

Analyzing Uber's Ride-sharing Economy

Farshad Kooti
USC ISI
Marina del Rey, CA

Mihajlo Grbovic
Yahoo Research
Sunnyvale, CA

Luca Maria Aiello
Bell labs
Cambridge, UK

Nemanja Djuric
Yahoo Research
Sunnyvale, CA

Vladan Radosavljevic
Yahoo Research
Sunnyvale, CA

Kristina Lerman
USC ISI
Marina del Rey, CA

ABSTRACT

Uber is a popular ride-sharing application that matches people who need a ride (or riders) with drivers who are willing to provide it using their personal vehicles. Despite its growing popularity, there exist few studies that examine large-scale Uber data, or in general the factors affecting user participation in the sharing economy. We address this gap through a study of the Uber market that analyzes large-scale data covering 59 million rides which spans a period of 7 months. The data were extracted from email receipts sent by Uber collected on Yahoo servers, allowing us to examine the role of demographics (e.g., age and gender) on participation in the ride-sharing economy. In addition, we evaluate the impact of dynamic pricing (i.e., surge pricing) and income on both rider and driver behavior. We find that the surge pricing does not bias Uber use towards higher income riders. Moreover, we show that more homophilous matches (e.g., riders to drivers of a similar age) can result in higher driver ratings. Finally, we focus on factors that affect retention and use information from earlier rides to accurately predict which riders or drivers will become active Uber users.

Categories and Subject Descriptors

H.4.3 [Information Systems]: Information systems applications

Keywords

Sharing economy; Uber; user characterization; prediction

1. INTRODUCTION

The rapid growth of the sharing economy, exemplified by ride-sharing platforms Uber and Lyft, as well as home-sharing platforms Airbnb and Couchsurfing, is changing the patterns of ownership and consumption of goods and services. In a sharing economy, consumers exchange services in a peer-to-peer fashion, through matching markets facilitated by social networks and online applications. Instead of owning a car or hailing a taxi, ride-sharing services enable consumers to request rides from other people who own private

vehicles, or in turn, become drivers offering rides. Similarly, home-sharing services enable consumers to stay in private homes while on travel, or offer rooms in their homes as short-term rentals. The various benefits provided to consumers, such as convenience, cost savings, possibility for extra income, and new social interactions, have fueled the sharing economy's dramatic growth [8].

Uber, along with Airbnb, is one of the most successful sharing economy markets. Founded in 2009, Uber is an online marketplace for riders and drivers. Riders use a smartphone app to request rides. Ride requests are assigned to Uber drivers, who use their own vehicles to provide the rides. Low prices, short wait times, as well as the convenience of simplified ride request and payment are considered the main reasons contributing to Uber's popularity among the riders [9]. On the other hand, the flexibility of work schedule and higher compensation rates are among the main reasons making Uber attractive to drivers [7].

Uber has grown wildly popular, providing more than a million daily rides as of December 2014¹ and is the most valued venture-backed company as of December 2015². Uber's popularity makes it attractive for studies aimed at understanding participation in the sharing economy. However, the system is still not well-understood. Specifically, what are the characteristics of Uber riders and drivers? What effects do different factors such as promotions, rider-driver matching, and dynamic (or *surge*) pricing have on user participation and retention? Can these factors and characteristics be used to accurately predict users' behavior on Uber, particularly whether a new user will become an active user?

We study Uber data that contains information about 59 million rides taken by 4.1 million people over a seven month period, along with data about 222 thousand drivers. This information is extracted from the email confirmation messages sent by Uber to riders after each ride, as well as weekly reports sent to drivers, collected on Yahoo servers. By analyzing usage and demographics of the population of Uber users, we find that an average active Uber rider is an individual in his or her mid-20s with an above-average income. Different demographic groups exhibit differences in their behavior: younger riders tend to take more rides, older riders take longer and more expensive rides, and the more affluent riders take more rides and are more likely to use more expensive types of Uber services, such as Uber Black.

We present a detailed analysis of Uber riders and drivers, in terms of age, gender, race, income, and times of the rides. Our main findings are as follows:

©2017 International World Wide Web Conference Committee (IW3C2), published under Creative Commons CC BY 4.0 License. WWW'17 Companion, April 3–7, 2017, Perth, Australia. ACM 978-1-4503-4914-7/17/04. <http://dx.doi.org/10.1145/3041021.3054194>



¹newsroom.uber.com/our-commitment-to-safety, accessed January 2017

²nyti.ms/1XD9cdT, accessed January 2017

- Uber is not an “all-serve-all” market. Riders have higher income than drivers and differ along racial and gender lines.
- Rider and driver attrition is very high, but the influx of newcomers leads to an overall growth in the number of rides. We identify characteristics of riders and drivers who become active users.
- Better matches of riders to drivers result in higher ratings.
- Surge pricing does not favor more affluent riders, but mostly affects younger riders (who use the service during peak times, including weekend nights).
- Drivers with many surge rides receive lower ratings, on average, suggesting the riders’ dislike of surge pricing.
- Based on a rider’s or driver’s initial activity, we can predict whether she or he will become active or leave Uber.

This work presents an in-depth analysis of the ride-sharing market from a large-scale Uber data including both riders and drivers. Our analysis reveals the demographic and socioeconomic factors that affect participation in the ride-sharing market, and enables us to predict who will become an active market participant. Since consumer retention is generally much cheaper than consumer acquisition [18], detecting customers who are likely to stop using Uber could help improve consumer retention.

2. RELATED WORK

Crowdsourcing platforms have emerged as solutions to cheaply execute large amount of independent micro-tasks that are easy to solve by human workers [11]. As opposed to virtual crowd-markets, in which tasks are executed fully online, more recently a number of platforms that support mobile crowdsourcing have sprung up (e.g., TaskRabbit, OpenStreetMap, Uber). These services have an inherently different structure, as they specifically address tasks that need physical presence in a place [6, 22, 23]. Drawing from the existing literature on characterizing crowdworkers [2, 14], we focus on the ride-sharing service Uber, studying the factors that are linked with the rate of participation.

