Complaints about this inequality have led some companies to obscure the racial identity of users by removing profile photos (*3*). However, this strategy can also backfire, by shifting discrimination to the point where racial identities become apparent and by undermining trust when transactions are between anonymous users (*3*, *29*).

Instead, our study suggests that it may be possible to algorithmically harness racial bias to combat racial inequality. Same-race booking preferences limit opportunities for Black hosts to be booked by, and thus reviewed by, guests of other races. Instead of hiding profile photos, platforms should selectively display the same racial composition of front-page endorsements for all providers, creating a level playing field in the positive effects of same-race endorsements on the willingness of white guests to book with hosts of a different race. Increasing the exposure of white guests to white-authored endorsements of Black hosts may lead to more white bookings, thereby making the algorithmic correction less necessary over time.

Our study has several limitations. First, despite the high (68%) reviewing rate (*30*), guests who did not leave reviews are missing from the data. This imposes a selection bias in favor of previous guests who were motivated to write reviews. Although this makes the endorsements unrepresentative of all user experiences, what matters for host selection is the racial distribution observed by the
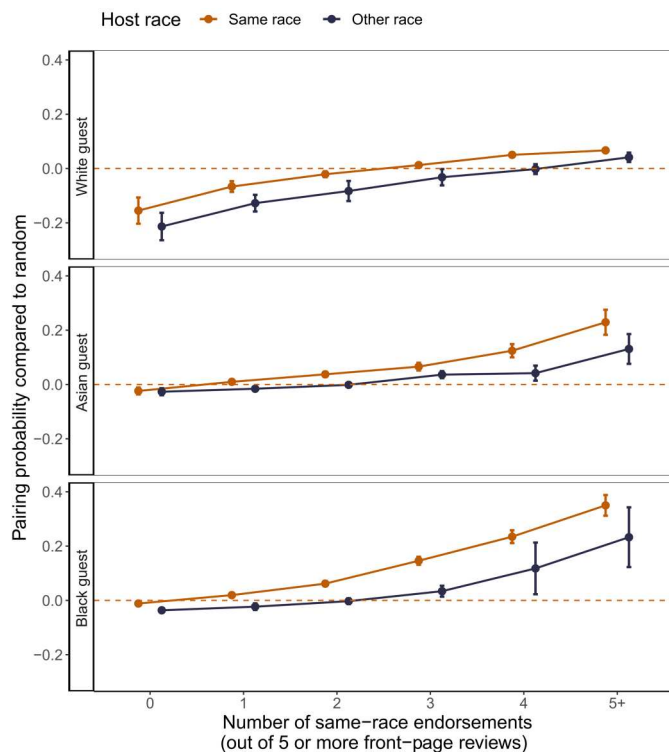
Fig. 3. Response to same-race endorsements of same- and other-race hosts. The x axis is the number of same-race endorsements of the host (out of five or more front-page reviews). The y axis is the probability that a white (top), an Asian (middle), or a Black guest (bottom) chooses a host, compared to chance. Color indicates whether the race of the host is the same as the guest. As same-race endorsement increases from zero to five or more, the normed probability of booking with the host increases, regardless of the race of the guest or host. However, among white guests, the increase in booking probability is larger for other-race hosts (D = 0.254, compared to D = 0.222), while among Black and Asian guests, the increase is larger for same-race hosts (D = 0.361 and D = 0.253), compared to other-race hosts (D = 0.269 and D = 0.158), among Black and Asian guests, respectively.

guest, not the true underlying distribution which the guest has no way to see.

Second, hosts also review guests, but we were unable to obtain these data. Future research is needed that uses host-authored endorsement of their guests to test for racial bias among hosts in the endorsement of guests and how these endorsements might affect racial diversity of reviews written by guests.

Third, we cannot rule out the possibility that our measure for same-race preference is confounded by unmeasured listing attributes that make guests more likely to book with and recommend the host. We tested the robustness of the results when controlling for privacy, amenities, price, and neighborhood racial composition, but a property can have special attributes (e.g., proximity to the beach and view of the mountains) that we cannot measure and therefore cannot control.

Fourth, we used time slices (before and after the start of 2017) to confirm the robustness of our temporally aggregated results (see the "Robustness check for temporal differences" section in the Supplementary Materials). However, our analyses were static and do not examine trends in racial bias. Changes to the Airbnb booking

system and user interface may have altered the behavior and racial composition of users during the time period for which we have data (2009 to 2018). Future research could leverage these changes to open up opportunities for causal inference that go beyond the descriptive account presented here.

Fifth, the racial composition and distribution of guest-host pairings could differ markedly from one neighborhood to another. We confirmed the robustness of our results by replicating the analyses using neighborhood-specific measures (see the "Robustness check for spatial differences" section in the Supplementary Materials), but future research could go further by using neighborhood variation to investigate ecological and contextual constraints on same-race preferences.

Sixth, our data are limited to users in New York City, which may understate the bias that might be observed in less cosmopolitan locations. However, New York City afforded the population density and diversity needed for the neighborhood-specific measures used to confirm the robustness of the spatially aggregated results.

Last, we did not investigate gender differences in racial bias. Although gender distributions do not vary by race in the larger population, gender is known to strongly influence pairing probabilities in the sharing economy (31). We controlled for gender to confirm the robustness of the reported results (see the "Robustness check for intrinsic appeal" section in the Supplementary Materials), but future research should focus on the intersection of gender and race in guests' booking behavior, reviews, and responses to endorsements.

MATERIALS AND METHODS
Dataset description
All data used in this study were obtained from Inside Airbnb (http://insideairbnb.com/), an independent third-party platform that specializes in the collection and distribution of Airbnb data. The data were collected for bookings in New York City between 12 March 2009 and 6 August 2018. We only collected listings that were "instant bookable," i.e., a listing that a guest can book immediately without needing to send a reservation request to a host for approval. Instant bookings are needed to isolate racial bias in the selection of a host from bias by hosts in approving of a guest. The dataset contains 13,663 instant bookable listings from 11,633 hosts and 407,295 time-stamped reviews from 377,086 guests. These data represent 34.6% of total listings and 40.4% of total reviews before excluding those that allowed hosts to reject guest applications. The rates of instant bookings are nearly identical across racial groups. After removing guests who could not be racially identified, the data used for the reported analyses consisted of 150,570 time-stamped guest reviews of 7147 hosts across 7903 instant bookable listings. To confirm the robustness of our results based on instant bookings, we replicated the analyses only with requested bookings and the results are almost identical (see the "Robustness check for booking type" section in the Supplementary Materials).

Racial classification using facial recognition
Airbnb guests cannot look up a host's racial identity in their user profile, nor can guests see the host's last name, which previous studies have used for labeling the race of online users (31, 32). However, guests can infer the race of a host or reviewer from their profile photo. (Airbnb recommends that users upload their

real profile photos and informs users that many hosts require that guests have a profile photo.) We have no way to know if the host's race, as perceived by the guest, matches the host's own racial identity. However, it is the perceived race of the host that should be used to measure a guest's racial bias. For example, we regard the choice of a host as racially biased if a white guest books with an other-race host mistakenly perceived to be white. It is the guest's perception that matters, not the host's own racial identity.

Like the guests, we also do not know the race with which the host self-identifies, and like the guests, we used profile photos to ascertain a user's race. The photos were labeled on Face++ (www.faceplusplus.com/), an online facial recognition service, because the number of users in our dataset was too large to use human coders. Face++ provides a reliable pretrained facial recognition algorithm that identifies demographic attributes including age and gender as well as race. Face++ classifies users into three racial categories: white, Black, and Asian. The racial distribution among the profile photos in our dataset was as follows: 59.82% of guests and 57.24% of hosts were white; 16.21% of guests and 22.01% of hosts were Black; and 23.97% of guests and 20.75% of hosts were Asian. As a reference point, the racial composition of New York City is 41.3% white, 23.8% Black, and 14.3% Asian in 2021 and 44.6% white, 25.1% Black, and 11.8% Asian in 2010, according to U.S. Census Bureau QuickFacts (www.census.gov/quickfacts/). The comparison with Census data indicates that white users are greatly overrepresented on Airbnb, while Black users are underrepresented (assuming random error in the visual classifications).

Face++ racial classification has been validated using the Chicago Face Database (82.8% accuracy) and the 10k US Adult Faces Database (72.9% accuracy) (33). However, those tests did not use Airbnb profile photos. We therefore tested for accuracy using a stratified random sample of 3000 profile photos matched with user IDs in our dataset. The photos were obtained using Airbnb URLs and were not locally stored. Each profile picture was evaluated by three independent coders recruited from Amazon Mechanical Turk. Each coder was asked to assign a perceived race with an option of "Other/None of the Above." Each profile picture was labeled with the race identified by at least two coders ($N = 2872$) and otherwise labeled as ambiguous ($N = 128$). The average agreement between Face++ and human coders was 83.3%. Although the agreement is less than ideal, the errors appear to be random, which means that the racial bias that we report is, if anything, underestimated.

