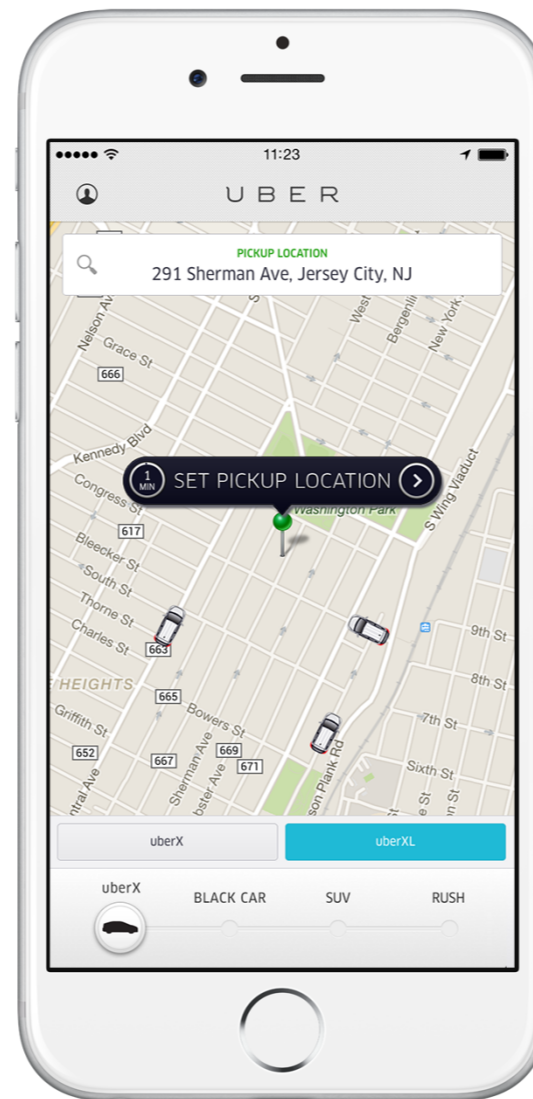


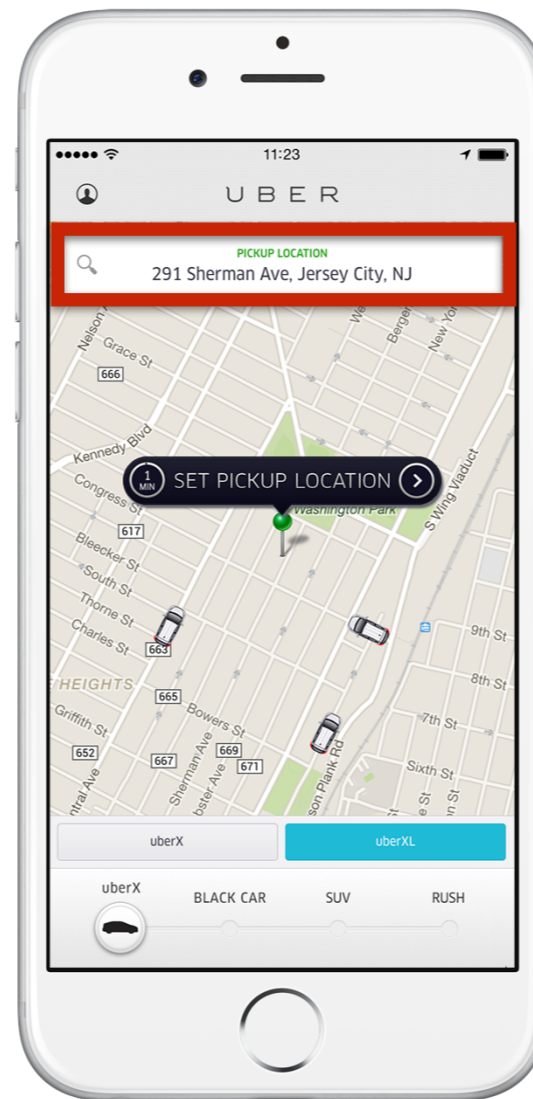
Peeking Beneath the Hood of Uber

Le Chen, Alan Mislove, Christo Wilson
