

# Analyzing Uber's Ride-sharing Economy

**Farshad Kooti**  
USC / Facebook

**Mihajlo Grbovic**  
Yahoo Research

**Luca Maria Aiello**  
Bell Labs

**Nemanja Djuric**  
Yahoo Research

**Vladan Radosavljevic**  
Yahoo Research

**Kristina Lerman**  
USC - ISI

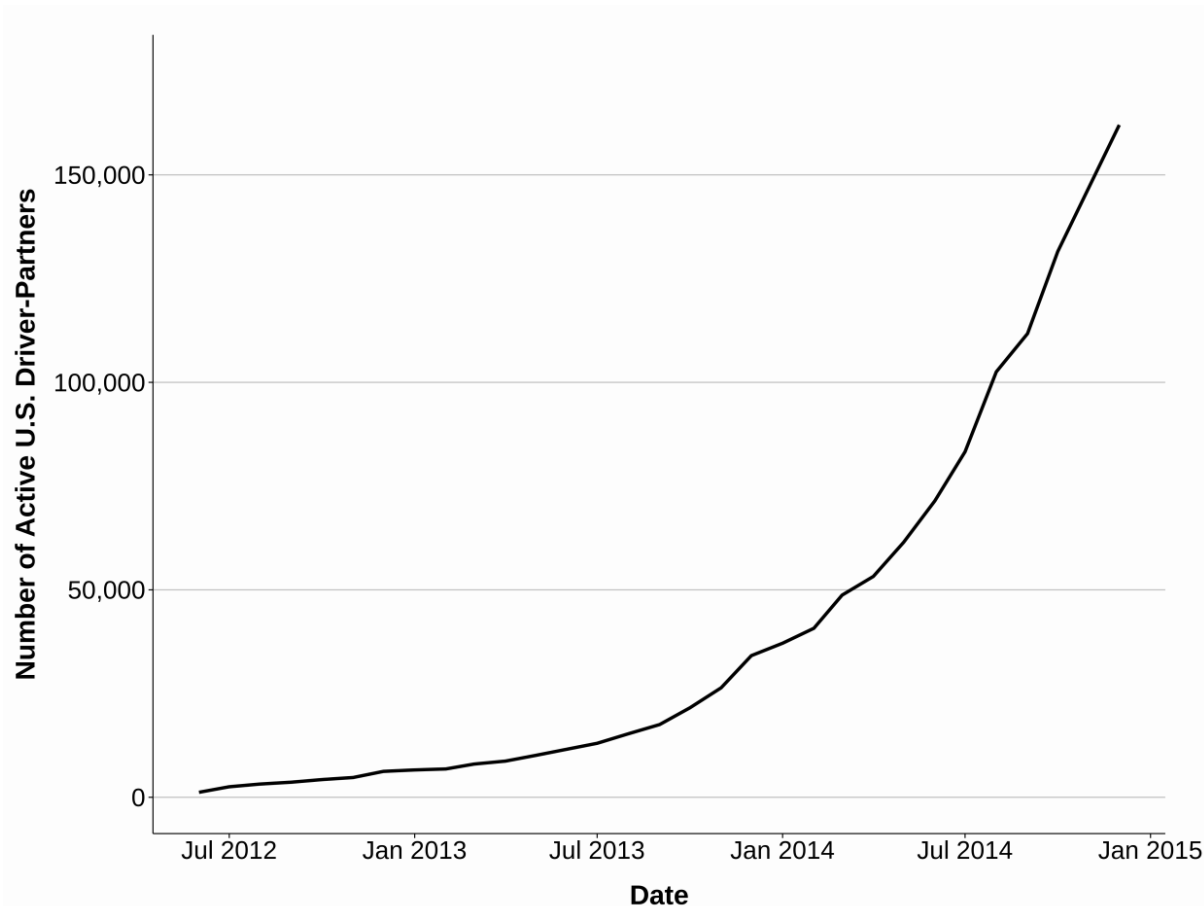


**YAHOO!**



**NOKIA** Bell Labs

# Popularity of ride-sharing



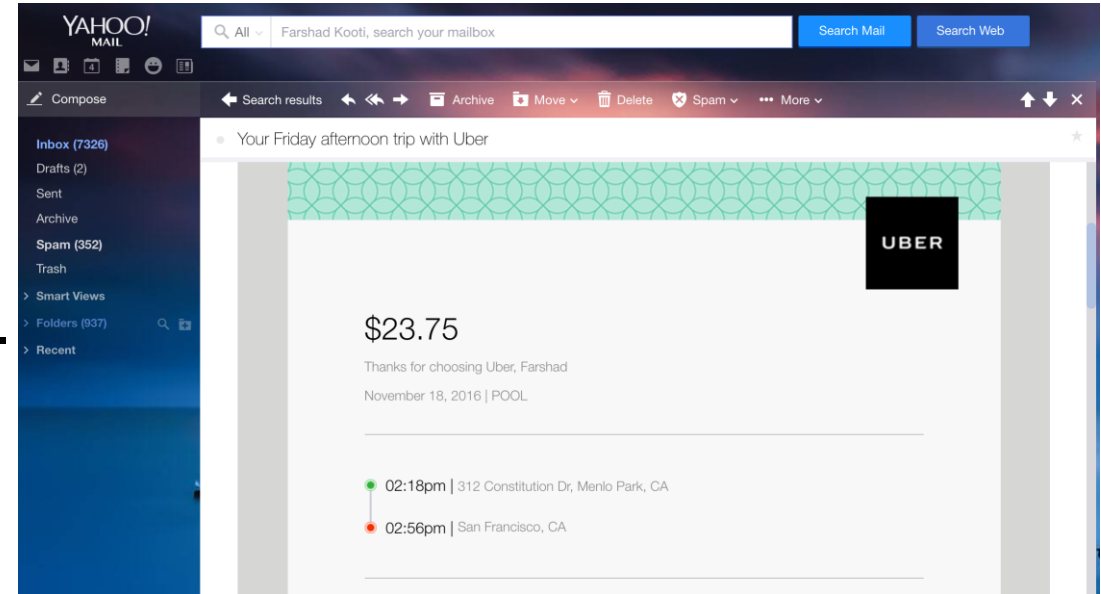
Number of riders and drivers is increasing exponentially [Hall and Krueger, 2015].

# Related Work

- Impact of sharing economy on traditional markets, in particular Airbnb on hotel industry [Zervas et al'15].
- Survey-based studies:
  - 15% of population use ride-sharing apps [Pew'16].
  - Ease of payment and shorter wait time are main reasons people prefer Uber over taxis [Rayle et al'16].
- Uber drivers are more similar to US workforce than taxi drivers [Hall and Kreuger'15]

# Dataset

- All activity on Uber
  - # of rides, length, duration, surge, etc.
- Extracted from confirmation emails
- Oct 2015–May 2016 (8 months)
- 59M rides, 4.1M riders, 222K drivers
- Demographics: age, gender, zip code