We labeled users as racially unidentified if they did not have a profile photo or their photo did not contain an algorithmically detected individual human face; 89.69% of guests and 80.97% of hosts in our data used images of people with identifiable human faces as their profile photos. Among users with algorithmically identifiable photos, 20.62% of guests and 11.79% of hosts had profile photos with more than one face, nearly all of which show the faces of a same-race couple or family. We excluded users with more than one face in their profile photos for three reasons: Multiple faces in profile photos are susceptible to greater error in algorithmic labeling due to the inability to distinguish between background and foreground faces, reduced facial resolution in the background, and more opportunities for error (e.g., if the algorithm has p probability to correctly label a single face, the probability to correctly label all n faces is p^n). Instead, we assume that there are no qualitative differences in how guests respond to a host or reviewer between a single

individual and a couple. Photos with multiple faces suggest an opportunity for future research to test whether guest preferences for interracial hosts more closely match their preferences for same-race or other-race individual hosts.

Front-page reviews
Airbnb posts up to six of a host's most recent reviews on the host's front page, without the need for the guest to click through to see the others. Previous studies have shown that front-page reviews are the most influential (34, 35). However, the data we obtained do not identify which reviews are front page. We therefore used the review date to label up to six of the most recent reviews as front page. We validated the labeling method using 100 manually coded front-page reviews. Ninety-seven of 100 were correctly labeled using recency of the review as a proxy.

Our analysis of racial preference for same-race reviews is based on an equal number of reviews for all hosts, regardless of race—the six most recent reviews posted on the host's front page. Hosts with fewer than six reviews were excluded from the analysis and we pooled five or more same-race endorsements together to ensure sufficiently large numbers for statistical reliability. All analyses in the main text used front-page reviews, but replication using all endorsements confirms the robustness of the reported results (see fig. S3).

Observed and expected host-selection probability
We measured racial bias by comparing the observed probability of host selection with the probability expected if race were randomly assigned. Random assignment was implemented by randomly shuffling the guests to produce a set of random pairings while preserving the original racial distribution. The difference between the observed and expected probabilities then captures the effects of racial preference that are over and above the distributional effects. Statistical significance of the difference was estimated using the magnitude of the difference in standard errors (the SD of the expected probability over 1000 reshuffled samples).

Sentiment analysis
The number of same-race endorsements fails to capture differences in the level of enthusiasm expressed in each endorsement and may include reviews that are critical of the host. We used VADER to measure the difference in mean review sentiment between endorsements from same- and other-race guests. VADER is a lexical sentiment analysis tool that is tuned for short texts in online social media (21). We validated the lexical measures by randomly selecting 50 negative reviews and 50 positive reviews as coded by VADER. Two human coders then independently classified the reviews as positive or negative. The intercoder reliability measured by Krippendorff's alpha was 0.94 ($P < 0.001$) and the average agreement between human coders and VADER classification was 0.93, where 1.0 indicates perfect agreement. We excluded the small number of negative reviews (1.44%) and labeled the remaining reviews as endorsements.

VADER's limited number of sentiment measures led us to also measure enthusiasm using LIWC (the 2022 edition). LIWC includes 131 "certitude" words (including "really," "actually," and "of course"), 356 "moralization" words (e.g., "wrong," "honor*," "deserv*," and "judge") that are used to convey moral judgment, 142 "politeness" words (e.g., "please" and "thank you"), 242 "prosocial" words (e.g., "care" and "help"), 128 "risk" words (e.g., "secur*,"

"protect*," "pain," and "risk*"), 89 "lack" words (e.g., "want," "hope," "wanted," "wish," and "didn't have"), 56 "want" words, and, lastly, 49 "fulfill" words (e.g., "enough," "full," "complete," and "extra") that convey dissatisfaction. We binned all the reviews for each of the nine host-by-guest combinations of white, Black, and Asian users. For each bin, we counted the number of instances in which any word in the dictionary appeared in each bin and divided this number by the number of words in the bin. Each of the nine bins was analyzed using the following word lists: certitude, moralization, politeness, prosocial, risk, and dissatisfaction (combining "want," "lack," and "fulfill," aligned for similar direction) and we calculated the standard errors by bootstrapping 1000 randomly drawn samples of the same size (with replacement).

Ethical considerations

This research was reviewed by the host university's Institutional Review Board and classified as exempt because the data were publicly available (under CC BY 4.0), although the research involved the use of automated face classification (Face++). Face recognition algorithms have lower accuracy for members of underrepresented and marginalized groups that can exacerbate racial inequalities in applications that rely on face recognition as inputs for active user decision-making. However, our study only relies on face recognition to label data passively and therefore is not susceptible to the ethical problems that arise when the algorithms are used in user-driven applications.

Supplementary Materials

This PDF file includes:
Supplementary Text
Figs. S1 to S9
Tables S1 to S8
References

REFERENCES AND NOTES

1. K. Levy, S. Barocas, Designing against discrimination in online markets. *Berkeley Technol. Law J.* **32**, 1183 (2017).
2. B. Edelman, M. Luca, D. Svirsky, Racial discrimination in the sharing economy: Evidence from a field experiment. *Am. Econ. J. Appl. Econ.* **9**, 1–22 (2017).
3. J. Mejia, C. Parker, When transparency fails: Bias and financial incentives in ridesharing platforms. *Manage. Sci.* **67**, 166–184 (2021).
4. J. Santana, P. Parigi, Risk aversion and engagement in the sharing economy. *Games* **6**, 560–573 (2015).
5. J. D. Tjaden, C. Schwemmer, M. Khadjavi, Ride with me—Ethnic discrimination, social markets, and the sharing economy. *Eur. Sociol. Rev.* **34**, 418–432 (2018).
6. J. Kleinberg, J. Ludwig, S. Mullainathan, C. R. Sunstein, Discrimination in the age of algorithms. *J. Leg. Anal.* **10**, 113–174 (2018).
7. J. Kleinberg, J. Ludwig, S. Mullainathan, C. R. Sunstein, Algorithms as discrimination detectors. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 30096–30100 (2020).
8. Z. Obermeyer, B. Powers, C. Vogeli, S. Mullainathan, Dissecting racial bias in an algorithm used to manage the health of populations. *Science* **366**, 447–453 (2019).
9. P. Törnberg, How sharing is the "sharing economy"? Evidence from 97 Airbnb markets. *PLOS ONE* **17**, e0266998 (2022).
10. Y. Ge, C. R. Knittel, D. MacKenzie, S. Zoepf, Racial discrimination in transportation network companies. *J. Public Econ.* **190**, 104205 (2020).
11. A. Marchenko, The impact of host race and gender on prices on Airbnb. *J. Hous. Econ.* **46**, 101635 (2019).
12. B. G. Edelman, M. Luca, "Digital discrimination: The case of Airbnb.com," *Harvard Bus. Sch. NOM Unit Work. Pap.* (no. 14-054) (2014).
13. B. Abrahao, P. Parigi, A. Gupta, K. S. Cook, Reputation offsets trust judgments based on social biases among Airbnb users. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 9848–9853 (2017).
14. R. Cui, J. Li, D. J. Zhang, Reducing discrimination with reviews in the sharing economy: Evidence from field experiments on airbnb. *Manage. Sci.* **66**, 1071–1094 (2020).
15. P. Racherla, M. Mandviwalla, D. J. Connolly, Factors affecting consumers' trust in online product reviews. *J. Consum. Behav.* **11**, 94–104 (2012).
16. M. Zhao, J. Xie, Effects of social and temporal distance on consumers' responses to peer recommendations. *J. Market. Res.* **48**, 486–496 (2011).
17. K. Munger, Tweetment effects on the tweeted: Experimentally reducing racist harassment. *Polit. Behav.* **39**, 629–649 (2017).
18. J. C. Turner, in *Rediscovering Social Identity*, T. Postmes, N. R. Branscombe, Eds. (Psychology Press, 2010), pp. 243–272.
19. H. Tajfel, J. C. Turner, W. G. Austin, S. Worchel, An integrative theory of intergroup conflict. *Organ. identity A Read.* **56**, 9780203505984--16 (1979).
20. H. Tajfel, Social identity and intergroup behaviour. *Soc. Sci. Inf.* **13**, 65–93 (1974).
21. C. J. Hutto, E. Gilbert, Vader: A parsimonious rule-based model for sentiment analysis of social media text, in *Proceedings of the 8th International AAAI Conference on Weblogs and Social Media* (vol. 8, 2014), pp. 216–225.
22. D. A. Strauss, Discriminatory intent and the taming of brown. *Univ. Chicago Law Rev.* **56**, 935 (1989).
23. J. De Houwer, Implicit bias is behavior: A functional-cognitive perspective on implicit bias. *Perspect. Psychol. Sci.* **14**, 835–840 (2019).
24. L. Cosmides, J. Tooby, R. Kurzban, Perceptions of race. *Trends Cogn. Sci.* **7**, 173–179 (2003).
25. M. R. Banaji, A. G. Greenwald, *Blindspot: Hidden Biases of Good People* (Delacorte Press, 2013).
26. B. Hofstra, R. Corten, F. van Tubergen, N. B. Ellison, Sources of segregation in social networks: A novel approach using Facebook. *Am. Sociol. Rev.* **82**, 625–656 (2017).
27. G. Zervas, D. Proserpio, J. W. Byers, A first look at online reputation on Airbnb, where every stay is above average. *Mark. Lett.* **32**, 1–16 (2021).
28. J. Bridges, C. Vásquez, If nearly all Airbnb reviews are positive, does that make them meaningless? *Curr. Issues Tour.* **21**, 2057–2075 (2018).
29. A. C. Morales, Giving firms an "E" for effort: Consumer responses to high-effort firms. *J. Consum. Res.* **31**, 806–812 (2005).
30. A. Fradkin, E. Grewal, D. Holtz, Reciprocity and unveiling in two-sided reputation systems: Evidence from an experiment on Airbnb. *Mark. Sci.* **40**, 1013–1029 (2021).
31. G. M. Castex, Providing services to hispanic/latino populations: Profiles in diversity. *Soc. Work* **39**, 288–296 (1994).
32. N. Garg, L. Schiebinger, D. Jurafsky, J. Zou, Word embeddings quantify 100 years of gender and ethnic stereotypes. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E3635–E3644 (2018).
33. B. Jaeger, W. W. A. Sleegers, A. M. Evans, Automated classification of demographics from face images: A tutorial and validation. *Soc. Personal. Psychol. Compass* **14**, e12520 (2020).
34. B. Pan, H. Hembrooke, T. Joachims, L. Lorigo, G. Gay, L. Granka, In Google we trust: Users' decisions on rank, position, and relevance. *J. Comput. Commun.* **12**, 801–823 (2007).
35. T. Joachims, L. Granka, B. Pan, H. Hembrooke, F. Radlinski, G. Gay, Evaluating the accuracy of implicit feedback from clicks and query reformulations in Web search. *ACM Trans. Inf. Syst.* **25**, 7 (2007).
36. M. J. Salganik, K. E. C. Levy, Wiki surveys: Open and quantifiable social data collection. *PLOS ONE* **10**, e0123483 (2015).

