Ray Fisman is the Slater Family Professor in Behavioral Economics at Boston University and a coauthor of *The Inner Lives of Markets: How People Shape Them—and They Shape Us.* **Michael Luca** is an assistant professor at Harvard Business School and a visiting assistant professor at Stanford University.

Airbnb, Uber, and others are facing the unintended consequences of their platforms' design choices.

FIXING DISCRIMINATION IN ONLINE MARKETPLACES

STEPHEN CHAN

IN THE LATE 1980S, LAW PROFESSORS IAN AYRES AND PETER SIEGELMAN SET OUT TO LEARN WHETHER BLACKS AND WOMEN GOT THE SAME DEALS AS WHITE MEN WHEN BUYING A NEW CAR. THEY TRAINED 38 PEOPLE—SOME WHITE AND SOME BLACK, SOME MALE AND SOME FEMALE—TO NEGOTIATE A PURCHASE USING A FIXED SCRIPT, AND UNCOVERED DISTURBING DIFFERENCES: ACROSS 153 DEALERSHIPS, BLACK AND FEMALE BUYERS PAID MORE FOR THE SAME CARS THAN WHITE MEN DID, WITH BLACK WOMEN PAYING THE MOST—ON AVERAGE, NEARLY $900 MORE THAN WHITE MEN. ALTHOUGH THE FINDINGS WEREN'T A SURPRISE TO MOST PEOPLE, LEAST OF ALL TO BLACKS AND WOMEN, THEY WERE A COMPELLING DEMONSTRATION OF JUST HOW DISCRIMINATORY MARKETS CAN BE.

Fast-forward a dozen years to the early days of internet commerce. Entrepreneurs were experimenting with web-based sales of everything, including automobiles. Economists Fiona Scott Morton, Florian Zettelmeyer, and Jorge Silva-Risso analyzed this new mode of selling cars and found that it did away with the racial and gender discrimination that, they also found, persisted in off-line automobile sales.

Indeed, the first generation of online marketplaces, including eBay, Amazon, and Priceline, made it hard for sellers to discriminate. Transactions were conducted with relative anonymity. A user could negotiate a purchase without providing any identifying information until the seller had agreed to the deal. As a *New Yorker* cartoon famously put it, "On the Internet, nobody knows you're a dog."

Except that platforms—and now their users—do know whether you're black or white, male or female, human or canine. And the internet has recently been revealed as a source of discrimination, not an end to it: With their identities uncovered, disadvantaged groups face many of the same challenges they have long confronted in the off-line world, sometimes made worse by a lack of regulation, the salience photos give to race and gender, and the fact that would-be discriminators can act without ever personally confronting their victims.

What happened, and what can we do about it?

The Emergence of Digital Discrimination

In the early days of e-commerce, shopping online often required a leap of faith. An eBay seller in Florida might post, say, a Topps baseball card for Nolan Ryan's 1974 season with the California Angels, along with a description of its condition. A collector in Massachusetts could bid on the card sight unseen, on the basis of that description. A card in mint condition might be valued at $60, but a dog-eared one would be worth just a fraction of that. What was to

Idea in Brief

THE PROBLEM
Online marketplaces such as eBay, Uber, and Airbnb have the potential to reduce racial, gender, and other kinds of discrimination that affect transactions in the off-line world. But recent research shows that the opposite has occurred.

THE REASON
Early platforms kept the identities of buyers and sellers relatively anonymous. But the addition of photos, names, and other means of identification to listings has inadvertently encouraged discriminatory behavior.

THE ANSWER
To create markets that are both efficient and inclusive, platform designers need to be mindful of the potential for discrimination and open to experimentation as they make choices about automation, algorithms, and the use of identifying data.

prevent the seller from passing off a well-worn card as pristine? Very little: A study by economists Ginger Jin and Andrew Kato found that in the early 2000s, eBay merchants often misrepresented the quality of sports trading cards.

The problem with early e-commerce was that one side of the market tended to know things the other side didn't—the condition of a baseball card, the reliability and care with which goods would be packaged, and so on. These challenges arise in all markets, but they were particularly severe for online platforms, for two main reasons. First, it's harder to overcome information asymmetries when you can't hold a product in your hand. Second, online sellers were, almost by definition, new to the business, since the business itself had been around for just a few years. There were no established brands, such as Sotheby's and Sears, to assure buyers they wouldn't be cheated.