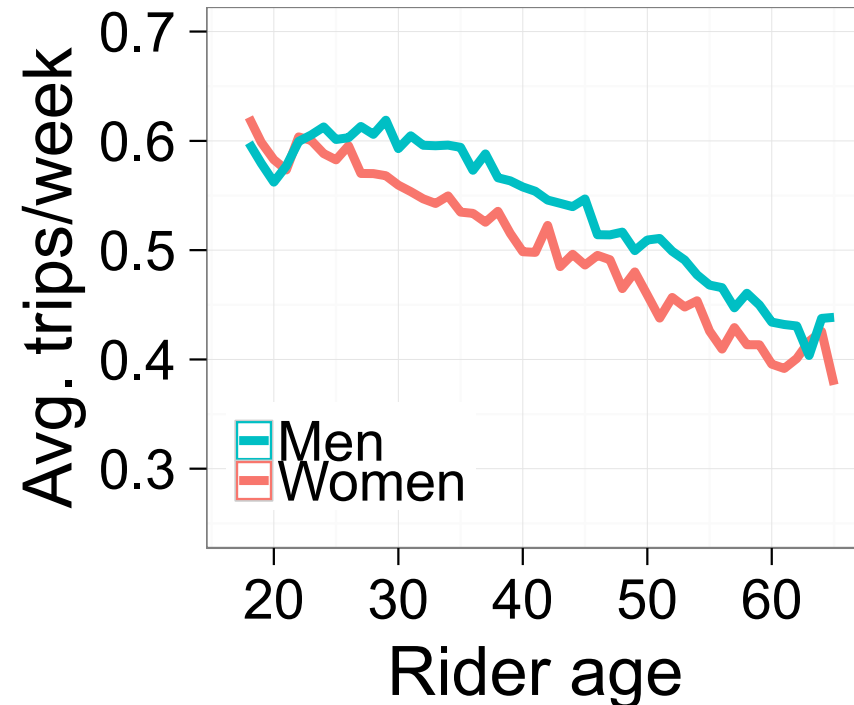




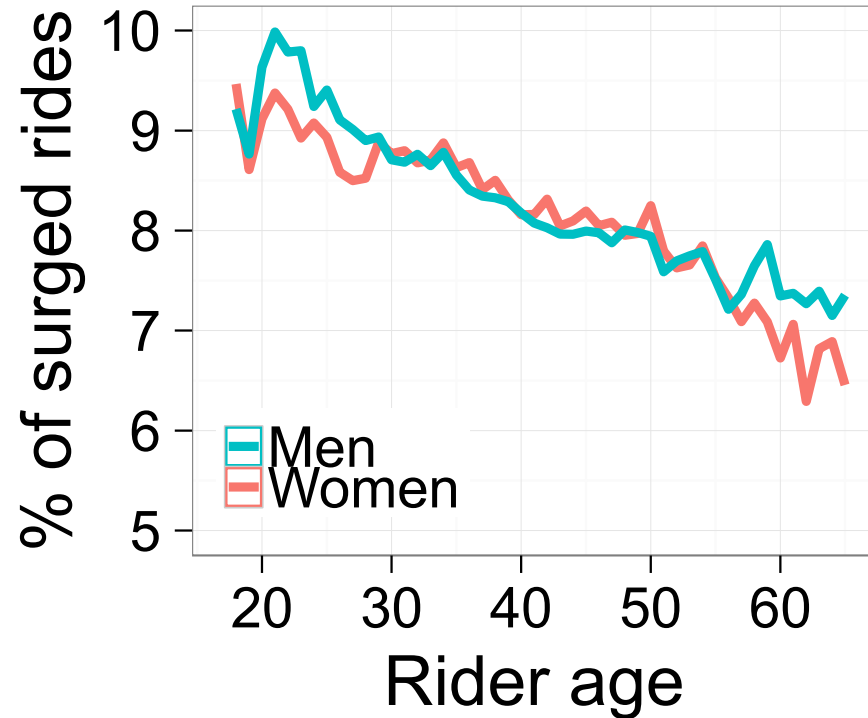
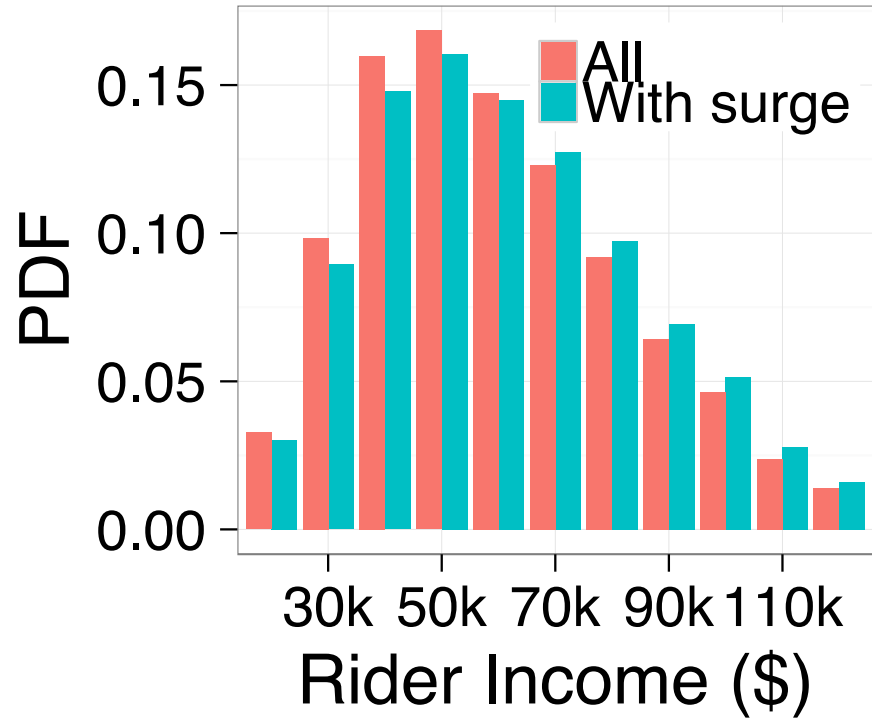
Riders

# Rider Demographics

- Men and younger riders are more active
- By age:
  - Average age of 34yrs, younger than US population
- By gender:
  - 51% women, 49% men
- By race:
  - White: 80.5% vs 72.4% (US)
  - Hispanic: 8.5% vs 16.3%
  - Afr.-Amr.: 8.2% vs 12.6%
  - Asian: 2.8% vs 4.8%



# Surge Pricing



Riders charged surged price are slightly more affluent and younger.

# Rider Attrition

- Know rider's first ride (received welcome email).
- Created monthly activity vector for riders with 4+ months data and 1+ rides (30k riders).
- Clustered using k-means, with k=3 (best choice based on mean square error).