Several studies measured the overall impact of sharing economies on traditional markets. For example, Zervas et al. showed that 8-10% of the hotel industry revenue is impacted by Airbnb [24, 25]. Rayle et al. conducted a survey [17] aimed at finding the reasons why people use ride-sharing services, as opposed to taxis or public transportation. They found that the ease of payment, shorter wait time, and faster service are the top three reasons for using ride-sharing services. In addition, 39% of survey participants stated that taxi would be their first alternative if Uber did not exist and 33% would use public transport, indicating that Uber is impacting the economy of the public transportation services as well. The clash between traditional and sharing economies has sparked a polarizing debate about the social cost of the introduction of new sharing economy services [20] and about the need for additional policies to regulate such emerging markets [5]. A recent study on the economy of Airbnb suggests that such regulations should be responsive to real-time demands [16], which in turn call for a data-driven analysis of the sharing platforms.

As ride-sharing services continue to gain popularity, the interest of the scientific community in analyzing its success factors grows as well. Several small-scale studies have been conducted on Uber. A recent Pew study³ surveyed 4.7K Americans and found that 15% of the population have used ride-sharing applications. Hall and Krueger studied Uber drivers’ activity along with the results of 601

surveys administered to drivers [7]. They found that the age distribution and education of the drivers is more similar to the general workforce than to the taxi drivers and chauffeurs, and showed that Uber’s popularity grew faster in certain cities such as Miami, Austin, and Houston. Other studies used Uber data along with data from taximeters to compare Uber rides to taxi rides [4]. Analyzing the data from the mobile application OpenStreetCab [21], the authors compared the cost of Uber rides to that of Yellow Cab rides and found that Uber effectively charges higher fares on average, especially for short but popular routes [15]. Finally, researchers analyzed the effects of Uber’s dynamic pricing, that adjusts the ride cost to the demand. Lee et al. interviewed 21 Uber and Lyft drivers [12] and found that drivers in the sample were not influenced by surge pricing information. Even though the surge pricing is an opaque mechanism that raises concerns about fairness and may cause frustration for the riders [3], in general it helps the ride-sharing marketplace [9].

We also study the changes in users’ engagement level on Uber. In the context of crowdsourcing platforms, previous work studied the main incentives that lead to user participation [13]. Airbnb users are motivated to monetize hospitality for a mixture of financial and social reasons [10]. Similarly, by looking at the relationship between activity of drivers and riders with socioeconomic indicators, we aim to find evidence that specific segments of the population may have different incentives to take part in the ride-sharing economy. Engagement level also has a direct impact on attrition or consumer churn. The importance of consumer attrition analysis is driven by the fact that retaining an existing consumer is much less expensive than acquiring a new consumer [18]. Several studies addressed churn in this context. For example, Ritcher et al. exploit the information from the users’ social network to predict consumer churn in mobile networks [19]. In this work, we study both the consumer (i.e., rider) and provider (i.e., driver) attrition.

3. DATA SET

Following each ride, Uber emails rider a receipt. This email includes pick-up and drop-off times, origin and destination addresses, duration of the ride, distanced traveled, type of the service (e.g., UberX, Uber Black), driver’s first name, and overall fare, along with a breakdown of the price, including whether or not a promotion code was used and whether the surge multiplier was applied (during peak hours the fare is multiplied by a value called the surge). We obtained information about Uber rides of Yahoo Mail users using an automated extraction pipeline that preserves the anonymity of both riders and drivers. In total, we study over 59M rides taken by 4.1M users from October 2015 to May 2016. There is a strong weekly pattern of Uber usage, with many rides taking place on weekends. Some holidays, such as New Year’s Eve and Halloween, result in large peaks in the number of rides, while others like Christmas result in a sharp drop.

Drivers receive two separate weekly emails. One includes the money earned each day of the week. And the other email includes the hours they worked each day of that week, percentage of busy hours worked, riders’ feedback text (if any), their average rating over the week, and whether the rating is higher or lower than an average. The riders rate drivers on a scale 1-5, with 1 being the worst and 5 being the best. Our data set includes more than 1.9M weekly summaries for 222K drivers. Moreover, whenever a person joins Uber they receive a welcome email. Thus, besides the ride information, we know when a user joined Uber either as a rider or a driver.

In addition to Uber emails, we relied upon the Yahoo Mail network graph during the same period of time. The email graph G

³www.pewinternet.org/2016/05/19/on-demand-ride-hailing-apps, accessed January 2017

consists of pairs of hashed user IDs that communicated with each other. For the present analysis, we consider the ego network of the users who are Uber riders and/or drivers.

Finally, we also collected users’ demographic information as specified in their Yahoo Mail account. That includes age, gender, and location at the ZIP code level. We conducted our analysis only on users from the US, unless otherwise stated. Further, only for purposes of this study we produced income and race *estimates* for all riders and drivers. Since Yahoo does not collect declared income or race information during sign-up process, we derived estimates using publicly available US census data that contains race and income distributions for each ZIP code. All drivers and riders from a specific ZIP code were assigned the median income and race associated with that ZIP code. As a result, the inferred incomes and race for users are aggregated estimates (we do not know the ground truth for any specific user) on which our presented results are based. Nevertheless, we will see that such coarse appraisal is enough to observe clear trends in the data.

Limitations. Our data set has a few limitations. First, our data only includes Uber users who are also Yahoo Mail users, and there may be a selection bias in the subset of users being studied. While this might happen to some extent, given the popularity of Yahoo Mail with over 300M users⁴, we believe the considered Uber population is large enough to be a representative sample and the findings could be generalized. Second, our data does not include ratings of individual rides. This information is shared neither with the riders nor with the drivers in the email, or even on their private profiles. To answer questions regarding the ratings, we take multiple steps to find the subset of drivers whose vast majority of rides were included in our data. We explain this in more details later in the paper.

4. ANALYZING RIDERS

In this section, we examine the relationship between the demographics and characteristics of Uber riders and their activity using the service. We answer questions such as: Who is a typical Uber rider? Who are the most active riders? Who is most affected by surge pricing? At what rate do riders stop using Uber?