Over time, buyer reviews and other feedback have allowed e-commerce sellers to build up reputations. But why stop at collecting feedback when so much potentially useful information could be mined from buyers' and sellers' identities? For example, in a 2012 study of peer-to-peer lending by Jefferson Duarte, Stephan Siegel, and Lance Young, subjects rated potential borrowers' trustworthiness after viewing photographs of them. It turned out that people who "look trustworthy" were more likely to have their loan requests granted. More surprisingly, they were also more likely to repay the loans. The implication was that if this type of fine-grained information could help market participants assess a transaction's prospects, it made sense to provide it.

On the websites of services ranging from freelancing to ride sharing to dog walking, many sellers now have discretion over whom they do business with on the basis of looks or even just a name. The availability of such information is platform-specific, with some sites preserving a fair amount of anonymity while others hark back to practices long banned in off-line markets. Similarly, on many sites, including Etsy and CustomMade, potential buyers see not only products but also the names and photos of sellers. Although having details about prospective transaction partners may make people more comfortable, a growing body of evidence shows that it facilitates discrimination.

The short-term-rental marketplace Airbnb is a case in point regarding the emergence of discrimination in online markets and the ways in which design choices influence the extent of it. When a would-be renter searches listings, he sees descriptions and pictures of both the property and the host. And hosts can see the names—and in many instances the pictures—of potential tenants before accepting or rejecting them.

One of us (Mike, working with Benjamin Edelman and Daniel Svirsky) has investigated racial discrimination on Airbnb. In a study focused on the U.S. market, the group 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 blacks.

Requests with black-sounding names were 16% less likely than those with white-sounding names to be accepted. And the discrimination was pervasive, occurring with cheap listings and expensive ones, diverse neighborhoods and homogeneous ones, rooms in the host's own dwelling and separate units rented out by landlords with multiple listings. 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. (In response to this study and to a growing chorus of criticism

from users and regulators, Airbnb commissioned a task force to identify ways to reduce discrimination, which proposed a series of changes in September 2016. We will discuss aspects of the announced policies below.)

Researchers have now documented racial discrimination in a variety of areas online, from labor markets to credit applications to housing. It is enabled by two features: markers of race, most obviously photographs but also subtler indicators, such as names; and discretion on the part of market participants over whom they transact with. As we will discuss in the next section, both are choices made by platform designers.

Another feature of online commerce has at times, also counterintuitively, nurtured rather than suppressed discrimination: the use of algorithms and big data. The search results Google serves up, the books Amazon suggests, and the movies Netflix recommends are all examples of machines' replacing imperfect human judgment about what customers want. It's tempting to assume that eliminating human judgment would eliminate human bias as well. But that's not the case.

In fact, algorithm-generated discrimination occurs in ways that humans would probably avoid. In an eye-opening study, computer science professor Latanya Sweeney sought to understand the role of race in Google ads. She searched for common African-American names—such as Deshawn and, well, Latanya—and recorded the ads that appeared with the results. She then searched for names, such as Geoffrey, that are more common among whites. The searches for black-sounding names were more

likely to generate ads offering to investigate possible arrest records.

Of course, Google didn't set out to show arrest-record ads to people who searched for African-American names. That happened because an algorithm "decided," on the basis of past searches, that someone searching for "Deshawn" is more likely than someone searching for "Geoffrey" to click on an arrest-related ad (and hence generate revenue for Google). That is, the choice was made, if unwittingly, by Google's algorithm designers.

Toward Smarter Market Design

Platforms—even when they're in the same industry—often differ in their design features, which can lead to different levels of vulnerability to discrimination. Take the decision whether and when to post user pictures. Uber does not provide drivers with photos of potential passengers, but its competitor Lyft does. This makes Uber less vulnerable than Lyft to discrimination by drivers. Similarly, 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 (if it shows them at all), whereas Airbnb requires that hosts include photos of themselves on its main search-results page.

Companies also have varying approaches to investigating possible discrimination and taking remedial action. For example, eBay worked with a team of social psychologists to explore whether male sellers get higher prices than female sellers for similar items (they do). More commonly, though, businesses avoid the issue. Although many executives acknowledge that discrimination occurs and express interest in reducing it, we've seen few earnest efforts like eBay's to gauge its extent. So researchers looking to study online discrimination must run their own experiments or scrape decidedly imperfect data from websites. (And we know of cases where company lawyers have gone after such efforts in an attempt to block race-related research.)