Clusters	% riders	Month 1	Month 2	Month 3	Month 4
Inactive	90.9%	2.1	0.4	0.4	0.5
Low activity	8.0%	8.5	5.8	5.6	5.6
High activity	1.1%	18.0	21.6	23.3	22.1

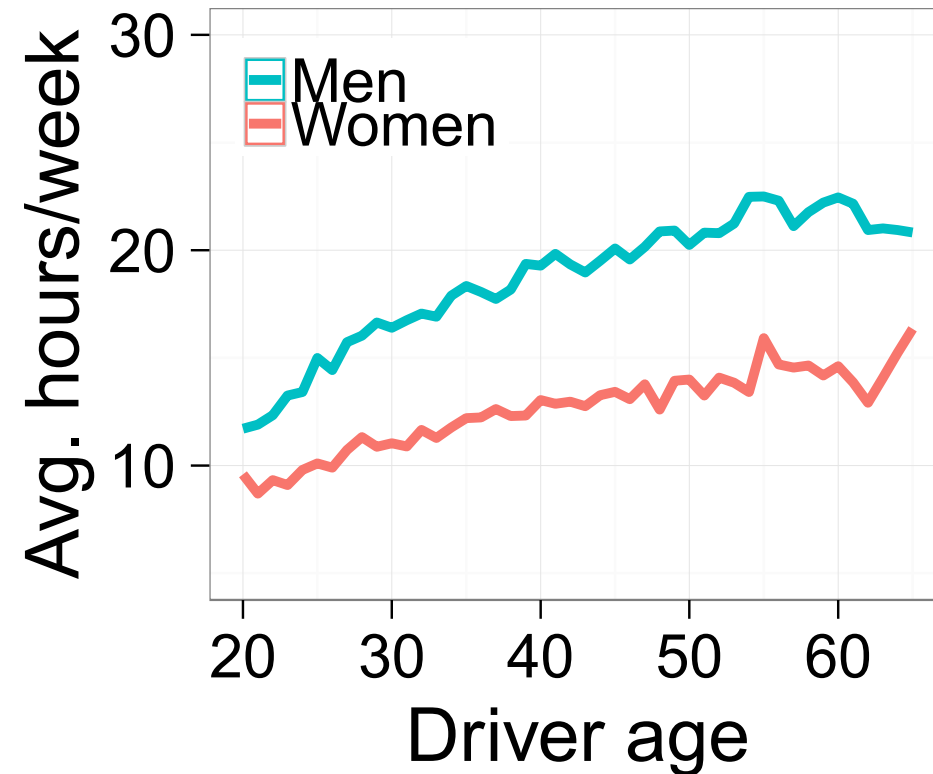


# Driver Demographics: Not an all-serve-all economy

## Compared to riders:

- Drivers are **less likely to be female**
  - 51% (riders) vs 24% (drivers)
- Drivers are **less likely to be white**
  - 81% (riders) vs 60% (drivers)
  - 52% (US taxi drivers), 75% (US workforce)
- Drivers are **less likely to be affluent**
  - median income \$62K (riders) vs \$53K (drivers)