4.1 Demographics and number of rides

A typical Uber rider is young (38% of riders are 18–27 years old) and slightly more likely to be a woman (51% are women). Female riders are somewhat younger than males (mean age of men is 34.6 years vs. 33.1 years for women). The vast majority of riders are white (80.5%), followed by Hispanic (8.5%), African-American (8.2%), and Asian-American (2.8%). Table 1 breaks down riders by race, age, and gender. Hispanic and African-American riders are younger than white and Asian-American riders, but the median number of rides is 3 for all races.

We consider the average number of rides per week as a measure of riders’ activity. We found that in general, older riders use the service less frequently, e.g., 30-year-old men use Uber 20% more than 50-year-old men, see Figure 1(a). Although young men and women (aged less than 25 years) use Uber at about the same rate,

⁴www.comscore.com, accessed January 2017

Table 1: Comparison of riders by race, age, and gender.

Race	%riders	%women	Avg. age
White	80.5%	49.6%	36.1
Hispanic	8.5%	52.2%	31.1
African-American	8.2%	61.1%	32.4
Asian-American	2.8%	50.4%	35.5

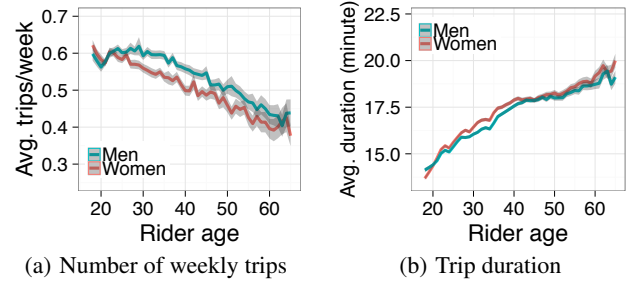


Figure 1: Rider activity as a function of age and gender along with 95% confidence interval.

older men use it slightly more than older women. The values shown in the figure are the averages for a given age and gender. The frequency of rides has a heavy-tailed distribution: most riders have a very low activity, while a few riders are very active. The median number of rides overall is only 0.2 rides per week, and the top 10% most active riders take 1.3 rides or more per week.

4.2 Duration, length, and cost of rides

The duration, cost, and length of rides are all significantly correlated. However, studying each individually helps us understand the type of rides that are taken by users in different demographic segments.

Figure 1(b) shows the average duration of rides taken by riders of a given age and gender. In general, rides tend to be relatively short (median duration is 14 minutes). Older riders take longer trips on average (average length of rides of 60-year-old men is 30% longer than those of 20-year-old men). Women and men do not vary significantly with respect to the length of the rides. The distance traveled shows an almost identical trend. Most of the rides are relatively short, with 50% of the rides being shorter than 4 miles, but 10% of the rides are longer than 36 miles.

Since travel time largely determines the fare, the trend in the cost of rides is similar to that of the duration of the rides. The average fare ranges from \$13–\$21, with older riders spending more per ride. Men and women are very similar, except middle-aged men spend significantly more than middle-aged women on the rides. The median fare of all rides is \$10, and the total money spent by each rider has a heavy-tailed distribution, with a small fraction of the riders responsible for a considerable fraction of the total spending. The top 1% of riders account for 18.8% of all money spent on fares in our data set. The Gini coefficient for rider spending, which measures inequality of a distribution on a scale from 0 (perfect equality) to 1 (maximum inequality), is 0.785 for total fares, showing a very high heterogeneity in spending among riders (compare this to Gini coefficient of 0.394 for 2014 US income inequality⁵).

4.3 Income, surge, and car type

Next, we examine the impact of rider income, surge pricing, and car type, on rider activity. First, we are interested to find who is most affected by surge pricing: lower income riders, who may be priced out by the increase in fares during peak hours, or the more affluent riders who are willing to pay more for rides during times of high demand. We estimate the household income based on the ZIP code of their self-declared home location. We use median income for a given ZIP code as the user’s estimated income. Older

⁵www.oecd.org/social/income-distribution-database.htm, accessed January 2017

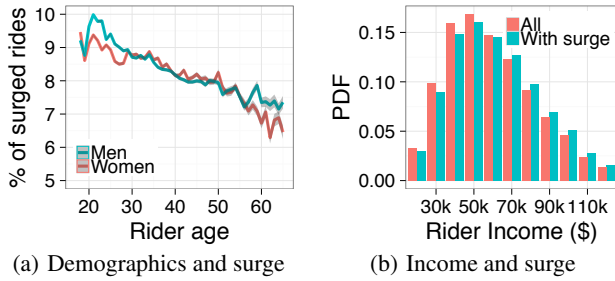


Figure 2: Riders and surge pricing along with 95% confidence interval. (a) Percentage of rides with surge pricing as a function of rider age and gender. (b) Comparison of income of riders who had at least one ride with a surge fare and rest of the riders.

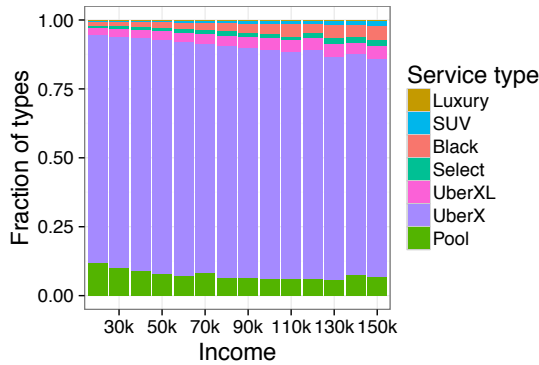


Figure 3: Type of Uber car requested by riders given their income.

riders have higher income compared to younger riders. However, the percentage of the rides with surge pricing for a given age and gender has exactly the opposite trend: older riders are less likely to pay the surge price, see Figure 2(a). This might be due to younger riders using Uber more during peak times, e.g., weekend nights, when there is surge pricing due to high demand. Even though these trends seem to suggest that riders paying surge prices should have lower income than riders not paying the surge prices, we find that riders with at least one surge ride have slightly higher income than riders who never paid surge pricing, see Figure 2(b). These plots seem to be conflicting, but they can be explained simply by the large heterogeneity among users. In short, people with higher income are more likely to take rides with surge pricing, but age plays a much more significant role.