Science Advances

Fighting bias with bias: How same-race endorsements reduce racial discrimination on Airbnb

Minsu Park, Chao Yu, and Michael Macy

View the article online
https://www.science.org/doi/10.1126/sciadv.add2315
Permissions
https://www.science.org/help/reprints-and-permissions

Use of this article is subject to the Terms of service

Science Advances

Supplementary Materials for

Fighting bias with bias: How same-race endorsements reduce racial discrimination on Airbnb

Minsu Park, Chao Yu, Michael Macy

Corresponding author: Minsu Park, minsu.park@nyu.edu; Chao Yu, cy395@cornell.edu; Michael Macy, mwm14@cornell.edu

This PDF file includes:

Supplementary Text
Figs. S1 to S9
Tables S1 to S8
Reference

Supplementary Text

Robustness check for number of reviews

Although the expected probability of same-race pairings depends on the racial distribution of users, the total number of bookings and reviews of a given host does not. Hosts accumulate more bookings (and thus more reviews) the more often they list their property. A White host who has just joined Airbnb will have fewer bookings and fewer reviews than a Black host who has listed their property for many years, regardless of the number of White and Black users on the platform. Moreover, guest preferences for same-race hosts cause the probability of a same-race booking to decrease, not increase, with the number of same-race listings, holding constant the number of same-race guests, due to greater competition among hosts for a limited supply of same-race guests.

This is confirmed in **Figs. S4-6**, which compare the number of reviews of White, Black, and Asian hosts. Most hosts have fewer than 50 reviews (**Fig. S4**). Across all the listings used to measure same-race preference in host selection (**Figs. 1-2** in the main text), while Black hosts have significantly more reviews than Asian ($D = 2.164$, $p = 0.050$), there is no significant difference in the number of reviews between Black and White hosts ($D = 1.439$, $p = 0.140$; see **Fig. S5**). Among the listings with six or more reviews used for analyzing the effects of same-race endorsement (**Fig. 3** in the main text), there is no significant difference in the number of reviews between hosts of different racial groups (see **Fig. S6**). In short, White hosts do not have significantly more reviews than Black or Asian hosts.

Robustness check for intrinsic appeal

Racial inequalities in wealth could be reflected in the attributes of properties listed by White hosts that make the listing more attractive and therefore more likely to be recommended and

chosen by guests, thereby confounding the effects of same-race preference. We tested for differences in intrinsic appeal by comparing the property attributes of White, Black, and Asian listings, including amenities, private vs. shared room, price, guest rating, neighborhood racial composition, host gender, and host reputation, as detailed below:

Amenities. Listings on Airbnb offer a wide range of amenities, from shampoo and smoke detectors to swimming pools, an indoor fireplace, and a doorman. We used a survey administered to a convenience sample ($N = 64$) to score all 103 amenities listed across the bookings in our dataset (excluding 21 accessibility-related accommodations such as "roll-in shower" and features without clear indications of intrinsic appeal, such as "long-term stay allowed"). Each participant was asked to choose which of two randomly paired amenities was more indicative of a "premium" listing. We used the "up" votes to score each amenity using the algorithm originally developed for Wiki Surveys (*36*) (see **Table S1** for the individual scores). Finally, we rated each Airbnb listing by summing the scores over all amenities listed for that property and then took the log of the summation (to correct for the skewed distribution).

Room type. Airbnb listings can be for the entire unit, a private room (with bathroom, kitchen, and living spaces shared with the host and/or other guests), or a shared room (dormitory style). Among the 7,903 listings, 52.12% were entire homes, 45.97% were private rooms with shared spaces, and 1.91% were shared rooms. We included binary variables for entire home and private room, with shared room (the least desirable) as the baseline.

Average price. Prices vary with the demand (e.g., around holidays), but the data from Inside Airbnb (http://insideairbnb.com/) only contain a listing's average per-night price in US dollars ($M = 137.13$, $SD = 177.31$). We used the natural log since the price distribution was highly skewed.

Guest rating. Airbnb hides the rating provided by each individual review but displays the average rating of each listing across all reviews. Airbnb displays the numerical scores as 0 to 5 stars with two decimal places ($M = 4.70$, $SD = 0.33$). We log-transformed the numerical scores to correct for the skewed distribution in which nearly all reviews are positive (*27, 28*).

Neighborhood racial composition. White preference for White hosts could be confounded by a preference for White neighborhoods. We obtained the racial distribution of each ZIP code in the New York metropolitan area from the 2016-2020 American Community Survey (https://data.census.gov/). Each property was scored with the proportion White in the ZIP code in which the property was located.

Gender. Guests may prefer female hosts, especially if the guest is also female. We therefore included binary variables to indicate the gender of the host ($N = 75,971$ female, $N = 74,599$ male) and the gender homophily of the host ($N = 78,647$ same gender).

Host reputation (cumulative review count). The number of reviews is a proxy for the number of bookings, and this "branding" can reduce guest uncertainty about a host's reputation. We measured the host's cumulative review count at the time of each booking ($M = 28.21$, $SD = 34.08$).

Table S2 reports the mean and standard deviation for each attribute, broken down by race of the host. While listings from White hosts charge a higher average price ($p < 0.01$) and have a higher guest rating ($p < 0.001$) compared to listings from Black or Asian hosts, their listings have a lower amenity rating (979.76) compared to listings from Black hosts (1006.47, $p < 0.001$).

We also tested for a potential confound by estimating pairing probabilities for White, Black, and Asian hosts while controlling for intrinsic appeal in a multinomial logistic regression.

For easier interpretation of the results, we computed estimated marginal means. **Table S3** reports same-race pairing probabilities with and without controls for intrinsic appeal.

We tested for confounding effects of intrinsic appeal on the response to same-race endorsements as well. **Table S4** reports estimated marginal means based on a multinomial logistic regression for the effect of same-race endorsements on the pairing probability, with and without controls for listing attributes. Same-race endorsements were measured as the number on the front page and the proportion out of all reviews. The results show that the response to same-race endorsements is largely independent of differences in intrinsic appeal.

Robustness check for booking type

Airbnb allows hosts to offer instant bookings (which the host cannot confirm) or requested bookings (which allow the host to refuse a guest's request). To confirm the robustness of our results based on instant bookings, we replicated the analyses using multinomial logistic regression for listings that include both booking types, i.e., instant and requested (see **Tables S5-6**) and listings with only requested bookings (see **Tables S7-8**). The results are almost identical to the results based on instant bookings (see **Figs. 1-3** in the main text and **Tables S3-4**).

Robustness check for temporal differences

The listings used in our analysis appeared as early as 2009 and were still active in 2018, and the results we report in the main text are based on data that are aggregated over this nine-year period. The aggregation could obscure two secular changes that may have occurred during that time. First, the results we report could overstate current levels of same-race preference due to earlier time periods when Airbnb was targeted with criticism for racial discrimination. This criticism led to changes in the user interface and booking system that may have altered user behavior as well

as the demographic composition of the user base. Second, the quality of listings may have improved as less marketable properties failed to survive competition for guests. In short, less biased guests and more marketable properties may be under-represented in the aggregate data on which our results are based, compared to data for the current period. However, due to the platform's exponential growth, over 80% of the data points are from the last two and half years of our data collection (from January 2016 to August 2018) and more than 60% are from the last 18 months. Hence, even if there were more racially biased guests and fewer marketable properties in earlier years, it seems unlikely that cross-temporal aggregation would have obscured a qualitative decline in racial bias or increase in listing quality.

Nevertheless, we replicated our analyses with two time slices of approximately the same size – one that covers one and a half years closest to the time of our data collection and a second slice that covers bookings made before 2017. **Fig. S8** confirms the robustness of our findings over different time periods. The figure shows almost identical results between the two time periods (panels **A-C** and **D-I** for the earlier and later time periods, respectively) and also between each period and the results we reported in the main text.