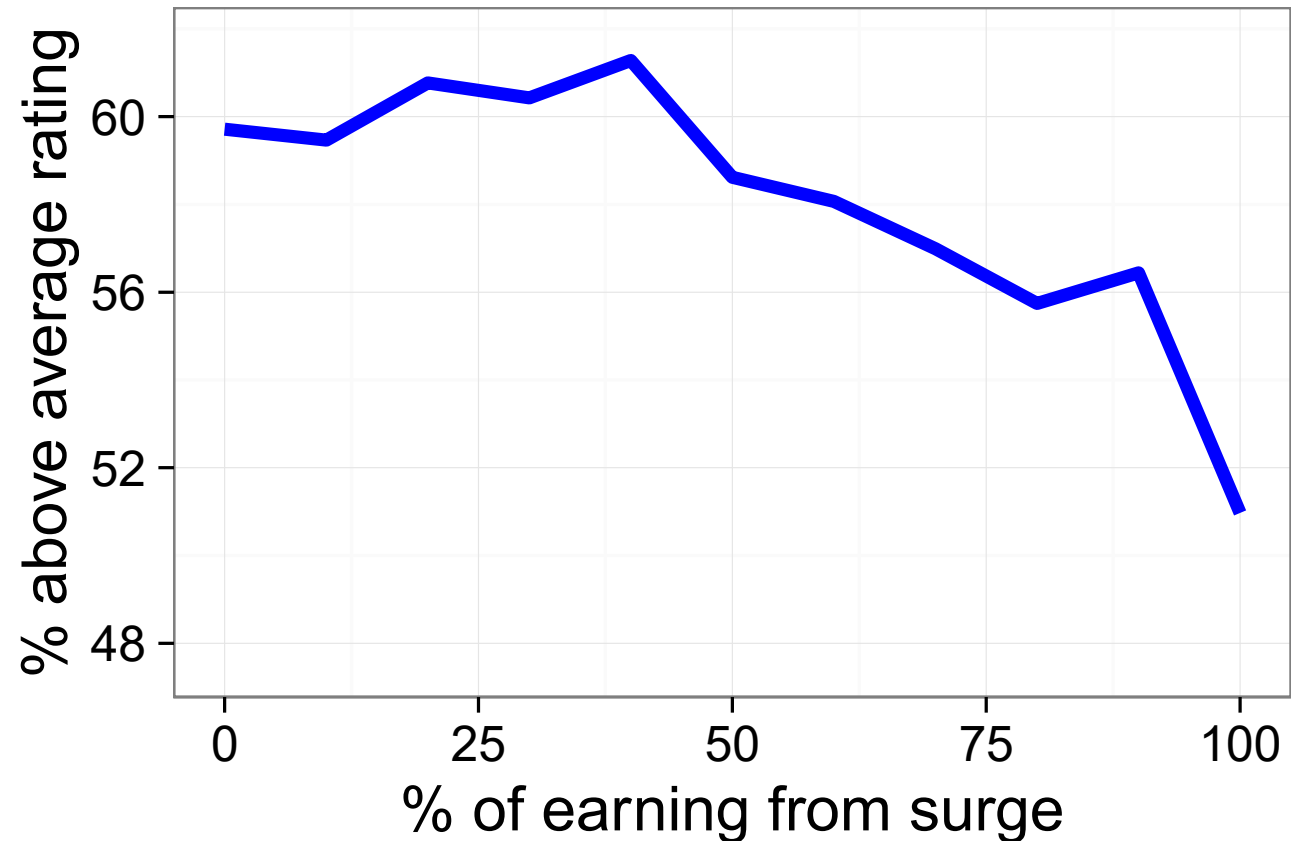
# Hours Worked



Men and older drivers work more, on average. Highly correlated with the earnings.

Only 19% of drivers work 40hrs/week or more.

# Drivers' Ratings and Surge



Surge pricing negatively affects ratings.

# Drivers vs Riders: Matching

- Match riders and drivers using location, time, and cost of rides and first names.
- Consider drivers with at least 10 rides and 75% of rides matched.
- Age: average age difference of 11.4yrs ( $\pm 0.47$ ) for above-average ratings and 13.1yrs ( $\pm 0.56$ ) for below-average ratings.
- Role of gender:

Women Drivers		Men drivers	
% Women Riders	% Above Avg. Weeks	% Men Riders	% Above Avg. Weeks
0%-45%	62.6%	0%-45%	60.2%
45%-55%	53.4%	45%-55%	57.2%
55%-100%	50.2%	55%-100%	61.9%

# Driver Attrition

- Know driver's sign up day (welcome email).
- Cluster based on **monthly hours worked**.
- Drivers have higher retention than riders.

Clusters	% drivers	Month 1	Month 2	Month 3	Month 4
Inactive	73.3%	20.1	4.7	3.0	2.2
Low activity	21.0%	89.9	45.1	16.4	16.4
High activity	5.7%	150.3	133.8	126.8	94.1



Activity Prediction

# Rider Activity

- Predict whether a rider will remain active, based on early activity.
- Features:
  - Rider features: age, gender, etc.
  - Ride features: # of rides, avg. distance, time of rides, etc.
  - Driver features: age, gender, rating, etc.
  - Social features: Number of contacts who are riders or drivers
- Training on 80% of users and balanced classes.

# Rider Activity Results

- Used C5.0 algorithm for classification.
- Used Logistic Regression on independent features for feature ranking.

Prediction Results	
Accuracy	75.2%
Relative Improvement	50.4%
Precision	0.786
Recall	0.687

Top Features	
Total # of rides	0.340***
# cities with rides	0.113***
Gender (men)	0.054***
Average fair	-0.009***
Age	-0.020***



# Driver Activity Prediction

- Use first two weeks to predict if the driver would become active driver in weeks 3-8.
- Similar features and approach as riders

Prediction Results	
Accuracy	83.1%
Relative Improvement	66.2%
Precision	0.775
Recall	0.689

Top Features	
Gender (men)	0.371***
Hours drove	0.157***
Age	0.037***
Earning rate	0.029***
Acceptance rate	-0.015*

