**Observational study:**

**Summary:**

The observational study conducted by Edelman and Luca tested racial discrimination against landlords in Airbnb.com. Researchers first constructed a data set combining the information on the quality of the apartment and pictures of New York City landlords. Data on quality of listed apartments on Airbnb in New York City as of July, 2012 (Edelman and Luca, pp3) contain the asking price, the characteristics of the host and of the apartment, number of guests leaving reviews, and the average rating for each characteristic of the host; photos of the listed apartments are collected to assess the characteristics of the apartment. In order to quantify the characteristics of the apartments, researchers hired Amazon Mechanical Turk workers to rate the apartments on a seven-point scale. Profile pictures were downloaded from Airbnb to collect the data on the race of the hosts, and researchers again hired Amazon Mechanical Turk workers to code the racial data. After collecting the data, researchers performed a regression of price on the listing information including housing and host characteristics to examine the factors other than racial information that might affect the price. Then, researchers performed another regression of price on the listing information, this time including the racial information (Black vs Non-black). In this regression, researchers combined different apartment and host characteristics with the racial variable to eliminate other factors that may affect guests’ perception of the housing, and therefore other apartment and host characteristics in this regression serve as controls. Researchers then calculated the racial gap after adding these controls.

**Computational methods used:**

Several observational methods discussed by Salganik are used in this research, namely, counting things and matching. From the histograms demonstrated in Figure 2, we can see that researchers counted rental price of Airbnb homes in the whole New York City, and from those in Figure 3 we can see that researchers counted rental price of black and non-black hosts separately. The observational data used for counting are massive information that comes from Airbnb. Matching is used when researchers compare black hosts with non-black hosts with similar apartment and host characteristics, the observational data that are used to compare and contrast the two groups is also extracted from online, i.e., listing information and profile pictures from Airbnb.

Researchers also adopted mass collaboration methods discussed by Salganik, i.e., human computation when they hired Amazon Mechanical Turk workers to examine pictures of apartments, to rate the apartments, and to examine profile pictures of hosts.

**Evaluation of Effectiveness**

Researchers of this observational study found that non-black hosts are able to charge 12% higher than black hosts even if location, characteristics and quality of their homes are similar, highlighting the risk of discrimination that arises from trust-building mechanisms in online marketplace.

When introducing the counting method, Salganik emphasized on identifying the interesting and important question to ask, since “the key to the research is combining good question with good data. The data alone are not enough” (Salganik). Here, we can see that the researchers raised a question that is important both economically and socially, that is, whether racial discrimination is present in online sharing economics when profile picture viewing is open to potential customers.

The data used in this research fulfill the good characteristics of big data discussed by Salganik. From Figure 1-a, we can see that there are more than 1000 rentals in the New York City by the time the author took the screenshot, suggesting the data is big enough. There are hosts posting listings and new guests booking rooms and houses and leaving reviews on Airbnb every day, so the data is always on. Researchers were simply extracting and downloading these data without contacting the hosts and the guests, and therefore the data is non-reactive.

However, we must notice a limitation of this research acknowledged by the researchers, that is, limited information about demand (Edelman and Luca, pp 5). This limitation reflects some of the bad characteristics of big data mentioned by Salganik. Airbnb is not willing to share data on consumer demand for this project, indicating the data is likely sensitive. Therefore, limitations on the information about demand shows the inaccessibility of this part of data.

**Field Experiment**

**Summary:**

The field experiment by Edelman, Luca and Svirsky examined the presence and the extent of racial discrimination against African American guests on Airbnb marketplace based names. First, the researchers collected property data from five cities with different geographic backgrounds and different levels of Airbnb usage. Participants of this experiment are Airbnb hosts in these five cities who have posted listings on Airbnb online marketplace. Amazon Mechanical Turk workers were hired to analyze profile pictures and pages of these hosts to collect information on hosts’ race, gender and age, on property information, and on past accommodation of African American guests. Researchers then created four treatment groups based on race and gender. They created four groups of guest accounts with names that indicated African American females, African American males, White females and white males, with all other guest information and characteristics identical. After creating guest accounts, researchers sent inquiries from test guest accounts to the hosts, and each host was randomly assigned one and only one guest account. Researchers then tracked host responses over the 30 days that followed and categorized responses as “Yes”, “Conditional yes”, “No response”, “Conditional no”, and “No”. In addition, to examine the likelihood of acceptance, a regression of the acceptance rate on guest’s race, host’s race, host’s gender, and several host and property characteristics was performed. From the outcomes of the experiment, researchers found that the discrimination is evident.