Robustness check for spatial differences

When a guest searches for a property, Airbnb returns a set of available listings that satisfy the guest's search criteria, and the racial distribution of this set could then be aggregated across all bookings. The expected value would then measure the probability of a same-race booking if guests picked a property at random from the set of properties displayed on their screen by Airbnb in response to their search criteria. However, these search-specific distributions are not made available by Airbnb. Instead, we calculated the expected probability of same-race pairings based on the racial distribution of all listings in our dataset across all locations. This spatially

aggregated value can be interpreted as the probability of a same race pairing if a randomly chosen guest booked a property at random from the set of all properties listed in the New York metropolitan area.

There are two important differences between these two methods. On the one hand, most searches are spatially constrained, and the search-specific distribution reflects the spatially constrained choices viewed by each guest – the choices most directly relevant to a guest when choosing to book a given property. On the other hand, the search-specific distribution also reflects potential racial bias of the guest that is embedded in the choice of search location (e.g., a White guest who searches for properties only in neighborhoods the guest believes are predominantly White). A racially biased expected probability cannot then be used to measure racial bias.

This trade-off can be reconciled by calculating the location-specific normed pairing probability at the time of booking and then aggregating the location-specific results to obtain the population-level probability as the weighted mean of means. Accordingly, we selected bookings made after January 1, 2018 and mapped the ZIP code of each listing to the corresponding neighborhood using United Hospital Fund (UHF 34) Neighborhood Index. We then calculated the pairing probability for each neighborhood. Finally, we aggregated the neighborhood-level probabilities as a weighted mean of means (weighted by the number of bookings in each neighborhood). **Fig. S9** shows nearly identical results whether the expected probability is based on the global distribution used in the manuscript or the aggregated neighborhood-specific distributions.
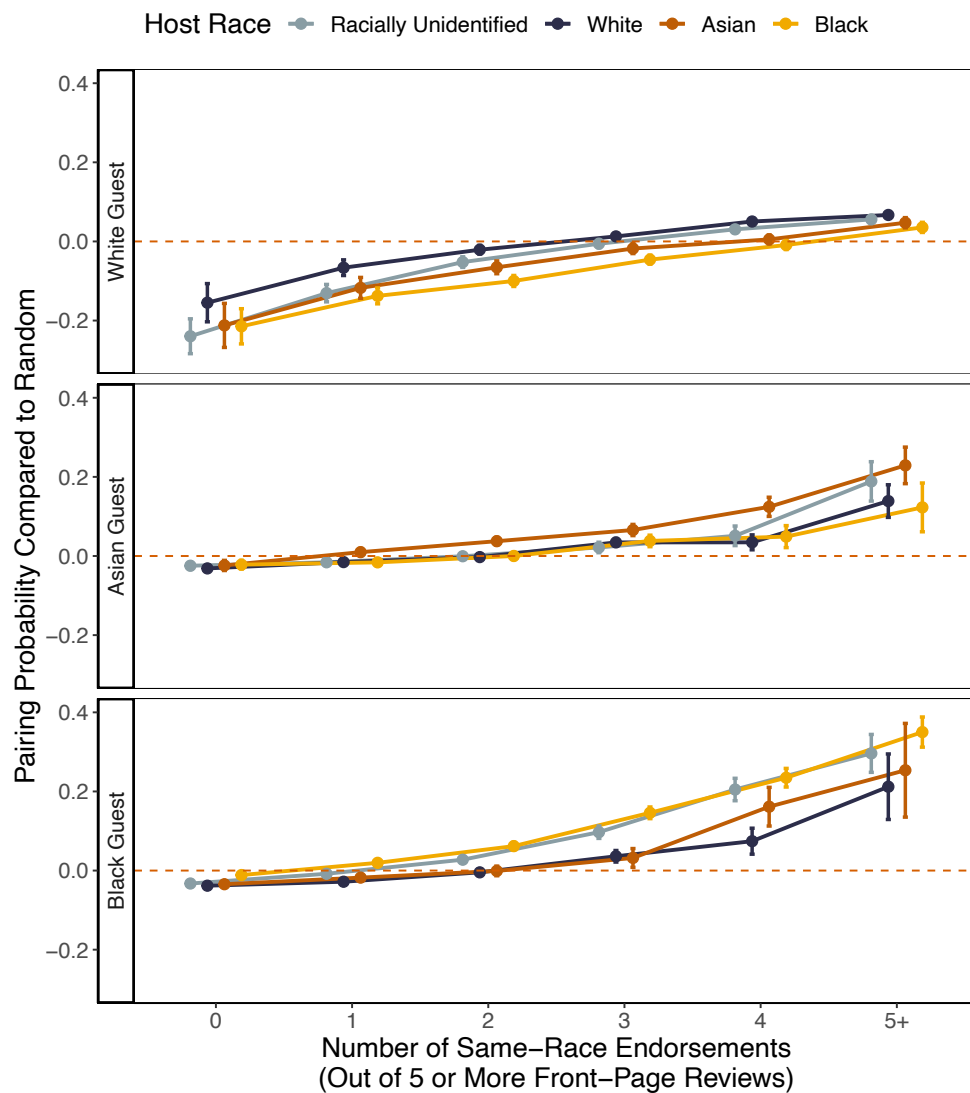
Fig. S1. Response to same-race endorsements of same- and other-race hosts by race of guests and hosts.
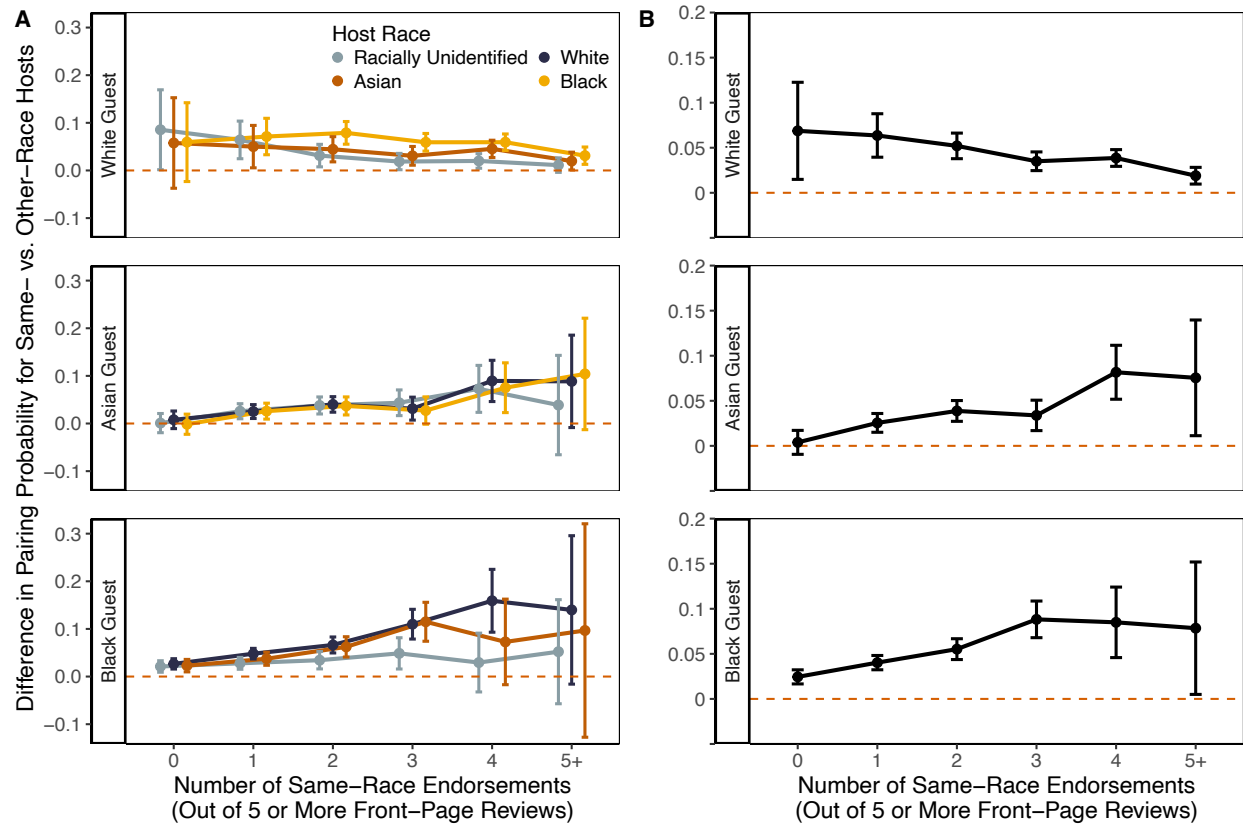
Fig. S2. Response to same-race endorsements of same- vs. other-race hosts. The y-axis is the difference in the pairing probability between same- vs. other-race hosts. The x-axis is the number of same-race endorsements on the host's front page. Panel **A** reports results for all combinations of the race of guest and host. Pairing probabilities for hosts of the same race are not shown since these are the baselines for measuring the difference with the pairing probabilities for other-race hosts. Panel **B** reports the average over all guest-host combinations (as a summary of Panel **A**).
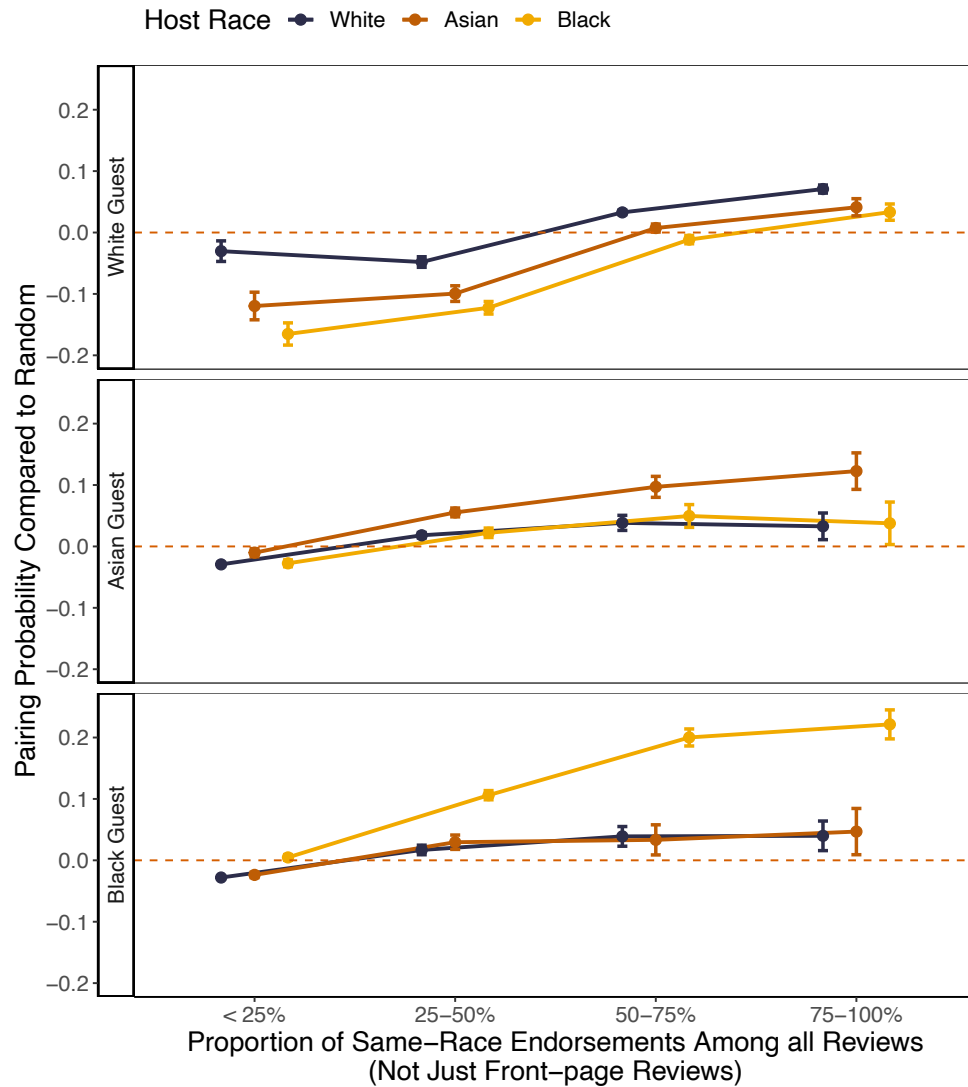