Northeastern University

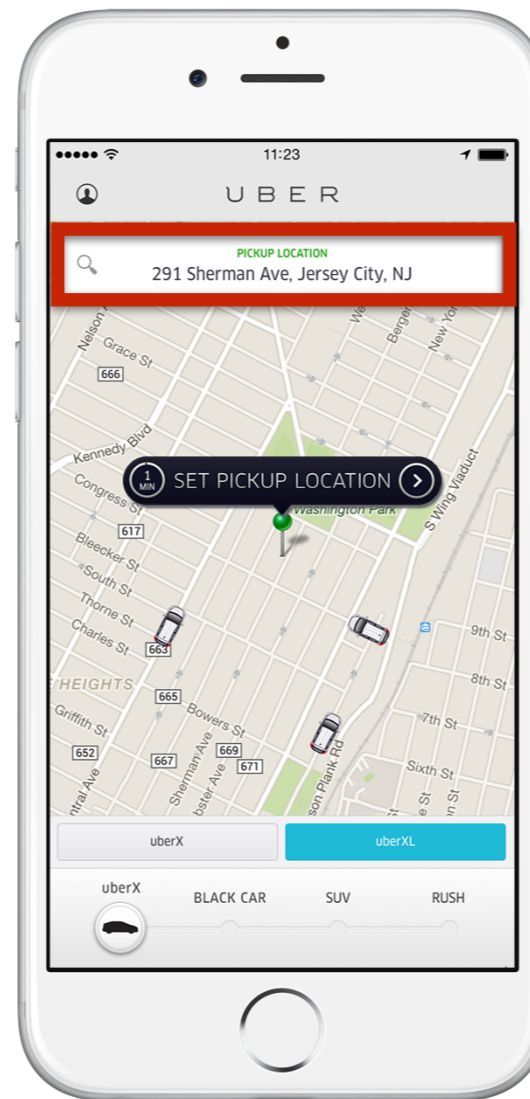
What is Uber?



What is Uber?

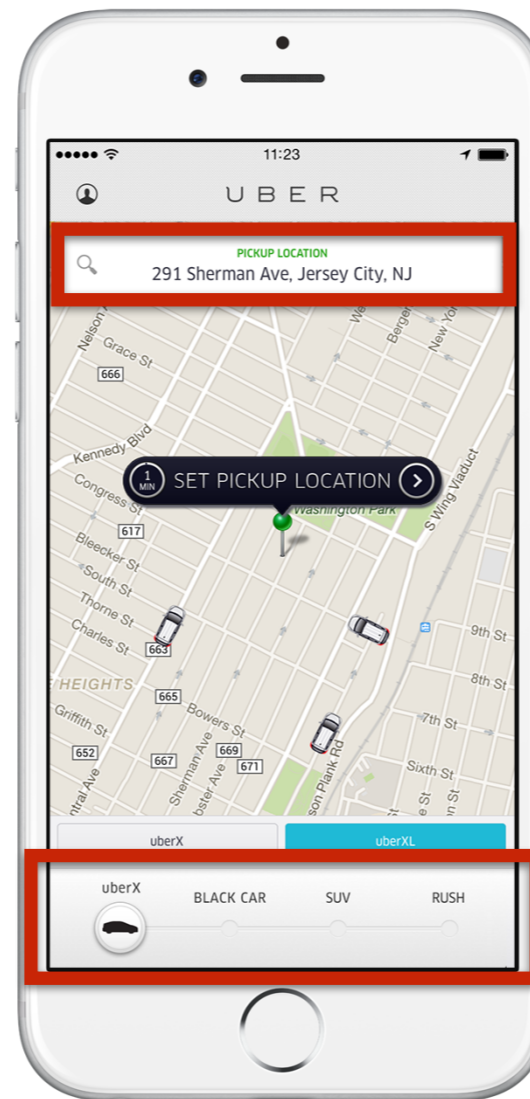


What is Uber?