**Computational Methods Used:**

Researchers have adopted digital field experiment for this research. As was stated in the title of this paper, the experiment was a field experiment because it was conducted in a natural, real-life setting, rather than in a lab setting, and the participants who interacted with test guest accounts were assuming that test accounts were real guests. The experiment was digital because the treatment was delivered online, and the outcomes were also collected online.

Researchers used digital methods for randomization of participant selection, as they used a random number generator to determine which listing of one host to select.

Researchers also used observational big data and counting methods prior to the experiment. Data of hosts and properties were collected from Airbnb’s online marketplace, and the number of reviews and prior experience of accommodating African Americans were counted.

In addition, researchers adopted mass collaboration methods, specifically, human computation methods, when they hired Amazon Mechanical Turk workers to analyze profile pictures and pages of the hosts.

**Evaluation of Effectiveness:**

**Validity:** In Table 2, researchers presented the result of regression of acceptance rate on guest race information and host information. While researchers provided p-values in raw data in Table 1a, they did not provide p-values in this table. Statistical conclusion validity could be more clear if p-values for the coefficient of “Guest is African American” were provided alongside this table. In addition, it would be helpful if researchers could explain the relevantly small Adjusted R-square.

The experiment procedure has several designs that aim to improve the internal validity. First, a random number generator was used to determine which listing of the host to choose, improving the randomization. This random number generator also improved the delivery of treatment, as it prevented the host from being contacted by more than one test guest account. Second, to ensure that host characteristics were coded correctly, researchers had three AMT workers to examine the same host profile picture to reduce the chance of human error.

In section D, researchers discussed their assessment of external validity. It is possible that these specific names have special impacts on host response, and researchers made additional explanation on the generalization of experiment outcomes. By comparing the response of hosts who had accommodated African Americans in the past and those of the hosts who had not, researchers stated that there was a drop in the racial gap and such a drop could not be explained by the specialty of the experiment design.

**Heterogeneity:** Researchers have taken the heterogeneity of treatment effects into consideration. By coding hosts’ race, gender, property information, and past accommodation information, they assessed whether there were different treatment effects on different groups of hosts. They examined whether hosts preferred guests with same race and gender, but the estimates did not indicate a significant difference. They did note that they found homophily among African American females, but not among other groups (pp 17). The significant heterogeneity emphasized by this paper was the difference in past accommodation information. The race gap dropped for those who had accommodated African Americans compared to those who had not.

**Mechanism:** Researchers have discussed the limitations on the identification of the causal mechanism. This research could not identify causal factors associated with the discrimination against the names, and it failed to reject either of statistical and taste-based experiment. Mechanism behind discrimination against names and alternative models of discrimination should be examined in the future.

**Value-added of both research projects**

The observational research collected massive data on host and property characteristics. These observational data provided information on the pricing in the past up to the time of the research, and the collection of data did not interfere with the subjects observed. Observational data in this research were large, and included different variables, enabling better examination across different variables. From the observational research, patterns can be identified (in this case, lower pricing for Black hosts) and interesting and important question can be raised based on the patterns identified from the data. While this observational research did include a matching method to approximate experiments, the treatment was beyond researchers’ control. Moreover, from this observational research alone we could not see the exploration of causal effects. Researchers of this observational research briefly mentioned that was not able to disentangle statistical and taste-based discrimination with no further exploration.