Fig. S3. Response to same-race endorsements of same- and other-race hosts. The endorsements are based on all reviews, not just the six front-page reviews.
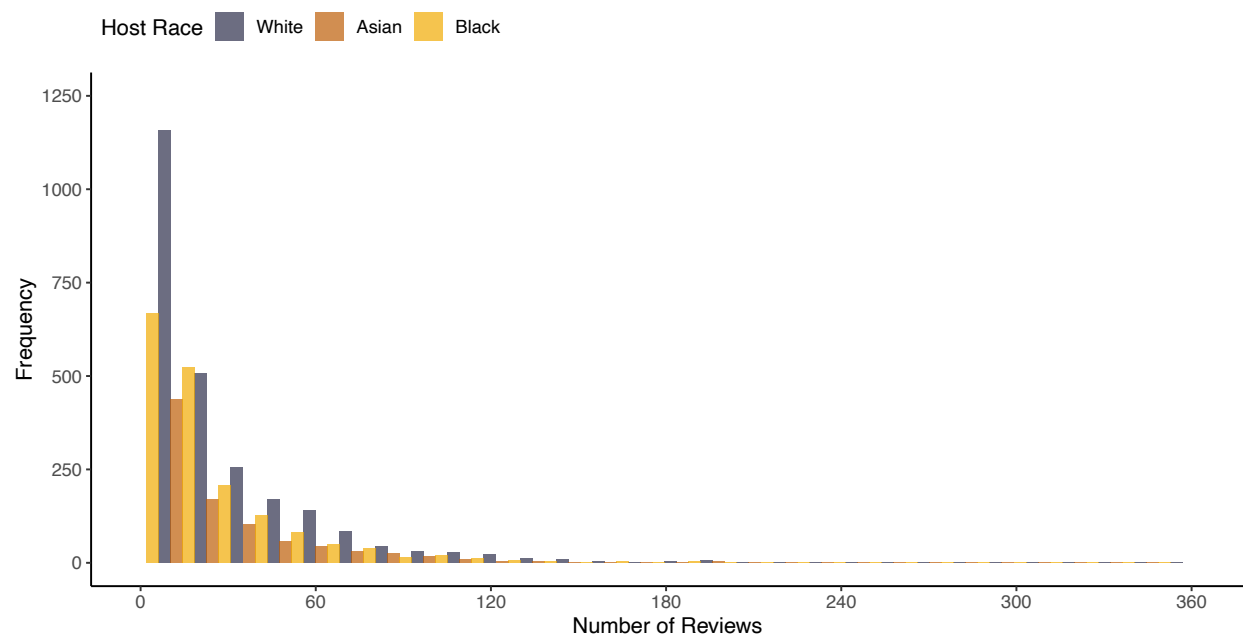
Fig. S4. The distribution of review count broken down by host race.
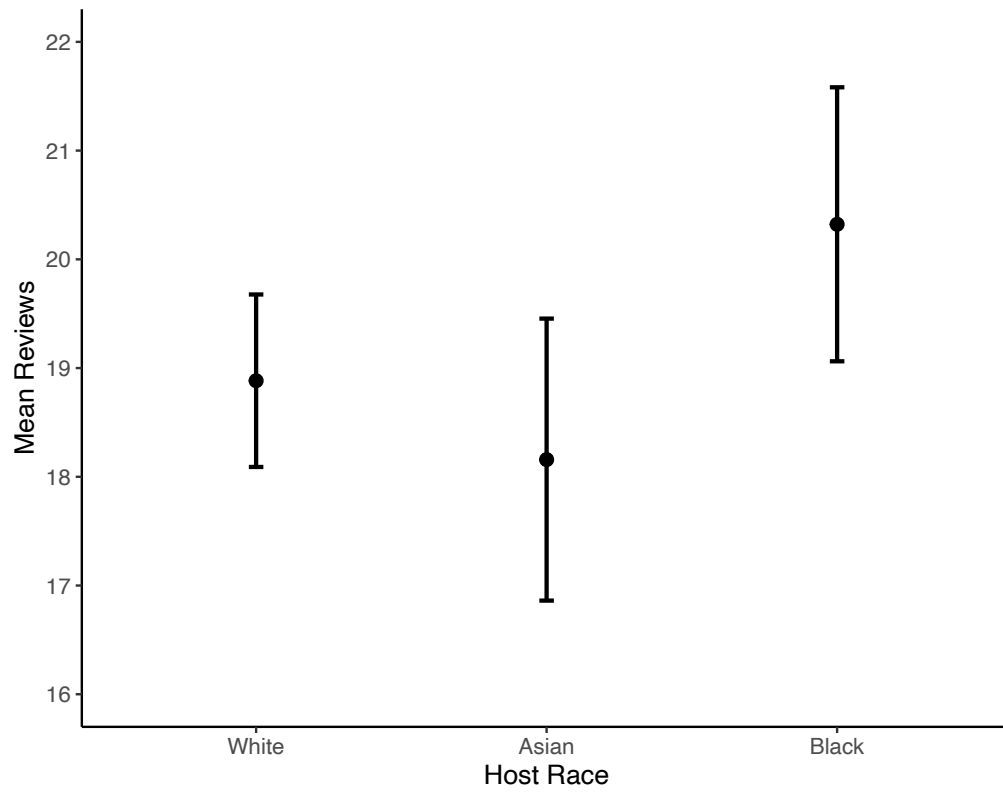
Fig. S5. Mean reviews by host race (among hosts with one or more reviews).
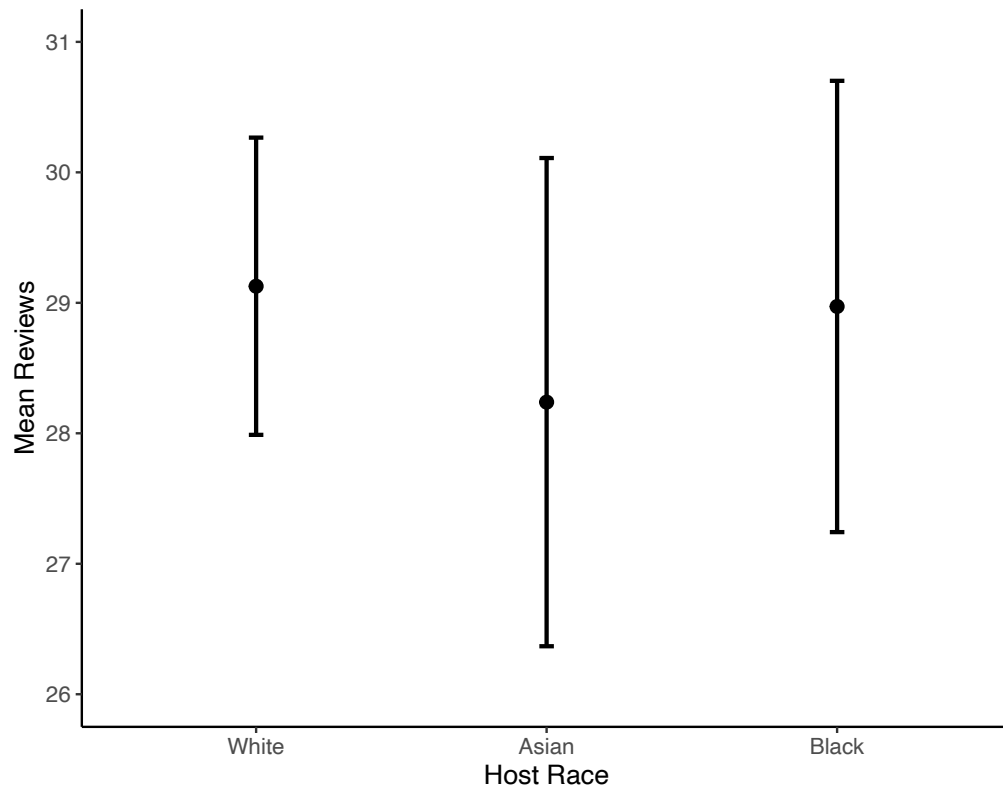
Fig. S6. Mean reviews by host race among hosts with six or more reviews.
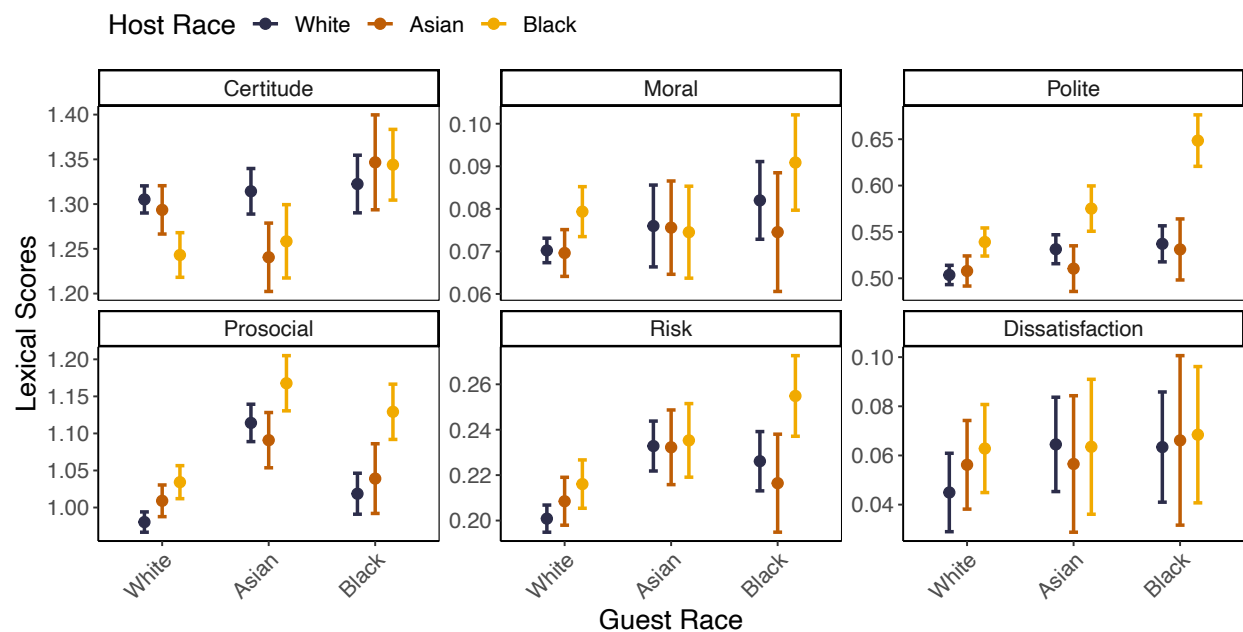
Fig. S7. Lexical scores for six attributes of review content by race of guest and host. All reviews were combined for each of the nine combinations of race of guest and host. Scores for each attribute were calculated as the percent of words in the combined reviews that match words in the LIWC dictionary for that attribute.
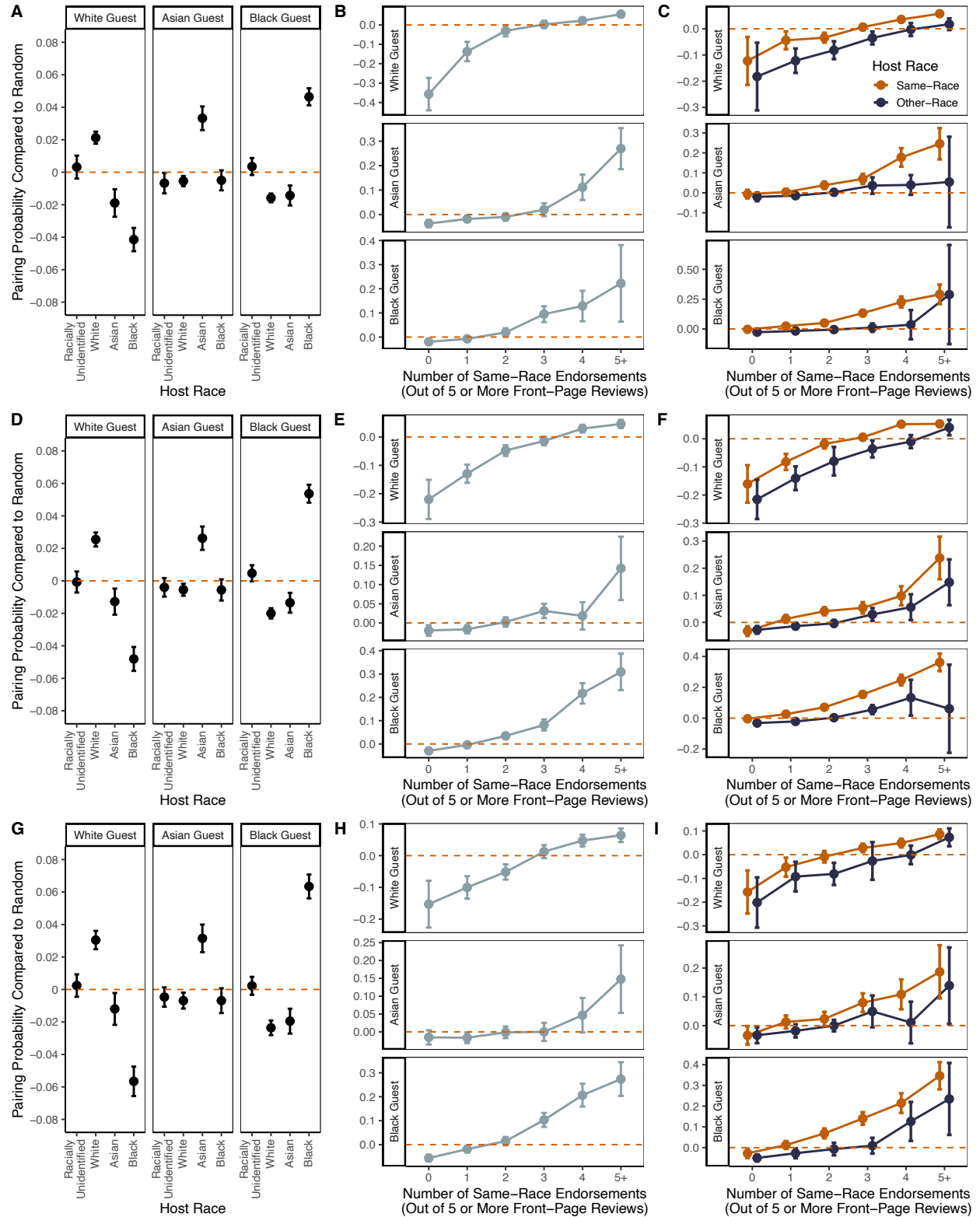
Fig. S8. The replication of Figs. 1-3 in the main text with bookings made before (A-C) and after (D-I) the start of 2017. The y-axis in all panels is the guest-host pairing probability relative to chance. Panels **D-F** are based on listings first appeared before 2017 while **G-I** are based on listing first appeared during or after 2017.
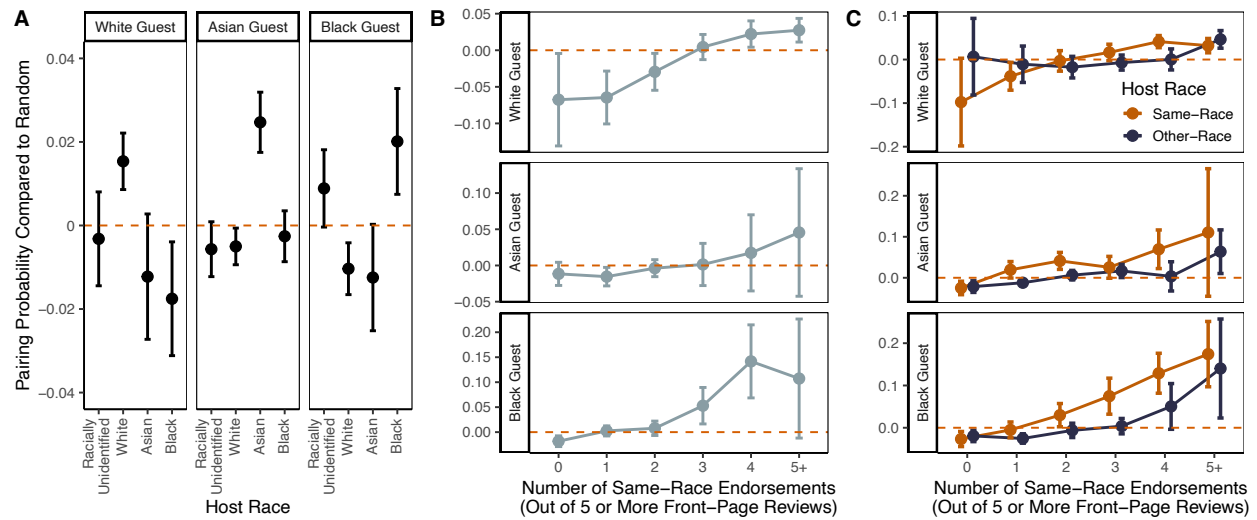
Fig. S9. Replication of Figs. 1-3 in the main text using location-specific pairing probabilities. The y-axis in all panels is the guest-host pairing probability relative to chance. The probability is the weighted average of location-specific pairing probabilities across 42 neighborhood locations (weighted by neighborhood population size). All bookings were made in 2018 (the final seven months of the data collection).