Uber offers different service options: budget options, such as UberX and UberXL, and more expensive luxury options, such as Black, Select, SUV, and Luxury. More recently, Pool ride is the cheapest option as it allows the rider to split the trip cost with another person headed in the same direction. Figure 3 compares the type of Uber cars requested by riders with different incomes. There is a clear trend, with more affluent riders requesting more expensive cars. For example, people with annual income of \$100k are 84% relatively more likely to take an Uber Black compared to users with annual income of \$50k.

4.4 Ride dynamics

To understand dynamics of rides, we start by examining diurnal variations in the number of rides taken. Human behavior generally

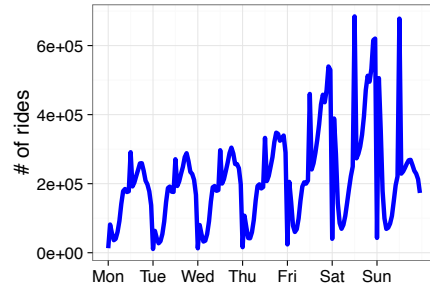


Figure 4: Number of rides at different times of the day and days of the week.

exhibits very strong daily and weekly trends, which is also reflected in Uber activity, see Figure 4. Weekday morning and afternoon peaks represent riders who use Uber to commute to work. Also, we see peaks during the lunch hour, showing increased usage of Uber for going to restaurants.

Next, we study round trip rides. Knowing which rides result in a return trip would be helpful for the system in predicting the needs of the rider and scheduling drivers accordingly. Overall, 9.1% of all rides have another ride from the exact same location back to the previous ride’s origin. About 2.5% of riders make round trip rides. If we consider the time the first ride took place, then 9.9% of rides starting 5am-12pm have a return ride, 10.2% of rides in 12pm-7pm, and 7.6% of rides in 7pm-5am, showing that the rides in the afternoon are more likely to have a return ride.

4.5 Promotions

Uber uses promotions to attract new riders, for example, offering them a free first ride. We extract all rides in which a promotion was used, and compare the characteristics of riders using promotions to the rest of the riders. While we cannot make any causal claims, any uncovered trends could suggest whether promotions are in fact a useful tool to attract new riders. Table 2 compares the characteristics of riders who used promotions to the rest of the population. Riders who used promotions are younger, less active, and have lower income. And interestingly, they are more likely to stop using Uber early on and drop out, if we define dropping out as not taking any rides after the first week.

4.6 Rider attrition

What happens after a rider’s first ride, whether or not a promotion was used? Does the rider become an active Uber user? Or does the rider stop using Uber and revert to his or her previous transportation options? Given the high costs of attracting new customers (advertising, promotions), retaining them is an economic priority for businesses.

Table 2: Comparison between riders who used promotions to those who did not.

	Promotion	No promotion
% men	44.3%	46.1%
Average age	31.5	34.0
Median # of rides	2	3
Median income	\$50.0K	\$62.5K
% drop out	59.9%	55.6%

To measure rider attrition, we focus on riders who took their first ride during our data collection period and measure changes in their engagement levels over time. Recognizing new riders is feasible due to the welcome email they receive from Uber upon signing up. We exclude riders who took their first ride during the last four months of our data collection period, to ensure that we have at least four months of rider activity records for the new riders. We also exclude riders who took only one ride during this period (11.5% of riders), because low activity rates could bias results. After filtering, we still remain with large number of riders, namely 295K riders. Next, we characterize each rider with a vector containing the number of rides taken in each month following their first ride. To identify different groups of riders who have similar behavior, we ran a k -means clustering algorithm over the riders.

To find the optimal number of clusters we perform a parameter sweep from $k = 2$ to $k = 15$. The mean square error (i.e., distance from the center of clusters) gradually decreases as k increases, but with diminishing returns; after $k = 3$ the error reduction becomes significantly smaller. We chose $k = 3$ to balance between compactness of the model and the quality of clustering. Table 3 shows the number of riders belonging to each cluster, as well as the centers of the three clusters. We see that the vast majority of riders (90.9%) belongs to the cluster that has almost no rides after the first month (labeled *Inactive*). The second cluster of riders (8.0%) has a medium level of activity, almost 1 ride a week (*Low activity*). Finally, the remaining riders (1.1%) are highly active and maintain high levels of activity over time.

The Inactive cluster includes riders who abandon the service quickly, while the remaining two clusters include more active riders. To characterize these riders, we break down each cluster by demographics in Table 3, showing that more active riders are younger than riders who eventually leave Uber, but we do not find a significant difference between the groups in their gender composition.

5. ANALYZING DRIVERS

In this section, we conduct an analysis of Uber drivers, focusing on their demographics and earnings, and identify factors that affect driver retention.

5.1 Demographics

In the US, there is a significant difference between the number of male and female Uber drivers, with 76% of the drivers being male and typically in their 30s. Other countries differ widely with respect to driver gender. The US has the highest percentage of women drivers (24.0%), followed by Malaysia (10.1%), Singapore (9.9%), and Canada (9.4%). Surprisingly, the UK has a much lower fraction: only 4.3% of all UK drivers are women.

