

Detecting and Mitigating Discrimination in Online Platforms: Lessons from Airbnb, Uber, and Others

Michael Luca and Dan Svirsky

The rise of online marketplaces raises the potential for markets that are both more efficient and less biased. Early research pointed to the arms-length nature and relative anonymity of online transactions as factors that might lead to less discrimination in online transactions. However, the extent to which this promise is realized depends on the design choices platforms make. As a growing share of markets and transactions have moved online, marketplaces have evolved and platform designers have sought ever newer ways to encourage trust between strangers. Platforms have made different design choices over time, across industries – and even within industries. These choices shape both the efficiency and inclusivity of markets.

When trust-building mechanisms facilitate discrimination ✕ Breaking with design choices made by many earlier online marketplaces, platforms such as Airbnb made names and pictures of market participants salient before deciding whether or not to transact. While this was presumably intended to encourage trust and ease commercial exchange among strangers, it also opened doors for discrimination in online marketplaces. Research has now documented racial or ethnic discrimination in a variety of areas online, from labor markets to credit applications to housing. It is enabled by two notable features. First, markers of race or ethnicity – most obviously photographs, but also subtler indicators, such as names – can trigger conscious or unconscious discrimination. The second feature is increased discretion on the part of sellers over which buyers they transact with. Both are choices made by platform designers.

KEYWORDS

Market Design, Platform Design, Discrimination, Field Experiments

THE AUTHORS

Michael Luca

Lee J. Styslinger III Associate Professor of Business Administration, Harvard Business School, Boston, MA, USA, mluca@hbs.edu

Dan Svirsky

Data Scientist, Uber Technologies, Inc, Boston, MA, USA, dsvirsky@uber.com





Platform managers should develop more awareness of discrimination and proactively address the problem.





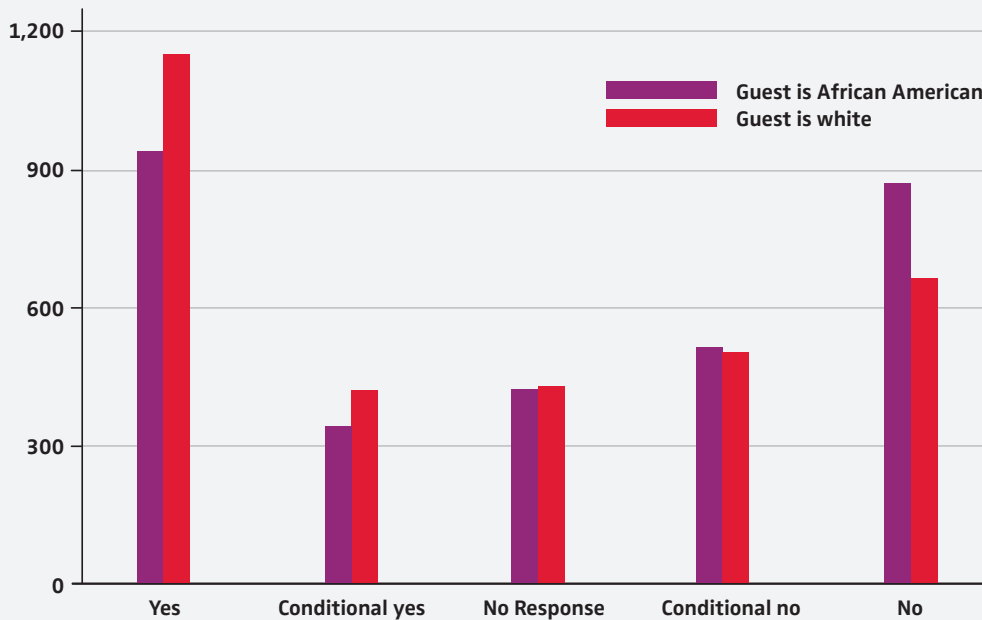
BOX 1

Investigating discrimination on Airbnb

In 2015, together with a colleague, we investigated racial discrimination on Airbnb. In a study focused on the US cities of Baltimore, Dallas, Los Angeles, St. Louis, and Washington, we constructed 20 user profiles and sent rental requests to roughly 6,400 hosts. The profiles and requests were identical except for one detail—the user’s name. Half the profiles had names that, according to birth records, are common among whites, while half had names common among African Americans.