The field experiment research randomly assigned guest accounts to different hosts. The assignment of treatment in this research is within the control of researchers. In addition, although researchers were not able to draw a conclusion on the causal mechanism, they were able to examine the factors that might have played a role in the causal mechanism. While no conclusion was drawn on the discrimination model, researchers were able to find evidence against pure taste-based discrimination and pure statistical discrimination, providing potential directions for future testings of mechanisms. However, before the field experiment started, researchers still needed observational data on race, gender, and property characteristics of hosts, without which heterogeneity of treatment effects could not be assessed.

A combination of the two projects will provide us with more insights. First, from massive observational data, we can identify and pose important and interesting questions, and in this case, the question was the presence and the extent of discrimination on Airbnb online marketplace. In addition, by using observational data that have good characteristics of big data, we can test for presence of certain patterns that are of economic and sociological interests. Furthermore, different variables of observational data will become critical in experiments that follows, because these variables allow for control of treatment and could provide answers to heterogeneity in treatment effects. The field experiment complements the observational research with randomized treatments. While the observational project counted pricing of all the Black hosts, the field experiment was able to obtain information on subjects’ response to names implying race. The experimental research could examine causal mechanisms, while researchers will need to refer to observational data in order to assess these causal effects. From the observational project, we could learn the presence of discrimination on Airbnb online marketplace, and the comprehensive data from the observational research provided us with the necessary information to start the field experiment; From the field experiment, we could learn the response of different groups of subjects to information that implied race information, and explorations relevant to causal mechanisms that are blocked in the observational research.

**Digital survey-based experiment**

One of the primary questions of interest drawn from these two papers is the causal mechanism that yields the discrimination. As indicated by both research projects, the discrimination is associated with people’s perceptions (perception of host images in the observational project, and perception of guest names in the field experiment project), so having data on people’s “internal state” (Salganik) is essential. This survey is going to ask questions on the factors that affect hosts’ responses and hosts’ perception on different names.

From the data collection process from the two research projects, we could see that it is possible to conduct probability sampling of hosts, since listing information includes hosts’ race, gender, age, location, property characteristics, and past accommodation information. Since there might be a significantly smaller number of African American female hosts than hosts of other races and/or gender, we could include a larger percent of African American female hosts while sampling and take a weighted mean or other statistics as necessary in later calculations. Similar methods could apply when we sample hosts with past reviews from African American guests and those without.

To deliver the survey, I am going to use the Airbnb built-in messenger. To increase the possibility of response to the survey, I am going to include coupons whose redemption code could only be visible after completion of the survey. I will ask closed survey questions, for example:

“What does name XYZ seem to imply about a guest?

1. Cultural background
2. Age
3. Race
4. Socio-economic status
5. It does not seem to imply anything”

The survey will be anonymous, and the race, gender, property characteristics and past accommodations of guests will be instead asked as questions at the beginning of survey. This might not be the perfect solution, since we could only rely on the validity of these response. It is possible that hosts might not want to disclose some of the information, resulting in measurement errors. One possible way to accommodate for this possibility is to compare the pool of profiles of respondents with the pool profiles of hosts. However, the discrepancy cannot be directly eliminated simply by comparing the information of respondents with that of the hosts, since different groups of hosts might have a different rate of response to the survey.

Another potential challenge of this survey is that Airbnb might block messages that are irrelevant to booking rooms or apartments. Under such circumstances, the alternative source to turn into would be Amazon Mechanical Turk, and the survey could be sent to AMT workers who are also Airbnb hosts. Again, there might be a gap between the whole population of hosts and the hosts who use Amazon Mechanical Turk, resulting in representation errors.

**References:**

Edelman, B. G., & Luca, M. (2014). Digital discrimination: The case of airbnb.com. Harvard Business School NOM Unit Working Paper, (14-054).

Edelman, B., Luca, M., & Svirsky, D. (2017). Racial discrimination in the sharing economy: Evidence from a field experiment. American Economic Journal: Applied Economics, 9(2), 1-22.

Salganik, Matthew J. 2017. Bit by Bit: Social Research in the Digital Age. Princeton, NJ: Princeton University Press. Open review edition.