Moreover, there are significant disparities in the racial composition of drivers and riders, and how they use the service (see Table 4). For example, while still the majority of drivers are white (60%), this is much smaller than the percentage of white riders (81%). The table also shows differences in number of hours worked and earnings of drivers from different races. The last two columns show the race distribution of the US workforce and of taxi drivers from census data⁶ for comparison. The racial distribution of Uber drivers is neither very close to the US workforce nor the taxi drivers, and still closer to the taxi drivers. Earlier studies have shown that Uber drivers are more similar to general US workforce than taxi drivers

⁶<http://census.gov/programs-surveys/acs/technical-documentation/pums.html>, accessed January 2017

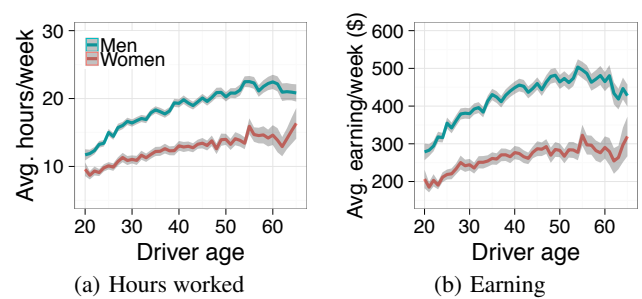


Figure 5: Average number of hours worked and weekly earnings of drivers, given their age and gender along with 95% confidence interval.

in terms of age and education [7], but our results show that this does not hold for the racial distribution.

5.2 Hours worked, income, and rating

Next, we examine the weekly number of hours drivers worked, their income, and ratings in the US. Figure 5(a) shows the average number of weekly hours worked by drivers of a given age and gender. Interestingly, older drivers work longer and are more likely to be full-time Uber drivers. The majority of drivers worked part-time, and only 19% worked 40 hours or longer in a week. The weekly hours worked is the length of the time the driver was active and received ride requests, and not necessarily the hours the driver was driving.

Our data set also includes the rate the driver was paid that week. Figure 5(b) shows the weekly earnings for drivers of a given age and gender. The main factor affecting the earnings is the surge pricing, and drivers working during the peak hours earn much more per hour. Younger drivers earn more, and men earn slightly more than women. Considering all weeks drivers worked, 25% of drivers were paid \$21-25 per hour and 1.5% of drivers had a rate lower than \$7.25 per hour, the federal minimum wage in the US. Please note that the earnings are not net earnings and costs of gas, insurance, and maintenance should be considered for calculating net earning. One report estimates these costs to be as high as one third of total earning⁷.

Figure 6(a) compares the number of hours worked by drivers with at least one surge ride with the rest of the drivers, and Figure 6(b) shows the earnings of these drivers. Drivers who have at least one surge ride, work almost as much as the rest of the drivers, but they earn significantly more: the median weekly earning for drivers with a surge ride is \$180, but the rest of the drivers have a median of only \$80. And drivers who have at least one surge ride, on average earn 60% more than the rest of the drivers, while working the same number of hours.

Some studies suggest that surge pricing is frustrating for riders [1, 3]. To verify that, we group the weeks based on the percentage of earning from surge, as an estimate of the percentage of the surge rides. Then, for each group, we calculate the percentage of the weeks that have an above average ratings. Figure 7 shows that initially the rating increases slightly as the drivers serve more surge rides, but the drivers who have many surge rides, receive worse ratings. This is in-tune with the earlier studies and suggests that earning more money could come in the expense of worse ratings.

Finally, we look at the ratings of drivers with respect to their age and gender. Drivers are rated by riders after each ride. Figure 8

⁷<https://goo.gl/u6ty6U>, accessed January 2017

Table 3: Size and centers of clusters of riders from their monthly number of rides along with their demographic breakdown.

Clusters	% riders	Month 1	Month 2	Month 3	Month 4	Avg. age	% women
Inactive	90.9%	2.1	0.4	0.4	0.5	35.1	53.3%
Lo activity	8.0%	8.5	5.8	5.6	5.6	31.9	51.3%
Hi activity	1.1%	18.0	21.6	23.3	22.1	31.2	52.1%

Table 4: Comparison of drivers of different races.

Race	% of drivers	% women	Avg. age	Avg. hrs worked	Avg. earning	US workforce	US taxi drivers
White	60.0%	21.9%	41.9	15.4hrs	\$355	75.2%	52.3%
African-American	21.6%	36.5%	40.8	14.8hrs	\$341	11.6%	25.2%
Hispanic	13.7%	23.9%	38.5	15.2hrs	\$378	7.6%	10.3%
Asian-American	4.7%	16.4%	41.6	18.2hrs	\$511	5.6%	12.2%

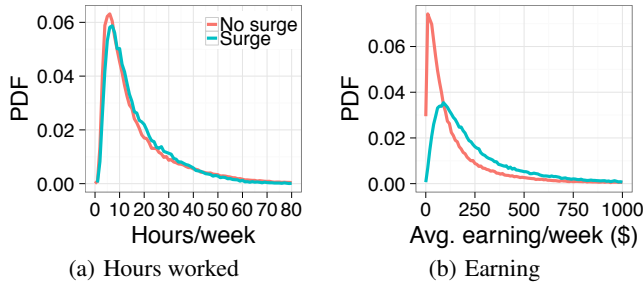


Figure 6: Comparison of drivers with at least one surge ride to the rest of the drivers.



Figure 7: Percentage of above-average ratings given the percentage of surge earnings in a week.

shows the fraction of weeks that the driver of a given age and gender had an above-average rating. Generally, older drivers tend to receive lower ratings and are more likely to receive a below average rating. Also, among 30-50 year-old drivers, women tend to receive higher ratings.

5.3 Driver retention

We study factors that correlate with driver activity. Similar to our analysis of riders, we cluster drivers based on the number of hours worked each month since joining Uber. With $k = 3$ clusters, a large fraction of drivers stop working almost completely, driving fewer than five hours during a period of a month. However, this fraction (73.3%) is much lower than the fraction of riders who stop using Uber (90.9%). The lower rate of driver attrition could be due to the higher effort required to become an Uber driver compared to an Uber rider [7]. About 21.0% of drivers have at least half an hour driving per day on average over the four months and the remaining 5.7% of drivers are very active, working longer than three

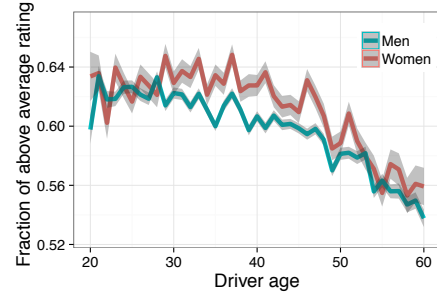


Figure 8: Percentage of above-average rating weeks for a given age and gender .

hours/day on average. The number of hours that the drivers worked drops with time across all three clusters.