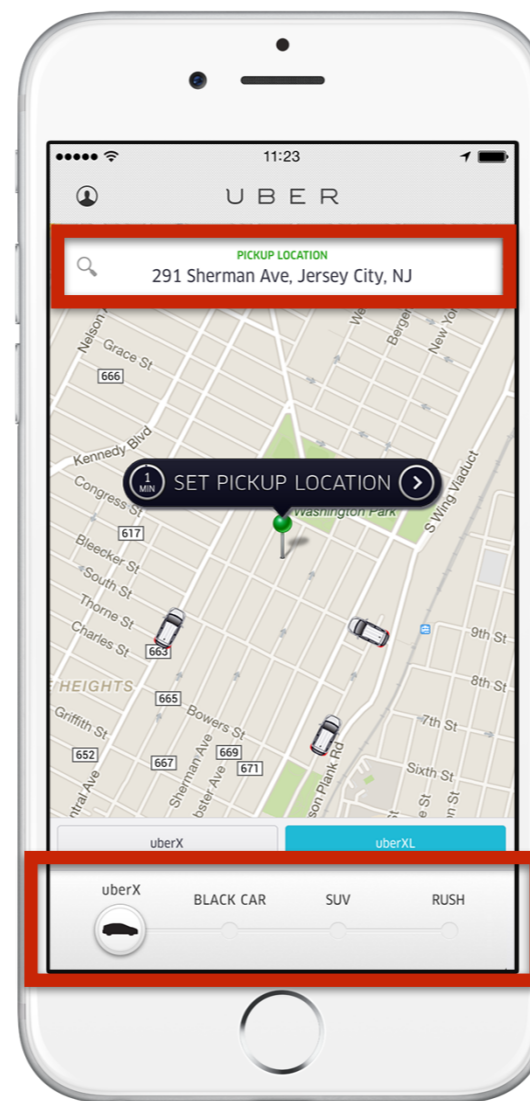
Type pickup location

What is Uber?



Type pickup location

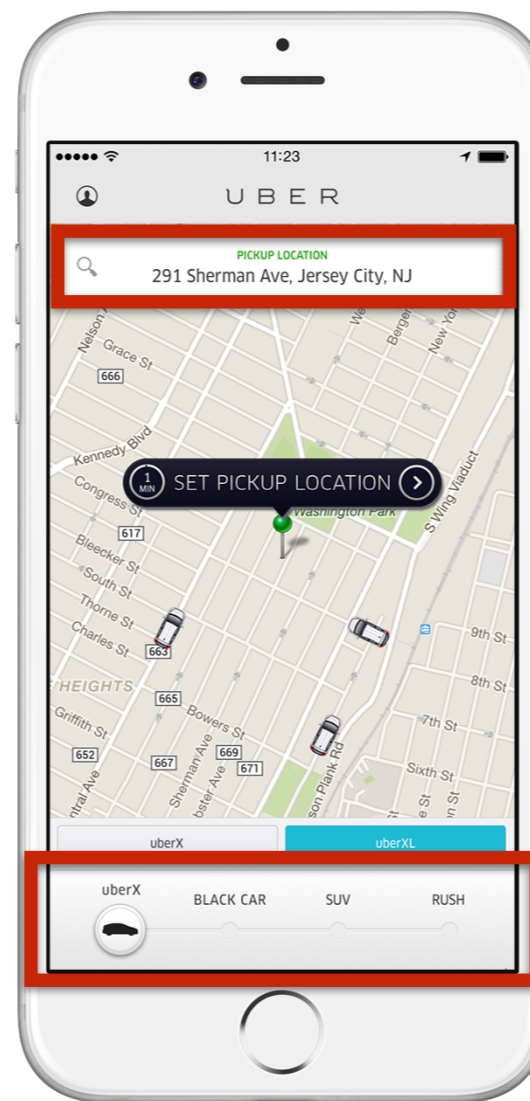
What is Uber?



Type pickup location

Choose type of car

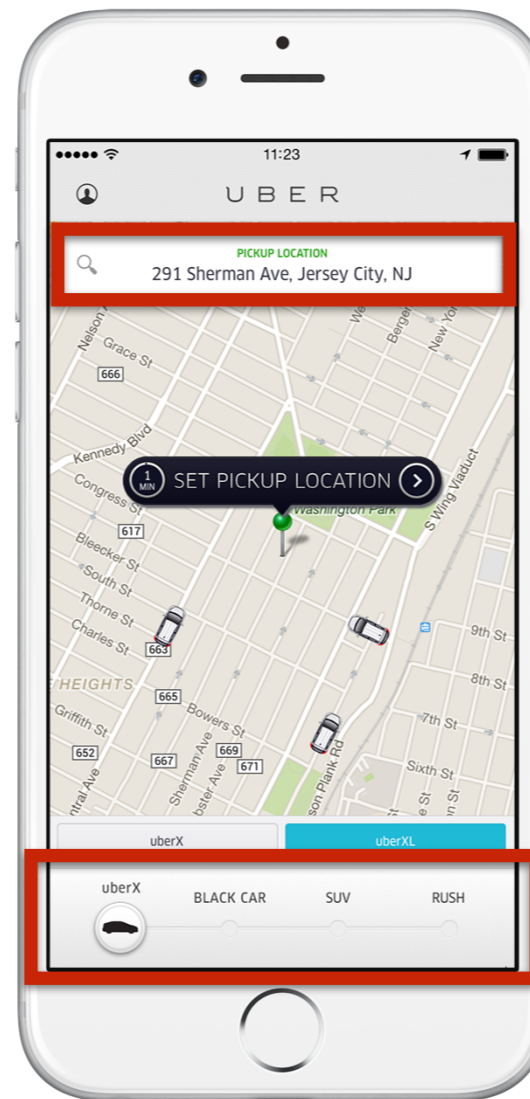
What is Uber?