In contrast with Airbnb’s claims, our work found evidence of extensive discrimination on the platform. Requests with African American-sounding names were approximately 16% less likely than those with white-sounding names to be accepted. Most of the hosts who declined requests from black-sounding profiles had never hosted a black guest – suggesting that some hosts are especially inclined to discriminate on the basis of race.

FIGURE 1 > Host response by race (as indicated by guest name)



from Edelman, B., Luca, M., & Svirsky, D. (2017)

While Airbnb did not cooperate on the study, they started to address the problem in response to its results and growing criticism from users and regulators. We also provided them with suggestions about changes they could make to reduce discrimination. Airbnb commissioned a task force to evaluate our suggestions and other ways to reduce discrimination. Airbnb eventually incorporated a set of design changes. One example of a potentially promising change is the recently expanded “instant book” feature, which enables qualified guests to book instantly and before hosts can have a look at their profile information. Besides preventing discrimination, this feature also makes booking simpler and more convenient. In response to our work, Airbnb committed to expanding the use of instant booking – and now has millions of listings around the globe that use this feature. However, to date, Airbnb has not revealed the extent to which discrimination persists on the platform – or the impact that their changes have had.

A managerial toolkit for reducing discrimination on platforms

✗ Even within an industry, platforms often differ in their design features, which can lead to different levels of discrimination. For example, the main search-results page of the vacation rental marketplace HomeAway displays photos only of the property for rent and withholds host photos until a later page or not at all, whereas Airbnb historically included host photos on its main search-results page. In response to our research, Airbnb has now changed that policy and shows host and guest photos only later in the process. This exemplifies the types of steps platforms can take to reduce discrimination. Drawing on our prior research as well as our experience with companies, we explore steps to mitigate discrimination (see Figure 2).

> Build awareness for potential discrimination on platforms

✗ Platforms should develop an understanding of the ways in which their design choices and algorithms can affect the amount of discrimination in a marketplace. By increasing awareness of this, managers can be proactive about investigating and tackling the problem. For example, Uber's central policy team created a Fairness Working Group to explore discrimination issues. Part of the group's value comes from its cross-functional nature – it brings together economists, data scientists, lawyers, and product managers from around the company to think through ways to address fairness challenges. Especially for large organizations, it can be useful to have a group dedicated to monitoring new projects solely for discrimination risks.

> Measure discrimination on the platform

✗ Currently, many platforms do not know the racial, ethnic, or gender composition of their transaction participants, and it's hard to address an issue if you aren't measuring it. A regular report or audit on the discrimination-endangered users, along with measures of each group's success on the platform, is a critical step toward revealing and confronting any problems. Following our research, Airbnb began to measure discrimination on the platform. The company now has economists and data scientists working on this topic. Similarly, Uber's working group helps to quantify discrimination on an ongoing basis.