Even companies with the best of intentions may not choose the best approach to fighting discrimination, because, to our knowledge, no system exists for thinking through the available design choices and their implications. Our aim in what follows is to offer a framework for companies that want to design and manage a thriving marketplace while minimizing the risk of discrimination.

We don't expect every market designer to make the same decisions. Just as competitors make differing design choices about other situations (for instance, Lyft lets riders tip through its app, but Uber doesn't), they will make differing choices about confronting discrimination; among other reasons, they place differing premiums on avoiding discrimination (although we believe that platforms should hold themselves to a high standard in this regard). Our goal is to help designers fully consider the implications and trade-offs of their design choices.

Below we offer two guiding principles for platforms struggling with this market-design challenge. We then evaluate four design choices that are likely to affect discrimination.

IT'S TEMPTING TO ASSUME THAT ELIMINATING HUMAN JUDGMENT WOULD ELIMINATE HUMAN BIAS AS WELL. BUT ALGORITHM-GENERATED DISCRIMINATION OCCURS IN WAYS THAT HUMANS WOULD PROBABLY AVOID.

PRINCIPLE 1: Don't ignore the potential for discrimination. Platforms should start with more-careful tracking. Currently, most don't know the racial and gender composition of their transaction participants. A regular report (and an occasional audit) on the race and gender of users, along with measures of each group's success on the platform, is a necessary (though not sufficient) step toward revealing and confronting any problems. It can shed light on areas where discrimination is an issue and reveal progress over time. It can also be a good-faith first step toward reducing discrimination. For example, Airbnb should regularly report the acceptance rates of guests broken out by factors such as race and gender. Making this information public would help raise user and regulator awareness and keep pressure on companies to deal earnestly with discrimination problems that arise as their platforms evolve. (Public disclosure of discrimination-related data is one dimension on which Airbnb's announced policies fall far short—but it's needed to ensure that

the company's broad, laudable goals translate into concrete results.)

PRINCIPLE 2: Maintain an experimental mindset. Platforms should do what they do best—experiment. Companies including Facebook, Yelp, and eBay have baked experimental thinking into their development of new products and features. To test design choices that may, along with other interventions, influence the extent of discrimination, companies should conduct randomized controlled trials. Airbnb should be applauded for a recent experiment in withholding host photos from its main search-results page to explore the effects on booking outcomes (although it has not made the results public).

IN MANY CASES, THE SIMPLEST, MOST 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.

DESIGN DECISION 1: Are you providing too much information? In many cases, the simplest, most 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. For many others, however, it would mean departing from the way they do business. An executive of a platform with a billion-dollar valuation told us that his firm would never consider eliminating photos or names.

In addition to choosing what information to reveal, platforms choose how salient to make it. And a large body of evidence has shown that salience matters. On some platforms, for example, shipping costs are separate from—and less salient than—the base price. In an influential experiment, economists Jennifer Brown, Tanjim Hossain, and John Morgan demonstrated that in this situation, a lower base price increases the chance that an item will sell, even when it is offset by a higher shipping charge. In other words, a customer is influenced not only by the information he sees but also by which information is most prominent.

To see how this insight might be applied, recall the comparison of Airbnb, which displays host photos on its main search-results page, and HomeAway, which does not. (In September, Airbnb stated that it will test alternative ways of presenting photos and other race-relevant information, although it did not commit to specifics.) By reducing the salience of race, platforms could reduce discrimination.

DESIGN DECISION 2: Could you further automate the transaction process? When using Uber, you tap the screen to order a ride; only after confirming do you learn who will pick you up. In theory, you can then cancel if you don't like the driver's rating or looks. But that takes effort, and this small "transaction cost" is probably just enough to deter most looks-based cancellations. Uber could just as easily have allowed riders to see the driver before tapping confirm or cancel, but it chose not to.

Having transactions occur before race and gender are revealed makes it more difficult for people to discriminate. Consider the Airbnb feature known as "instant book," designed to make booking simpler and more convenient. A host using it allows renters to book her property without her having first approved them. Instant book is an opt-in feature: Landlords must sign up for it. Research has shown that default bias is strong: Most hosts will use whatever option is set up as the default. If Airbnb switched its default to instant book, requiring hosts to actively opt out of it, discrimination would most likely be lessened. The company might even consider making hosts pay for the privilege of screening customers—for example, it could charge a premium for opting out of instant book. (In September the company announced that it would accelerate the use of instant book, although it did not specify how it would accomplish this.)