We also characterize the drivers in each cluster in Table 5. The first cluster (labeled *Inactive*) includes drivers with the lowest engagement levels, who eventually stop driving for Uber, while the other two clusters contain active drivers with different engagement levels. Active drivers tend to be older, mostly men, compared to the Inactive drivers.

6. RIDER VS. DRIVER

In this section, we answer the questions that involve both riders and drivers at the same time, including comparison of their demographics, and studying the effect of matching on ratings.

6.1 Demographic comparison

First, we are interested to see if Uber has a “all-serve-all” economy, or the riders have different demographics and income distribution. As shown in Figure 9, riders have higher income compared to drivers: median income for a rider is \$62.4k and the median income of drivers is \$55.3k. Also, riders are 51% more likely to be men, and 7.3 years younger than drivers on average. If we pick a random rider and driver, the rider is 34% more likely to be white than the driver and the driver is 5 times more likely to be African-American than the rider. Even though the profiles of riders and drivers differ significantly, a considerable 17.4% of drivers are also riders.

6.2 Effect of matching

We consider the age and gender of riders and drivers to see if there is any pattern in the ratings with respect to the match. We take the following steps to match a rider and a driver:

Table 5: Size and centers of the clusters of drivers from their monthly hours worked along with demographics of each cluster.

Clusters	% drivers	Month 1	Month 2	Month 3	Month 4	Avg.age	% women
Inactive	73.3%	20.1	4.7	3.0	2.2	37.1	40.3%
Lo activity	21.0%	89.9	45.1	26.3	16.4	43.2	31.3%
Hi activity	5.7%	150.3	133.8	126.8	94.1	43.6	25.3%

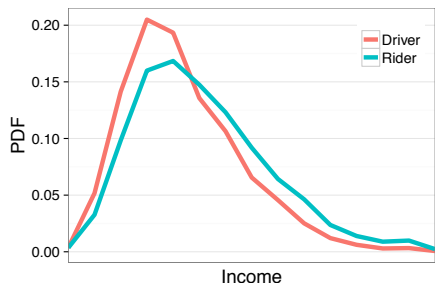


Figure 9: Comparison of income of riders and drivers.

1. To ensure that user privacy is always preserved, our automatic email extraction pipeline hashes any personally identifiable information from the email content. For example, Uber ride receipts contain a message with driver’s first name: “Thank you for driving with David”. Our pipeline detects and encrypts the first name of the driver (i.e., replaces it with a hashed value). The same procedure is applied on the first name of the rider.
2. Retrieve driver’s first name hash and date of the ride from the rider’s email receipts.
3. Match to all drivers with the same first name hash from the weekly summary emails.
4. Eliminate the drivers who did not make more than the fare of the ride in that day.
5. Eliminate drivers who are in a different state than the rider.
6. Consider a match if there is only one driver left.

The rating for each ride is not shared with the riders nor drivers, and we only get the weekly summary of ratings for drivers. So, we find the drivers that have large enough number of rides in a week (at least 10), and large enough fraction of their rides were matched with a rider (at least 75%). Then, we only consider these driver-weeks and compare the rating for the weeks that the rating is above average with the weeks that the rating is below average, given the age or gender difference among the riders and the driver. For quantifying the effect of age, we compare the average age difference among riders and drivers for the weeks with an above-average rating, and compare that with the average age difference for the weeks that the ratings were below average. We find that age plays a considerable role. The average age difference is 1.7 years smaller in the weeks with above average rating, 11.4 years (± 0.47 , 95% confidence interval) years vs. 13.1 (± 0.56).

To measure the effect of gender, we consider women and men drivers separately, and compare the percentage of above average weeks given the fraction of riders who are men/women. Gender plays an interesting role for women drivers, where lower percentage of women riders resulted in higher ratings. The weeks that the majority of the riders were men, were 12.4% more likely to have an above average rating, see Table 6. The trend is more complicated for men drivers, where having 0-45% or 55%-100% men riders,

results in higher rating compared to the cases where there are 45%-55% men drivers. The filtering and statistically significant result were only possible due to the large size of our data set.

7. PREDICTION

We use the findings presented in earlier sections to predict whether or not a new rider or driver will become an active Uber user.

7.1 Predicting rider activity

We define the prediction problem as follows: given all the information about a rider and his or her Uber activity during the *first two weeks* since joining Uber, will that person become an active rider or not? We define as active riders those who take six or more rides in weeks 3–8 of using the service (i.e., at least one ride a week on average). To this end we use the following sets of features:

- **Rider characteristics:** age, gender, location, income, education, and ZIP code.
- **Ride features:** # of rides, average distance, average price, average duration, fraction of rides in the second week, percentage of rides with surge pricing, number of cities Uber was used in, fraction of the rides in the weekend and weekday, fraction of rides in the morning, afternoon, or at night, fraction of rides with a promotion, and number of distinct origins and destinations.
- **Driver features:** for the rides that have been matched: driver demographics, driver ratings, and age and gender difference of riders and drivers.
- **Social features:** number of Uber rider and Uber driver friends based on the email network graph.

We extract all of the above features, and balance the classes by under-sampling the larger class. This results in half of the users in the data set being active users (50% baseline for random prediction). Then, we select a random set of 80% of the users for training and use the remaining 20% for testing. We use the C5.0 classifier [26] for our predictions and achieve accuracy of 75.2%, which is a 30.4% relative improvement over the baseline. The precision is 0.786 and the recall is 0.687. We also define a stronger baseline, which predicts that a user will become active if that user had more than two rides in the first two weeks (median number of rides taken by all riders in the first two weeks). This baseline performs much better than the random baseline and achieves an accuracy of 74.3%, which is still slightly lower than our classifier. This simple base-

Table 6: Percentage of above average weeks for women and men drivers, given the percentage of women or men drivers.