Table S1. Amenity scores (0 to 100 scale) derived from 8,949 pairwise comparisons of amenities using a convenience sample ($N = 64$).

| Amenity	Scores	Amenity	Scores
Patio or balcony	73.72	Building staff	49.20
Private bathroom	72.28	Oven	49.07
Private living room	71.26	Paid parking on premises	48.84
Pool	70.51	Carbon monoxide detector	48.68
Beachfront	64.52	Lockbox	48.54
Luggage drop-off allowed	64.37	Self check-in	48.26
Lake access	63.44	Hot water	48.24
Waterfront	63.10	Microwave	47.93
Private entrance	62.70	EV charger	47.64
Gym	62.50	Safety card	47.37
Internet	62.35	Bathtub	47.31
Balcony	61.22	Bed linens	47.13
Air purifier	60.47	Wide hallway clearance	46.91
Breakfast	60.00	Free street parking	46.54
Indoor fireplace	59.78	Washer	46.24
Garden or backyard	58.94	Coffee maker	46.07
Elevator	58.48	Essentials	45.95
Suitable for events	58.28	Bathtub with bath chair	45.73
Laptop friendly workspace	57.53	Kitchenette	45.55
Hot tub	57.41	First aid kit	45.28
Free parking on premises	57.22	Cooking basics	44.71
24-hour check-in	56.97	Dishes and silverware	44.57
Beach essentials	56.96	Hair dryer	44.57
Wi-Fi	56.68	Wide entryway	43.78
Doorman	56.28	Family/kid friendly	43.64
Heating	56.28	Fireplace guards	43.48
Air conditioning	56.02	Single level home	43.48
Kitchen	55.23	Iron	43.33
Extra pillows and blankets	55.15	Smoke detector	43.07
Room-darkening shades	55.03	Keypad	43.03
Refrigerator	54.44	Paid parking-off premises	43.02
Dishwasher	54.40	Loudspeaker	41.88
BBQ grill	53.42	Hangers	41.57
Electric profiling bed	52.57	Pack 'n play/travel crib	41.52
Lock on bedroom door	52.55	Sitter recommendations	41.50
Handheld shower head	52.33	Fire extinguisher	41.06
Ski in/out	52.30	Wide doorway	41.04
Pocket Wi-Fi	51.93	Shower chair	40.10
TV	51.91	Window guards	39.76
Toilet	51.81	High chair	38.25
Well-lit path to entrance	51.67	Hot water kettle	37.42
Stove	50.94	Children's dinnerware	37.28
Ethernet connection	50.86	Firm mattress	37.23
Host greets you	50.82	Outlet covers	36.97
Dryer	50.56	Shampoo	36.84
Pets allowed	50.28	Baby bath	36.31
Cable tv	50.26	Table corner guards	35.15
Changing table	50.00	Baby monitor	34.51
Game console	49.70	Crib	34.32
Buzzer/wireless intercom	49.69	Smoking allowed	34.20
Smart lock	49.47	Children's books and toys	34.01
Stair gates	49.44		

Table S2. Listing attributes broken down by host race.

	Host Race		
	Asian	Black	White
Average price	133.27	113.60	147.89
	(6.47)	(2.14)	(2.41)
Guest rating	93.33	93.43	94.49
	(0.18)	(0.16)	(0.09)
Amenity score	955.28	1006.47	979.76
	(9.75)	(9.99)	(6.08)

Std. error in parentheses

Table S3. Pairing probability with White, Black, and Asian hosts, broken down by guest race, with and without controls for listing attributes.

Host race	Guest race	Pairing probability			
		Without listing attributes		With listing attributes	
		Estimate	SE	Estimate	SE
White	White	0.625	0.002	0.682	0.006
Black	White	0.551	0.003	0.653	0.007
Asian	White	0.581	0.003	0.655	0.007
White	Black	0.143	0.001	0.177	0.004
Black	Black	0.217	0.002	0.236	0.006
Asian	Black	0.150	0.002	0.181	0.005
White	Asian	0.232	0.001	0.141	0.004
Black	Asian	0.233	0.002	0.112	0.004
Asian	Asian	0.269	0.003	0.164	0.005

Note: $p < 0.001$ for all estimates

Table S4. Association between same-race endorsements and pairing probability with White, Black, and Asian hosts, broken down by guest race and front page vs. all reviews, with and without controls for listing attributes.

| | | | Endorsement coefficient | | | |
| | | | Without listing attributes | | With listing attributes | |
	Host race	Guest race	Estimate	SE	Estimate	SE
Front page reviews	White	White	0.029	0.002	0.014	0.002
	Black	White	0.042	0.002	0.040	0.002
	Asian	White	0.037	0.002	0.031	0.003
	White	Black	0.021	0.001	0.018	0.002
	Black	Black	0.048	0.002	0.058	0.002
	Asian	Black	0.024	0.002	0.022	0.003
	White	Asian	0.020	0.002	0.011	0.001
	Black	Asian	0.020	0.002	0.008	0.002
	Asian	Asian	0.034	0.002	0.032	0.002
All reviews	White	White	0.363	0.014	0.174	0.017
	Black	White	0.487	0.019	0.462	0.022
	Asian	White	0.432	0.022	0.349	0.026
	White	Black	0.259	0.014	0.214	0.018
	Black	Black	0.443	0.012	0.532	0.019
	Asian	Black	0.292	0.023	0.276	0.030
	White	Asian	0.269	0.014	0.155	0.013
	Black	Asian	0.277	0.022	0.158	0.015
	Asian	Asian	0.395	0.021	0.377	0.020

Note: $p < 0.001$ for all estimates

Table S5. Pairing probability with White, Black, and Asian hosts, for both instant and requested bookings, broken down by guest race, with and without controls for listing attributes.

| | | **Pairing probability** | | | |
| | | Without listing attributes | | With listing attributes | |
Host race	Guest race	Estimate	SE	Estimate	SE
White	White	0.629	0.001	0.696	0.005
Black	White	0.558	0.002	0.673	0.005
Asian	White	0.592	0.002	0.677	0.005
White	Black	0.143	0.001	0.184	0.003
Black	Black	0.216	0.001	0.234	0.004
Asian	Black	0.147	0.001	0.185	0.004
White	Asian	0.228	0.001	0.120	0.003
Black	Asian	0.226	0.001	0.092	0.002
Asian	Asian	0.261	0.002	0.138	0.003

Note: $p < 0.001$ for all estimates

Table S6. Association between same-race endorsements and pairing probability with White, Black, and Asian hosts, for both instant and requested bookings, broken down by guest race and front page vs. all reviews.

| | Host race | Guest race | Endorsement coefficient | | | |
| | | | Without listing attributes | | With listing attributes | |
			Estimate	SE	Estimate	SE
Front page reviews	White	White	0.027	0.001	0.013	0.001
	Black	White	0.044	0.001	0.041	0.002
	Asian	White	0.035	0.002	0.029	0.002
	White	Black	0.019	0.001	0.014	0.001
	Black	Black	0.047	0.001	0.055	0.002
	Asian	Black	0.025	0.002	0.025	0.002
	White	Asian	0.018	0.001	0.010	0.001
	Black	Asian	0.019	0.002	0.008	0.001
	Asian	Asian	0.031	0.002	0.025	0.001
All reviews	White	White	0.354	0.01	0.157	0.012
	Black	White	0.492	0.013	0.473	0.015
	Asian	White	0.424	0.015	0.337	0.018
	White	Black	0.248	0.009	0.174	0.013
	Black	Black	0.443	0.008	0.528	0.014
	Asian	Black	0.296	0.015	0.299	0.021
	White	Asian	0.242	0.01	0.132	0.008
	Black	Asian	0.264	0.015	0.133	0.009
	Asian	Asian	0.365	0.014	0.295	0.012

Note: $p < 0.001$ for all estimates

Table S7. Pairing probability with White, Black, and Asian hosts, for requested bookings, broken down by guest race, with and without controls for listing attributes.