# Implications and Summary

- Overall understanding of the fast-growing ecosystem
  - Not an “all-serve-all” economy
- Age, gender, and (maybe) racial bias in ratings
  - People’s awareness
  - Improving the rider-driver matching
- Predictions and analyses help us find users more likely to stop using Uber
  - Promotions
  - Changing the experience

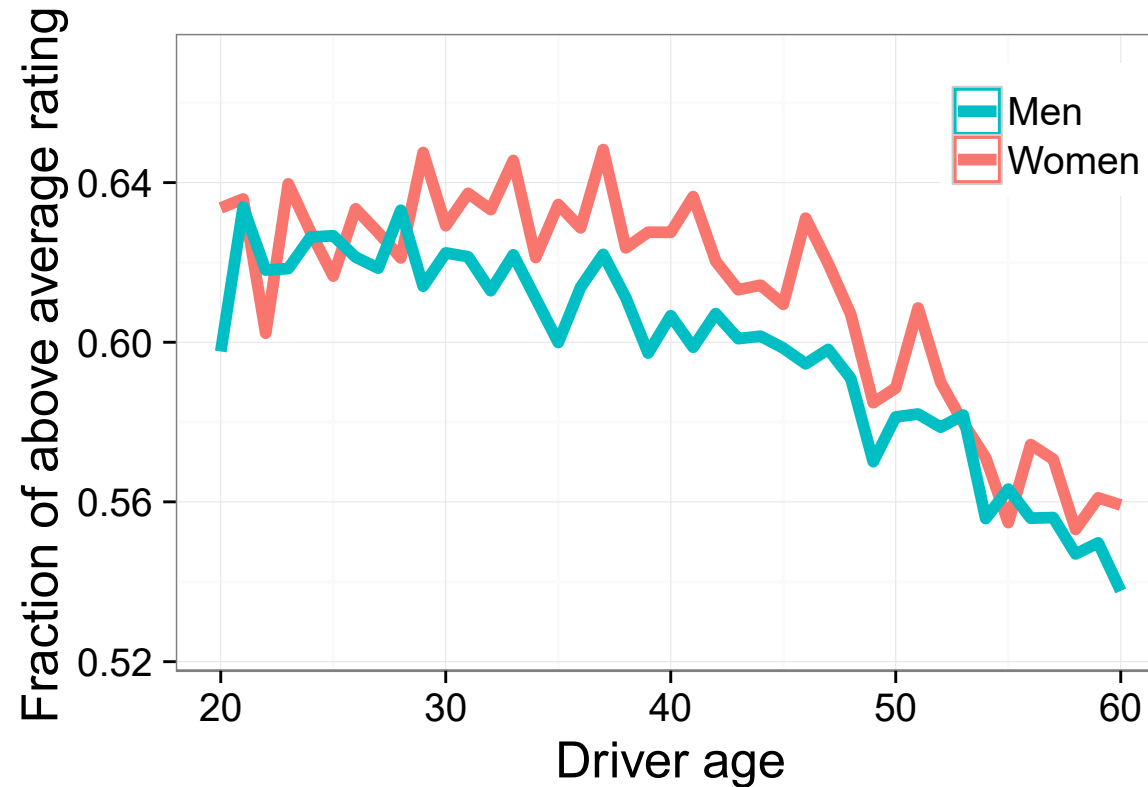
Thanks

?