WOMEN DRIVERS			MEN DRIVERS		
% women riders	% above avg weeks	Std error	% men riders	% above avg weeks	Std error
0%-45%	62.6%	3.0	0%-45%	60.2%	1.6
45%-55%	53.4%	3.2	45%-55%	57.2%	1.4
55%-100%	50.2%	3.2	55%-100%	61.9%	1.2

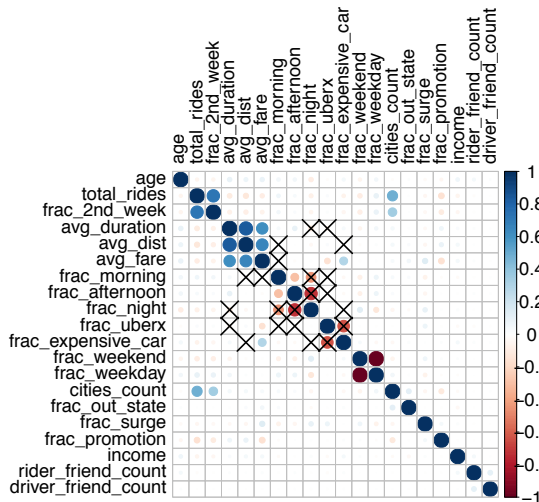


Figure 10: Correlation between the features of the riders. Pairs without statistically significant correlation are crossed (p-value < 0.05).

Table 7: Results of logistic regression on the independent variables for the riders. *p-value < 0.001, **p-value < 0.01, *p-value < 0.05**

Variable	Coeff.	Variable	Coeff.
Total # of rides	0.340***	Age	-0.020***
% out of state rides	0.330	% rides at night	-0.098***
# different cities	0.113***	# driver friends	-0.162
Gender (men)	0.054***	% rides weekends	-0.382***
# rider friends	0.006	% ride expensive cars	-0.645***
Average fare	-0.009***	% surge rides	-0.837***

line indicates that the signals from the activity of the users is strong enough to be an accurate predictor.

We use logistic regression to quantify the importance of the features. Since regression is sensitive to colinearities in the data, we first eliminate correlated features, by calculating pair-wise correlation coefficient and randomly removing one of the features that has high statistically significant correlation with another feature (> 0.7 or < -0.7), see Figure 10. Table 7 shows the results of the logistic regression on the remaining 12 independent variables, which we normalized first. Older users are less likely to become active Uber riders. Men and riders with more trips in the first two weeks are more likely to become active riders, but those who had more expensive rides, used more expensive car types (such as Uber Black), and had higher fraction of rides on the weekends, are less likely to become an active rider in the future.

7.2 Predicting driver activity

We conduct a similar prediction task for the drivers, based on the hours worked instead of the number of rides. We define active drivers as those who worked 10 hours or more per week in weeks 3–8 since joining Uber. Those who worked less than 10 hours a week in weeks 3–8 are deemed inactive drivers. We consider all the user characteristic and social features mentioned for the riders,

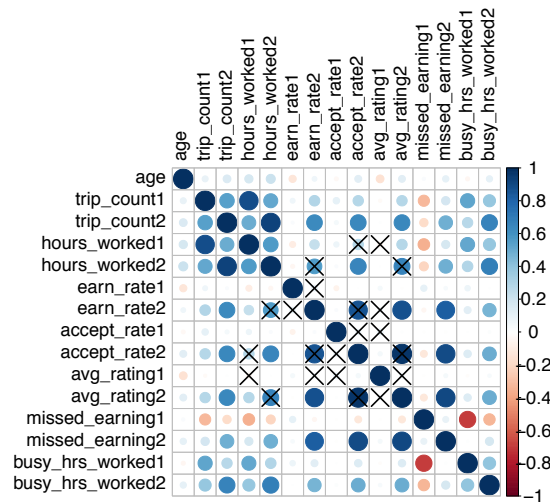


Figure 11: Correlation between the features of the drivers. Pairs without statistically significant correlation are crossed (p-value < 0.05). 1 and 2 at the end of feature name indicate the value from the first or second week. Earn rate: \$ earned per hour. Acceptance rate: % of ride requests accepted. Missed earning: Money the driver could’ve earned if drove during busy hours determined by Uber.

Table 8: Results of logistic regression on the independent variables for the drivers. *p-value < 0.001, **p-value < 0.01, *p-value < 0.05**

Variable	Coeff.	Variable	Coeff.
Gender (men)	0.371**	Earning rate	0.029*
Driver’s rating	0.227	Money missed in the week	-0.002
# hours drove	0.157***	Acceptance rate	-0.015*
Age	0.037***	# busy hours worked	-1.479

and add the following features that are specific to drivers: # of days worked, # of hours worked, # of rides given, ratings, rate of earning, % of busy hours worked, acceptance rate, and missed earnings for week 1 and 2 separately.

With the same setup as above we achieve 83.1% accuracy, which is 66.2% relative improvement over the 50% baseline. Precision is 0.775 and recall is 0.689. If we define a similar strong baseline as for riders, using the median hours worked in the first two weeks, the baseline achieves accuracy of 81.9%. This is significantly higher than the random baseline, yet again slightly lower than our trained classifier. Similarly to the case of riders above, early driver behavior is a very strong indicator of future engagement.

Moreover, we remove the correlated features (using Figure 11) and carry out logistic regression over non-correlated features. Table 8 shows that older users, men, and drivers who worked more and had higher earning rates are more likely to become active drivers, but the drivers who had a lower acceptance rate (% of rides they accepted to deliver) are less likely to become an active driver.

8. CONCLUSION

This work characterizes Uber’s riders and drivers. We consider age and gender, and race and show how different populations be-

have differently. For example, younger riders use Uber more frequently compared to older riders, but they take shorter rides. Considering gender, while the riders have balanced gender split, drivers have a very imbalanced split, with 76% of drivers being male. We also show that riders have about \$12k higher annual income than drivers. Study of surge pricing shows that drivers who take advantage of busy hours can earn on average 60% more, while working the same number of hours.