		Pairing probability			
		Without listing attributes		With listing attributes	
Host race	Guest race	Estimate	SE	Estimate	SE
White	White	0.633	0.002	0.711	0.007
Black	White	0.565	0.002	0.694	0.008
Asian	White	0.600	0.003	0.701	0.007
White	Black	0.143	0.001	0.191	0.006
Black	Black	0.215	0.002	0.232	0.007
Asian	Black	0.145	0.002	0.187	0.006
White	Asian	0.225	0.001	0.098	0.003
Black	Asian	0.220	0.002	0.074	0.003
Asian	Asian	0.255	0.002	0.112	0.004

Note: $p < 0.001$ for all estimates

Table S8. Association between same-race endorsements and pairing probability with White, Black, and Asian hosts, for requested bookings, broken down by guest race and front page vs. all reviews, with and without controls for listing attributes.

	Host race	Guest race	Endorsement coefficient			
			Without listing attributes		With listing attributes	
			Estimate	SE	Estimate	SE
Front page reviews	White	White	0.025	0.001	0.011	0.002
	Black	White	0.045	0.002	0.042	0.002
	Asian	White	0.033	0.002	0.027	0.003
	White	Black	0.017	0.001	0.010	0.002
	Black	Black	0.046	0.002	0.052	0.002
	Asian	Black	0.026	0.002	0.028	0.003
	White	Asian	0.016	0.001	0.008	0.001
	Black	Asian	0.019	0.002	0.007	0.001
	Asian	Asian	0.028	0.002	0.018	0.002
All reviews	White	White	0.345	0.013	0.141	0.016
	Black	White	0.493	0.017	0.477	0.020
	Asian	White	0.412	0.020	0.319	0.024
	White	Black	0.239	0.013	0.129	0.019
	Black	Black	0.444	0.011	0.517	0.020
	Asian	Black	0.299	0.020	0.324	0.031
	White	Asian	0.214	0.014	0.108	0.010
	Black	Asian	0.246	0.022	0.107	0.011
	Asian	Asian	0.336	0.019	0.214	0.014

Note: $p < 0.001$ for all estimates

REFERENCES AND NOTES

1. K. Levy, S. Barocas, Designing against discrimination in online markets. *Berkeley Technol. Law J.* **32**, 1183 (2017).

2. B. Edelman, M. Luca, D. Svirsky, Racial discrimination in the sharing economy: Evidence from a field experiment. *Am. Econ. J. Appl. Econ.* **9**, 1–22 (2017).

3. J. Mejia, C. Parker, When transparency fails: Bias and financial incentives in ridesharing platforms. *Manage. Sci.* **67**, 166–184 (2021).

4. J. Santana, P. Parigi, Risk aversion and engagement in the sharing economy. *Games* **6**, 560–573 (2015).

5. J. D. Tjaden, C. Schwemmer, M. Khadjavi, Ride with me—Ethnic discrimination, social markets, and the sharing economy. *Eur. Sociol. Rev.* **34**, 418–432 (2018).

6. J. Kleinberg, J. Ludwig, S. Mullainathan, C. R. Sunstein, Discrimination in the age of algorithms. *J. Leg. Anal.* **10**, 113–174 (2018).

7. J. Kleinberg, J. Ludwig, S. Mullainathan, C. R. Sunstein, Algorithms as discrimination detectors. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 30096–30100 (2020).

8. Z. Obermeyer, B. Powers, C. Vogeli, S. Mullainathan, Dissecting racial bias in an algorithm used to manage the health of populations. *Science* **366**, 447–453 (2019).

9. P. Törnberg, How sharing is the "sharing economy"? Evidence from 97 Airbnb markets. *PLOS ONE* **17**, e0266998 (2022).

10. Y. Ge, C. R. Knittel, D. MacKenzie, S. Zoepf, Racial discrimination in transportation network companies. *J. Public Econ.* **190**, 104205 (2020).

11. A. Marchenko, The impact of host race and gender on prices on Airbnb. *J. Hous. Econ.* **46**, 101635 (2019).

12. B. G. Edelman, M. Luca, "Digital discrimination: The case of Airbnb.com," *Harvard Bus. Sch. NOM Unit Work. Pap.* (no. 14-054) (2014).

13. B. Abrahao, P. Parigi, A. Gupta, K. S. Cook, Reputation offsets trust judgments based on social biases among Airbnb users. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 9848–9853 (2017).

14. R. Cui, J. Li, D. J. Zhang, Reducing discrimination with reviews in the sharing economy: Evidence from field experiments on airbnb. *Manage. Sci.* **66**, 1071–1094 (2020).

15. P. Racherla, M. Mandviwalla, D. J. Connolly, Factors affecting consumers' trust in online product reviews. *J. Consum. Behav.* **11**, 94–104 (2012).

16. M. Zhao, J. Xie, Effects of social and temporal distance on consumers' responses to peer recommendations. *J. Market. Res.* **48**, 486–496 (2011).

17. K. Munger, Tweetment effects on the tweeted: Experimentally reducing racist harassment. *Polit. Behav.* **39**, 629–649 (2017).

18. J. C. Turner, in *Rediscovering Social Identity*, T. Postmes, N. R. Branscombe, Eds. (Psychology Press, 2010), pp. 243–272.

19. H. Tajfel, J. C. Turner, W. G. Austin, S. Worchel, An integrative theory of intergroup conflict. *Organ. identity A Read.* **56**, 9780203505984--16 (1979).

20. H. Tajfel, Social identity and intergroup behaviour. *Soc. Sci. Inf.* **13**, 65–93 (1974).

21. C. J. Hutto, E. Gilbert, Vader: A parsimonious rule-based model for sentiment analysis of social media text, in *Proceedings of the 8th International AAAI Conference on Weblogs and Social Media* (vol. 8, 2014), pp. 216–225.

22. D. A. Strauss, discriminatory intent and the taming of brown. *Univ. Chicago Law Rev.* **56**, 935 (1989).

23. J. De Houwer, Implicit bias is behavior: A functional-cognitive perspective on implicit bias. *Perspect. Psychol. Sci.* **14**, 835–840 (2019).

24. L. Cosmides, J. Tooby, R. Kurzban, Perceptions of race. *Trends Cogn. Sci.* **7**, 173–179 (2003).

25. M. R. Banaji, A. G. Greenwald, *Blindspot: Hidden Biases of Good People* (Delacorte Press, 2013).

26. B. Hofstra, R. Corten, F. van Tubergen, N. B. Ellison, Sources of segregation in social networks: A novel approach using Facebook. *Am. Sociol. Rev.* **82**, 625–656 (2017).

27. G. Zervas, D. Proserpio, J. W. Byers, A first look at online reputation on Airbnb, where every stay is above average. *Mark. Lett.* **32**, 1–16 (2021).

28. J. Bridges, C. Vásquez, If nearly all Airbnb reviews are positive, does that make them meaningless? *Curr. Issues Tour.* **21**, 2057–2075 (2018).

29. A. C. Morales, Giving firms an "E" for effort: Consumer responses to high-effort firms. *J. Consum. Res.* **31**, 806–812 (2005).

30. A. Fradkin, E. Grewal, D. Holtz, Reciprocity and unveiling in two-sided reputation systems: Evidence from an experiment on Airbnb. *Mark. Sci.* **40**, 1013–1029 (2021).

31. G. M. Castex, Providing services to hispanic/latino populations: Profiles in diversity. *Soc. Work* **39**, 288–296 (1994).

32. N. Garg, L. Schiebinger, D. Jurafsky, J. Zou, Word embeddings quantify 100 years of gender and ethnic stereotypes. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E3635–E3644 (2018).

33. B. Jaeger, W. W. A. Sleegers, A. M. Evans, Automated classification of demographics from face images: A tutorial and validation. *Soc. Personal. Psychol. Compass* **14**, e12520 (2020).

34. B. Pan, H. Hembrooke, T. Joachims, L. Lorigo, G. Gay, L. Granka, In Google we trust: Users' decisions on rank, position, and relevance. *J. Comput. Commun.* **12**, 801–823 (2007).

35. T. Joachims, L. Granka, B. Pan, H. Hembrooke, F. Radlinski, G. Gay, Evaluating the accuracy of implicit feedback from clicks and query reformulations in Web search. *ACM Trans. Inf. Syst.* **25**, 7 (2007).

36. M. J. Salganik, K. E. C. Levy, Wiki surveys: Open and quantifiable social data collection. *PLOS ONE* **10**, e0123483 (2015).