Response to the FTC Workshop on Ridesharing:
Date: 20 May, 2015

In response to the public call put out by the FTC for consultation on various aspects of ridesharing business practices and how they impact consumers, businesses and employee/contractors, we chose to focus on one of the most contentious issues that afflicts ridesharing operations - reputation systems and trust mechanisms.

We are a group of academics and drivers employed in the ridesharing industry as independent contractors, seeking to make a collective representation to the FTC on the existing problems with the design of interactions within mobile applications of services such as Uber and Lyft but not limited to them. We conducted a survey to source responses from ridesharing drivers within the United States. We present a brief summary of the responses, key issues highlighted and changes suggested by the driver community in the designs of ridesharing platforms.

We have consolidated the responses along four verticals:

1) biases in the reputation system
2) passenger understanding of the reputation system
3) effectiveness of the reputation system
4) driver suggestions to mend the reputation system

Bias in reputation systems:

From over 115 responses received from across the United States, most drivers have identified "Lack of Awareness" as the biggest problem affecting transactions in Uber, Lyft and other ridesharing platforms. Drivers explain that TNC companies are not building upon the existing understanding of ratings, safety and public behavior in taxis. While traditionally, taxi drivers do not get "rated" for their services or behavior, excessive emphasis is laid on driver's friendliness, ability to make conversations and other aspects that are not related to driving work at all. These factors are crucial to how passengers understand and rate drivers.

Low awareness:
Also, from their past experience with other platforms where 5 star rating scales are deployed, all passengers do not have a uniform understanding of how ratings work. For instance, as in the restaurant rating platform Yelp or in the e-commerce platform Amazon, 4 stars out of 5 may indicate exceptional quality, in Uber, getting average ratings below 4.6 and in Lyft getting below 4.79 can lead to deactivation (and hence the inability to drive for the service at all). In terms of passenger biases, the drivers who participated in the survey said that reasons for low ratings (lower than 5) can be numerous and subjective. These factors are mostly beyond driver control.
For instance, whenever the "surge" feature gets activated in Uber and Lyft because of peak hour traffic or excessive demand, passengers have to pay more for a ride. Given this unpleasant experience and their inability to address this issue, they rate drivers low assuming that drivers represent the company. However, in reality, drivers are only independent contractors with the company. So, a damage to their reputation or deactivation does not directly affect the company's revenue or reputation. Apart from surge

pricing, other aspects like racism, gender bias, preferred route to the destination and arbitrary things like disagreement over sports or movies also contribute to low ratings.

Illegality:

The other big issue is the paradox of the one sided reputation system where passengers essentially have entire control over the ability of drivers to work, forcing them to often indulge in illegal activities in order to maintain their ratings.
One of the biggest complaints that drivers have is the "squeeze" problem or passengers insisting on taking more people than realistically possible and legally permissible. Also, since one of the biggest use cases of Uber is weekend cab services for inebriated passengers, drivers also land in trouble if the refuse to let passengers consume or carry open containers of alcohol inside their vehicles. Eventually, refusing the ride leads to a sharp decline in ratings and taking the ride means indulging in illegal activities.

Driving Minors:

Drivers have also reported that often adult passengers order a ride through their phone for minors. It is only upon their arrival that they realize the problem. In such a scenario, if drivers refuse to accept the ride, their total number of rides rejected goes up and they risk getting low ratings. If they accept the ride, it not only constitutes a legal violation but also places enormous risk on the lives of the minor and driver.

Passenger understanding of reputation:

In the same survey, when asked how well passengers understand the rating system, most drivers responded by saying "not well" or "not at all". They also added that since rating every ride is not mandatory for the passenger, many passengers, if they are in a hurry or sleepy or drunk, forget to or avoid rating. Similarly, for passengers who did not have exceptionally good or bad experiences, there is a lower incentive to rate. This has been proven in other reputation systems as well. Thus, there is a higher chance of outliers who give extreme ratings because of very good or bad memorable experiences. What is worse is that passengers who give exceptionally low ratings don't have to explain the rationale behind their actions. Drivers only get anonymized rating reports at the end of the day. This makes it difficult to seek feedback on, improve their service or understand what went wrong.

Effectiveness of ratings as reputation measure

When asked to rate the effectiveness of ratings in terms of promoting trust within ridesharing businesses, 47% drivers rated them as 1 (implying very poor), 18% rated them as 2 (poor), 20% rated them as 3 (average) and the rest 17% rated them as 4 or 5 (implying good or very good).
The issue that surfaces from the biases listed above is that Uber and Lyft ratings are failing to produce a reliable measurement of the actual quality of driving. Since an integral part of services like Uber and Lyft is to provide a pleasant social experience along with the basic service of transportation, the reputation allows customers to subjectively rate based on their "entire experience" without defining or even providing guidelines about what to base their ratings on.

If reputation systems are in place to reinforce and build trust for both the parties involved, the sharing economy model at large poses a grave problem - here there are 3, not 2 parties involved. So, while for the passenger, rating drivers amounts to rating Uber service, if a driver gets deactivated or gets low ratings, due to the sheer amount of available drivers and scale of Uber's operations, the loss of drivers is only their own and not of the company's. In that sense, the division of responsibility and control over one's engagement in ridesharing platforms is not equal. While the companies retain the maximum amount of control over algorithms, fares and disputed matters and passengers do not share any liability, drivers experience the maximum lack in control over their situation.
The sharing economy platforms, like Uber and Lyft in this case, have no means of verifying if customer ratings are "fair". They also do not actively mediate to average the ratings or set a minimum low ratings threshold in order to compensate for the skewed rating graph.

Suggestions:

- Currently passengers are being rated as well. However, it is unclear what impact those ratings have on encouraging better passenger compliance. Even so, low rated passengers still get rides and they have the option of choosing between multiple services.
- Moreover, there is an enormous lack of transparency in how the current rating system works in ridesharing services. We request the consultation to seek clarity on the effectiveness and purpose of the existing reputation design system.
- We also suggest that if loss of reputation can jeopardize employment for over 160,000 ridesharing drivers in the United States, companies like Uber and Lyft should redesign their applications to actively seek and encourage feedback on rides.
- Further, the biggest shortcoming of the existing system is the lack of tooltips, suggestive text or tutorials to inform passengers about how ridesharing works. We suggest that producing such guidelines will drastically improve driver-passenger communication and management of expectations, thus promoting more trust in the system.
- Similarly, since companies also control all the technological components of these applications, they should be tasked with mediating extremely high and low ratings to mitigate their effect on overall ratings.

Raw responses from the survey can be accessed here:
https://docs.google.com/forms/d/1D8Gse2_YkFZLIAeiFrXhNfDusIS3JhWIXRIRTosVi3U/viewanalytics

We look forward to the consultation process and request that these problems be seriously considered to not only improve the work experience within ridesharing but to also promote safety, trust and mutual respect between passengers and drivers.

For further queries, please contact .
Thank you,
Best,

Noopur Raval