We also study the ratings given to the drivers by riders. We find that older drivers tend to get lower ratings, and women drivers who are 30-50 years old tend to get higher ratings. Interestingly, the matching of riders and drivers has an effect on the ratings: rider and driver having smaller age difference results in a higher rating, and for women drivers the rating tends to be higher when larger fraction of riders are men. These findings could be used to perform a better matching and improve user experience.

Finally, we focus on users' engagement levels and show that vast majority of users become less active and drop out after just a few weeks. By leveraging our findings, we were able to predict the users who will become active riders or drivers with higher accuracy than alternate methods. Prediction of user attrition or abandonment can be helpful for Uber to focus on these users, as retaining existing users is much less expensive than acquiring new ones.

Acknowledgements

The work was funded in part by the Army Research Office under contract W911NF-15-1-0142.

9. REFERENCES

- [1] This is your brain on uber. <http://www.npr.org/2016/05/17/478266839/this-is-your-brain-on-uber>. Accessed: 2016-05-27.
- [2] M. Buhrmester, T. Kwang, and S. D. Gosling. Amazon's mechanical turk a new source of inexpensive, yet high-quality, data? *Perspectives on psychological science*, 6(1):3–5, 2011.
- [3] L. Chen, A. Mislove, and C. Wilson. Peeking beneath the hood of uber. In *Proc. of IMC*, pp. 495–508, New York, NY, USA, 2015. ACM.
- [4] J. Cramer and A. B. Krueger. Disruptive change in the taxi business: The case of uber. 2015.
- [5] B. G. Edelman and D. Geradin. Efficiencies and regulatory shortcuts: How should we regulate companies like airbnb and uber? *Harvard Business School NOM Unit Working Paper*, (16-026), 2015.
- [6] M. F. Goodchild. Citizens as sensors: the world of volunteered geography. *GeoJournal*, 69(4):211–221, 2007.
- [7] J. V. Hall and A. B. Krueger. An analysis of the labor market for uber's driver-partners in the united states. *Princeton University Industrial Relations Section Working Paper*, 587, 2015.
- [8] J. Hamari, M. Sjöklint, and A. Ukkonen. The sharing economy: Why people participate in collaborative consumption. *Journal of the Association for Information Science and Technology*, 2015.
- [9] J. Horpedahl. Ideology über alles' economics bloggers on uber, lyft, and other transportation network companies. *Econ Journal Watch*, 12(3):360–374, 2015.
- [10] T. Ikkala and A. Lampinen. Monetizing network hospitality: Hospitality and sociability in the context of airbnb. In *Proc. CSCW*, pp. 1033–1044, New York, NY, USA, 2015. ACM.
- [11] A. Kittur, E. H. Chi, and B. Suh. Crowdsourcing user studies with mechanical turk. In *Proc. of CHI*, pp. 453–456, New York, NY, USA, 2008. ACM.
- [12] M. K. Lee, D. Kusbit, E. Metsky, and L. Dabbish. Working with machines: The impact of algorithmic and data-driven management on human workers. In *Proc. CHI*, pp. 1603–1612, New York, NY, USA, 2015. ACM.
- [13] J. Lindqvist, J. Cranshaw, J. Wiese, J. Hong, and J. Zimmerman. I'm the mayor of my house: Examining why people use foursquare - a social-driven location sharing application. In *Proc. CHI*, pp. 2409–2418, New York, NY, USA, 2011. ACM.
- [14] W. Mason and S. Suri. Conducting behavioral research on amazon's mechanical turk. *Behavior Research Methods*, 44(1):1–23, 2011.
- [15] A. Noulas, V. Salnikov, R. Lambiotte, and C. Mascolo. Mining open datasets for transparency in taxi transport in metropolitan environments. *EPJ Data Science*, 4(1):1–19, 2015.
- [16] G. Quattrone, D. Proserpio, D. Quercia, L. Capra, and M. Musolesi. Who benefits from the sharing economy of airbnb? In *Proc. of WWW*, pp. 1385–1394. 2016.
- [17] L. Rayle, D. Dai, N. Chan, R. Cervero, and S. Shaheen. Just a better taxi? a survey-based comparison of taxis, transit, and ridesourcing services in san francisco. *Transport Policy*, 45:168–178, 2016.
- [18] F. Rechinheld and W. Sasser. Zero defections: Quality comes to service. *Harvard Business Review*, 68(5):105–111, 1990.
- [19] Y. Richter, E. Yom-Tov, and N. Slonim. Predicting customer churn in mobile networks through analysis of social groups. In *Proc. of SDM*, volume 2010, pp. 732–741. SIAM, 2010.
- [20] B. Rogers. The social costs of uber. *University of Chicago Law Review Dialogue*, 82, 2015.
- [21] V. Salnikov, R. Lambiotte, A. Noulas, and C. Mascolo. Openstreetcab: Exploiting taxi mobility patterns in new york city to reduce commuter costs. *arXiv preprint arXiv:1503.03021*, 2015.
- [22] S. A. Sheppard, A. Wiggins, and L. Terveen. Capturing quality: Retaining provenance for curated volunteer monitoring data. In *Proc. CSCW*, pp. 1234–1245, New York, NY, USA, 2014. ACM.
- [23] R. Teodoro, P. Ozturk, M. Naaman, W. Mason, and J. Lindqvist. The motivations and experiences of the on-demand mobile workforce. In *Proc. of CSCW*, pp. 236–247, New York, NY, USA, 2014. ACM.
- [24] G. Zervas, D. Proserpio, and J. Byers. The rise of the sharing economy: Estimating the impact of airbnb on the hotel industry. *Boston U. School of Management Research Paper*, (2013-16), 2015.
- [25] G. Zervas, D. Proserpio, and J. W. Byers. The impact of the sharing economy on the hotel industry: Evidence from airbnb's entry into the texas market. In *Proc. of EC*, pp. 637–637, New York, NY, USA, 2015. ACM.
- [26] R. Quinlan. Data mining tools see5 and c5. 0. 2004.