We believe that increased automation and standard economic incentives, carefully implemented, could both reduce discrimination and—by eliminating some of the back-and-forth needed to complete a transaction—increase profits on a variety of platforms.

DESIGN DECISION 3: Could you make discrimination policies more top-of-mind? In a 2012 study, the research team of Lisa Shu, Nina Mazar, Francesca Gino, Dan Ariely, and Max Bazerman set

out to test whether something as simple as the location of a signature on a form could affect honesty. They observed that people are often asked to fill out information and then sign at the end to attest to its veracity. They wondered whether people would be less likely to cheat if they signed at the very beginning of the form—before filling it out. Indeed, signing at the top led to less cheating in both a lab experiment and a real-world experiment with an auto insurance company. It also worked in the context of tax returns.

There's a lesson here for marketplaces: If you want people to do something, think carefully about when to prompt them. Most platforms have policies prohibiting discrimination, but they're buried in fine print. For example, Airbnb hosts must agree not to discriminate—but they do so when first signing up to be a landlord. By the time a host is deciding whether to accept a potential renter, she has probably forgotten that agreement. Marketplaces could present anti-discrimination policies at a more relevant moment—and have the host's agreement not to discriminate occur during the actual transaction process. Some people would still violate the policies, of course, but that would require a much more conscious choice.

DESIGN DECISION 4: Should your algorithms be discrimination-aware? Design choices also determine the extent to which an algorithm leads to discrimination. Thus far many algorithm designers have ignored factors such as race and gender and just hoped for the best. But in many cases the probability that an algorithm will unintentionally achieve equality is essentially zero; recall how Google's algorithms handled ads for arrest records.

If an algorithm designer cares about fairness, she needs to track how race or gender impacts the user experience and to set explicit objectives. Does she want to ensure that black customers are not rejected at higher rates than white customers? That women are offered the same prices as men?

Google tweaked its algorithm in response to the arrest-record study, but companies can proactively monitor and respond to such problems. That might entail compensating for some users' discrimination. For example, suppose Uber noticed that some passengers consistently gave low ratings to black drivers who received five stars from most of their other riders. The company could underweight ratings from those passengers—who have revealed themselves to be discriminatory—when calculating black drivers' overall feedback scores.

A Lesson from Symphony Orchestras

Platforms exist within a larger social context, of course; we can't create a color- and gender-blind world simply by designing platforms that are less apt to facilitate discrimination. And it would be wishful thinking to imagine that every platform designer aspired to that goal; sometimes enabling discrimination is good for business. When that's the case, we can only appeal to business leaders' sense of social responsibility or hope that government regulation will intervene.

But there are many instances in which the idea of "doing well by doing good" does hold—times when platform businesses could reduce discrimination at a low cost or even while increasing profits. It's also possible that a few enlightened businesses could start a virtual cycle that forces better behavior from other market participants.

Consider how the challenge of creating diversity in U.S. symphony orchestras was met. In the mid-1960s, less than 10% of the musicians in the "big five" U.S. orchestras (Boston, Philadelphia, Chicago, New York, and Cleveland) were women. In the 1970s and 1980s, as part of a broader diversity initiative, the groups changed their audition procedures to eliminate potential bias. Instead of conducting auditions face-to-face, they seated musicians behind a screen or other divider. In a landmark 2000 study, economists Claudia Goldin and Cecilia Rouse found that the screen increased the success rate of female musicians by 160%. In fact, they attributed roughly a quarter of the orchestras' increased gender diversity to this simple change. And with selection based more squarely on musical ability, the orchestras were undoubtedly better off.

When we first read this study, many years ago, we were intrigued by the rare glimpse it provided into discrimination's effects and by the outsize impact of a small change. But the solution felt frustratingly context-specific. It was hard to imagine gender- or race-blind interactions between buyers and sellers or employers and job candidates.

The online era has changed that. Early on we witnessed the internet's potential to create marketplaces free of race, gender, and age considerations. We've now evolved far enough that platform designers can choose where and when to place virtual screens. We hope they will use that power to create a more inclusive society.

HBR Reprint R1612G