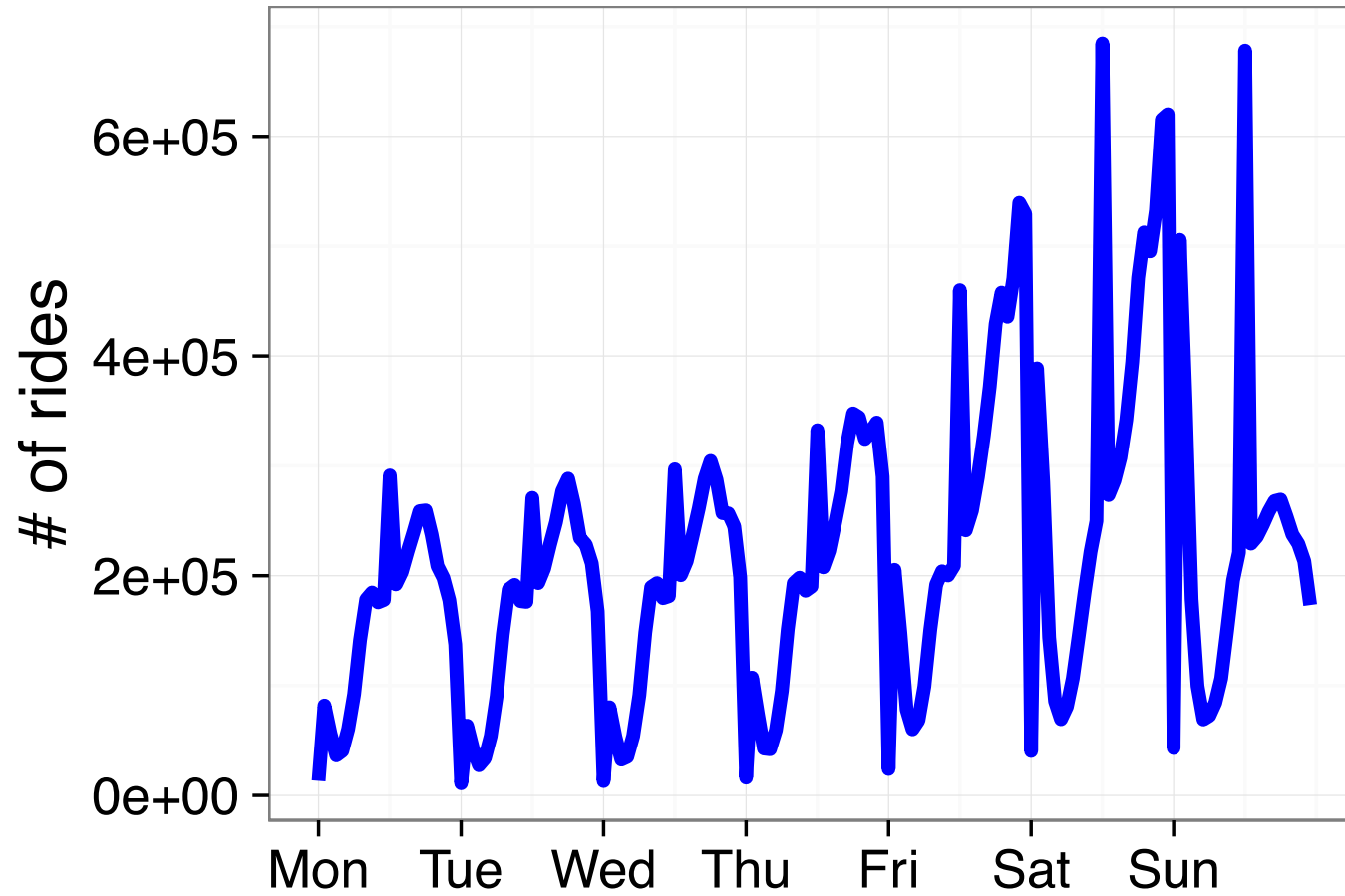
U B E R

# Drivers' Ratings



Younger drivers get better ratings compared to older drivers.

# Ride Dynamics



Clear daily and weekly patterns.

# Summary

- Uber is not an “all-serve-all” market.
- Rider and driver attrition is very high, but the influx of new-comers leads to an overall growth in the number of rides.
- Matching affects rating results.
- Drivers are “punished” for surge rides with lower ratings.
- Based on a rider’s or driver’s initial activity, we can better predict whether she or he will remain active or quit Uber.

# Driver Demographics: Not an all-serve-all economy

## Compared to riders:

- Drivers are less likely to be female
  - US: 24.0%, Malaysia: 10.1%, Singapore: 9.9%, Canada: 9.4%, ..., UK: 4.3% are women
- Drivers are less likely to be white
  - 81% (riders) vs 60% (drivers)
- Drivers are less likely to be affluent
  - median income \$62K (riders) vs \$53K (drivers)

Race	% Drivers	US Taxi Drivers	US Workforce	% Women	Avg. Age	Avg. hrs Worked / week
White	60.0%	52.3%	75.2%	21.9%	41.9yrs	15.4hrs
Afr.-Amr.	21.6%	25.2%	11.6%	36.5%	40.8yrs	14.8hrs
Hispanic	13.7%	10.3%	7.6%	23.9%	38.5yrs	15.2hrs
Asian-Amr.	4.7%	12.2%	5.6%	16.4%	41.6yrs	18.2hrs