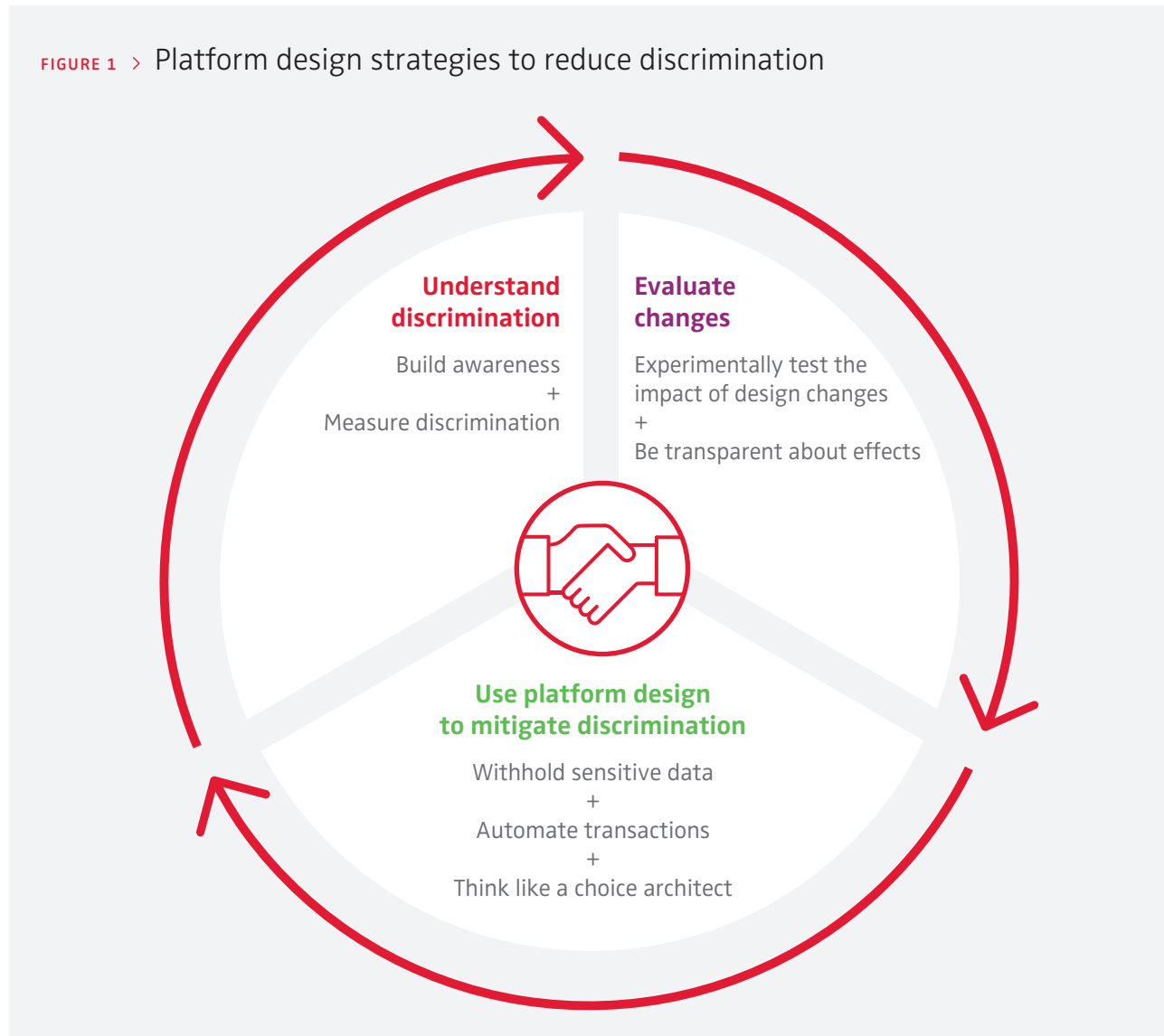
> Withhold sensitive data

✗ In many cases, a simple but effective change a platform can make is to withhold potentially sensitive user information, such as race and gender, until after a transaction has been agreed to. Some platforms, including Amazon and eBay, already do this. To see the impact of showing markers of race before a transaction occurs, consider eBay – which does not make race salient. In a recent study users hold baseball cards in their hand and post that as a photo in the listing – they found that even that subtle prompt led to racial discrimination in baseball card sales on eBay.

> Consider automating transactions, but be aware of algorithmic bias

✗ Automation and algorithms can be a useful tool in reducing bias. For example, return to the example of Airbnb's instant booking described in Box 1. This feature eliminates the step in which hosts look at the guest's name and picture and decide whether to approve or reject the guest. Airbnb has now greatly increased the number of users who use instant booking.

FIGURE 1 > Platform design strategies to reduce discrimination



A growing body of literature has begun to explore ways to debias algorithms as well. Changing humans' preferences or attitudes is often difficult, but changing the inputs of an algorithm, or its objectives, can at times be more straightforward. For example, LinkedIn redesigned its Recruiter tool – a search platform to find job candidates – to ensure that the gender breakdown of search results matches the gender breakdown for that occupation as a whole. If 30% of data scientists are women, then a recruiter searching for data scientists would see search results where 30% of the users are female candidates. This example highlights the ways in which algorithms can impact the equity of a market. It also highlights the nuance and managerial judgment involved in designing

an algorithm, as the target set for the tool is only one of multiple ways to think about fairness.

- > **Think like a choice architect** ✕ Further, platforms can use principles from choice architecture to reduce discrimination. For example, in a variety of contexts, research has documented the tendency for people to use whatever option is set up as the default. To see why this matters, return to the example of the instant book feature at Airbnb, which is an opt-in feature: landlords must sign up for it. If Airbnb switched its default to instant book, requiring hosts to actively opt out of it, the company could reduce the scope for discrimination. Airbnb has experimented with approaches such as this over time.



Marketplaces could present antidiscrimination policies at a more relevant moment – during the actual transaction process.