Type pickup location

Choose type of car
- UberX: basic sedans

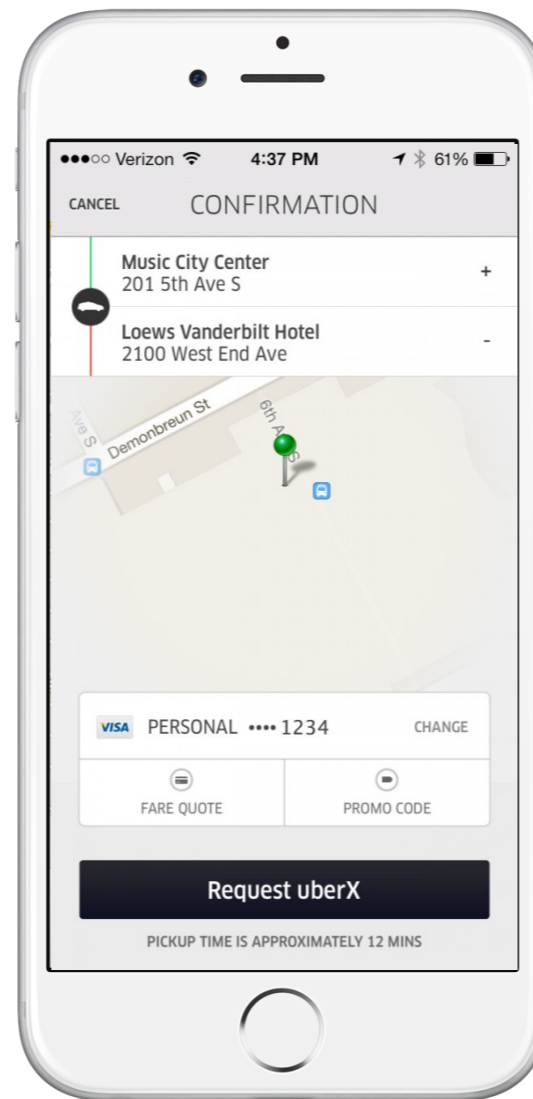
What is Uber?