# Driver Demographics: Not an all-serve-all economy

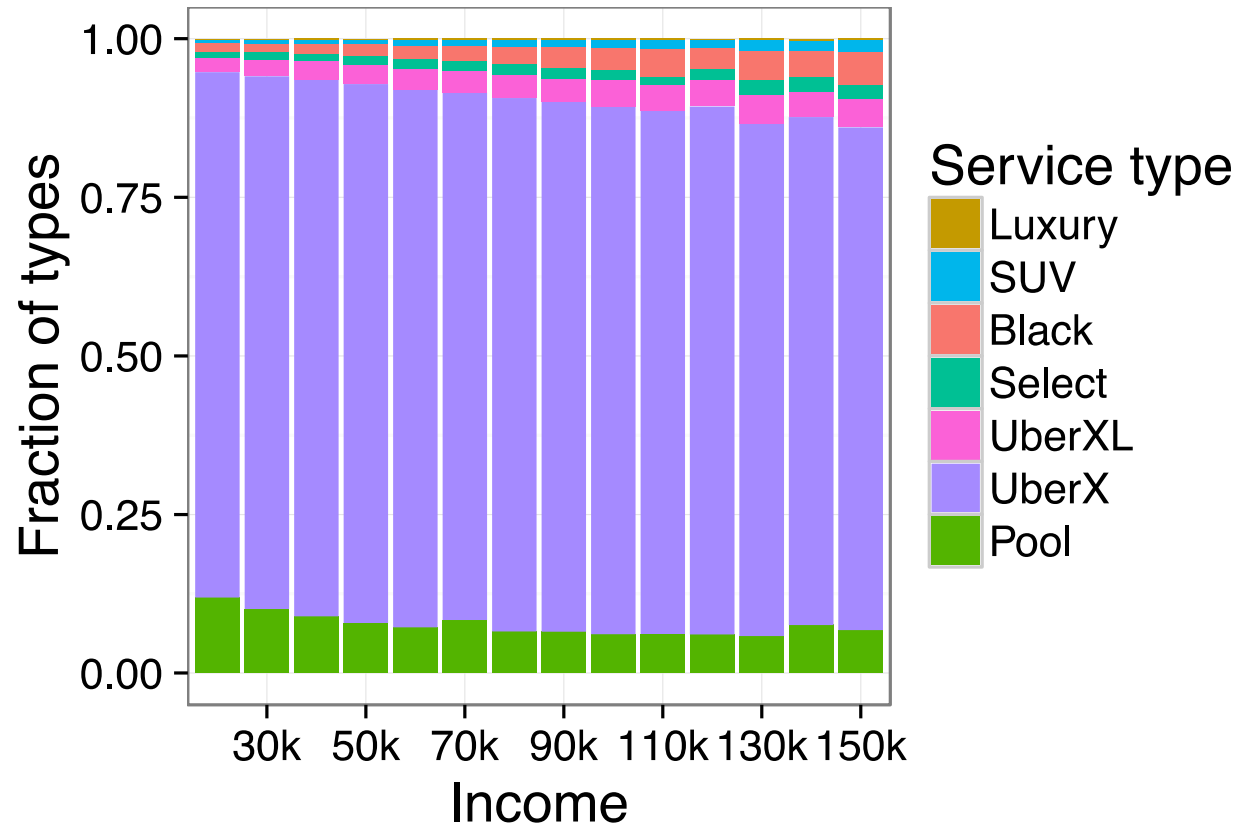
## Compared to riders:

- Drivers are less likely to be female
  - 51% (riders) vs 24% (drivers)
- Drivers are less likely to be affluent
  - median income \$62K (riders) vs \$53K (drivers)
- Drivers are less likely to be white
  - 81% (riders) vs 60% (drivers)
  - 52% (US taxi drivers), 75% (US workforce)

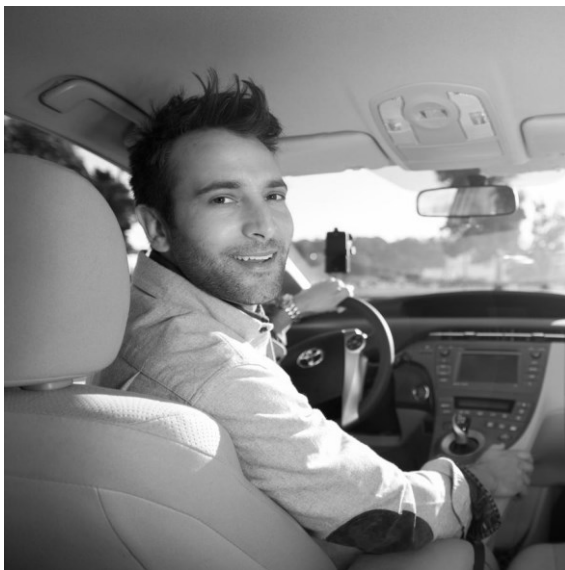
Race	% Drivers	US Taxi Drivers	US Workforce
White	60.0%	52.3%	75.2%
Afr.-Amr.	21.6%	25.2%	11.6%
Hispanic	13.7%	10.3%	7.6%
Asian-Amr.	4.7%	12.2%	5.6%



# Car Type



People with income of \$100k are 84% relatively more likely to take an Uber Black compared to users with annual income of \$50k.



Drivers