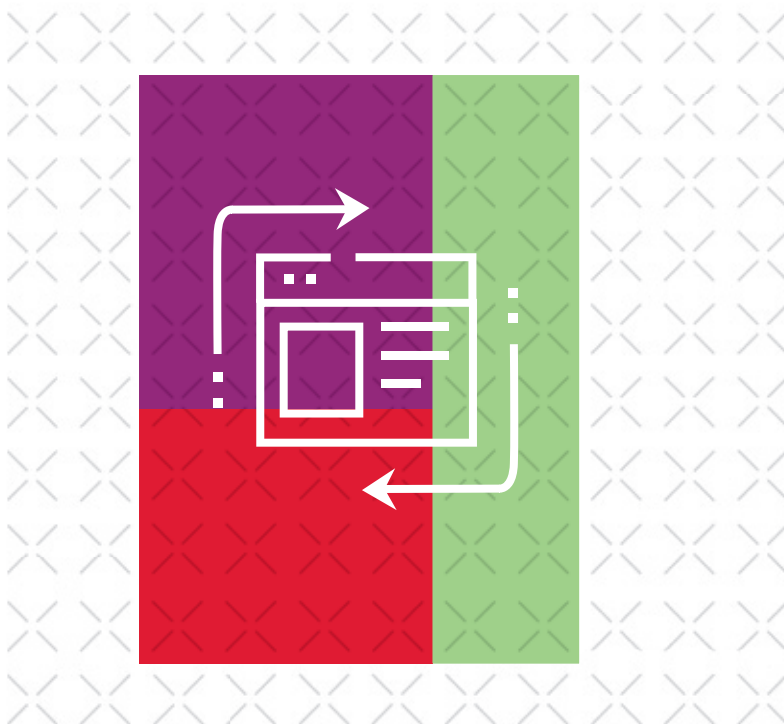
As a second example, consider discrimination policies. Most platforms have policies prohibiting discrimination, but they're buried in fine print and mostly just ticked off once without proper reading. Marketplaces could present antidiscrimination policies at a more relevant moment – during the actual transaction process. Some people would still violate the policies, of course, but others might be glad for the hint because they might not be aware of the problem otherwise.

- > **Experimentally test the impact of platform design changes on discrimination levels** ✕ Once platforms have a way to measure discrimination, they can incorporate this into their experimental testing. By incorporating such metrics, platforms can better understand the disparate impact of different designs and features. Airbnb has, for example, conducted an experiment in withholding host photos from its main search results page to explore the effects on booking outcomes. Following our research, they

now have a team that explores issues such as this. LinkedIn tested the effects of the changes to its recruiting search, finding that the changes did not impact the success rate of recruiters' outreach messages.

- > **Be transparent** ✕ Platforms should aim for transparency and work with a broad set of stakeholders to identify and solve issues of discrimination. This will help to facilitate discussions among platform managers and designers before it becomes a crisis. It will also allow the progress of measures to be evaluated over time. This is particularly important given the uncertain impact of changes platforms are making.

The rise of online marketplaces has dramatically changed the nature of many economic transactions. Our research has shown the profound impact that design choices can have on outcomes. By leveraging insights from platform design research, companies have the opportunity to create markets that are both efficient and inclusive. ✕



FURTHER READING

Ayres, I.; Banaji, M.; & Jolls, C. (2015): "Race effects on eBay", *The RAND Journal of Economics*, Vol. 46, 891-917. doi: 10.1111/1756-2171.12115

Edelman, B.; Luca, M.; & Svirsky, D. (2017): "Racial discrimination in the sharing economy: Evidence from a field experiment", *American Economic Journal: Applied Economics*, Vol. 9(2), 1-22. <https://doi.org/10.1257/app.20160213>

Fisman, R.; & Luca, M. (2016): Fixing Discrimination in Online Marketplaces", <https://hbr.org/2016/12/fixing-discrimination-in-online-marketplaces>

Luca, M. & Bazerman, M. (2020): *The Power of Experiments: Decision-Making in a Data-Driven World*, <https://mitpress.mit.edu/books/power-experiments>

Morton, F. S.; Zettelmeyer, F.; & Silva-Risso, J. (2003): "Consumer Information and Discrimination: Does the Internet Affect the Pricing of New Cars to Women and Minorities?", *Quantitative Marketing and Economics*, Vol. 1, 65-92. <https://doi.org/10.1023/A:1023529910567>