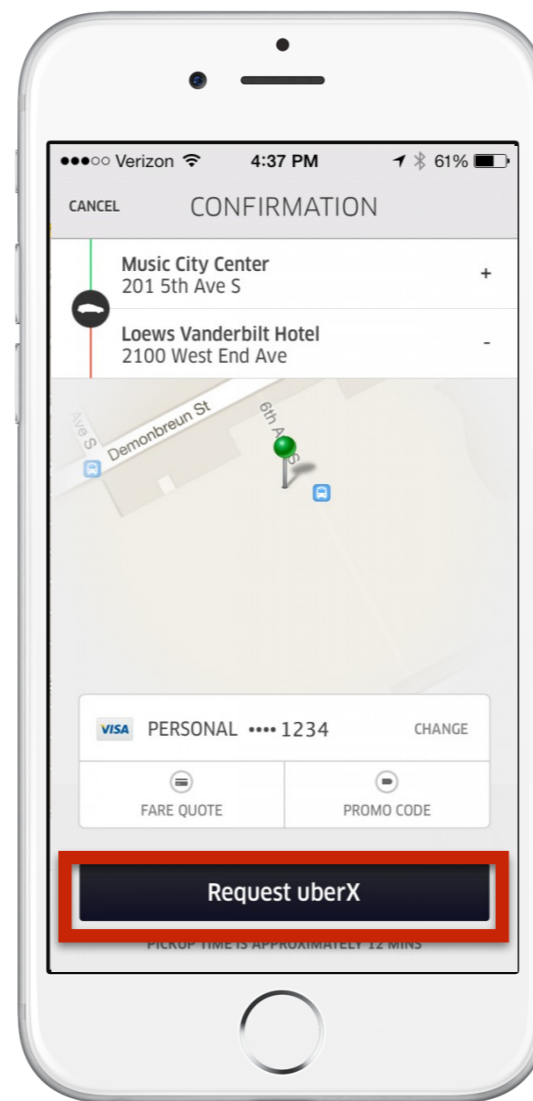
Type pickup location

Choose type of car
- UberX: basic sedans
- UberBlack: luxury cars

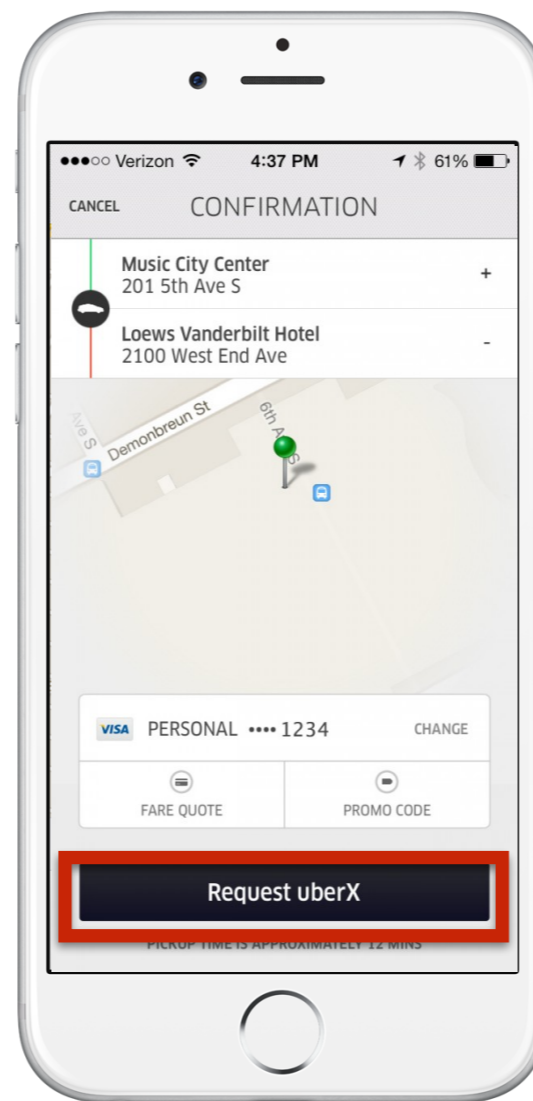
What is Uber?



What is Uber?



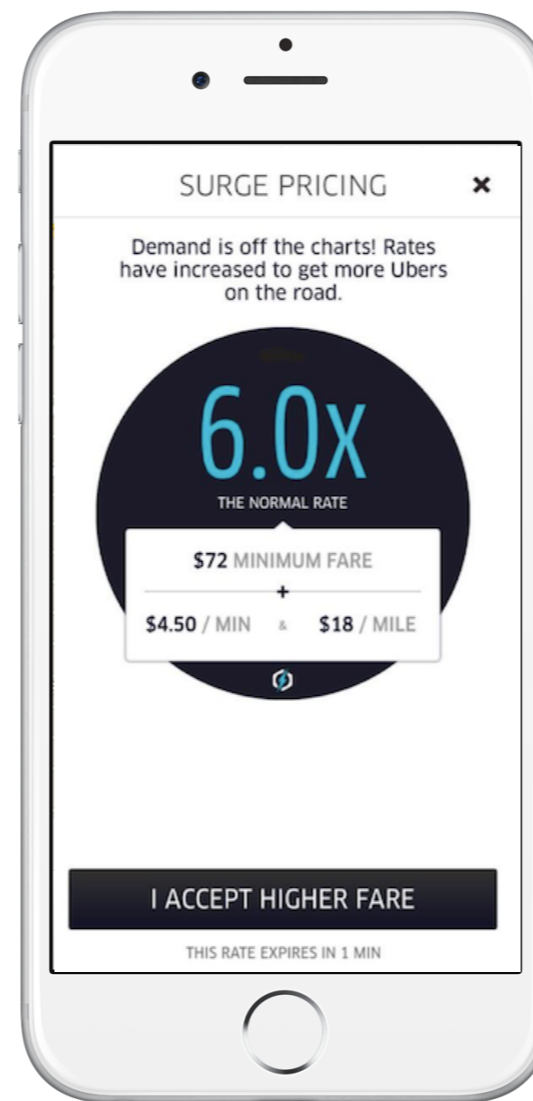
What is Uber?



Request!

Simple and smooth! But ...

Simple and smooth! But ...



Blackbox algorithm

Blackbox algorithm

- Transparent marketplaces

Blackbox algorithm

- Transparent marketplaces



Blackbox algorithm

- Transparent marketplaces

  **Prices are set
by the suppliers**

Blackbox algorithm

- Transparent marketplaces

  **Prices are set by the suppliers**

- Uber is not a transparent marketplace



Blackbox algorithm

- Transparent marketplaces



- Uber is not a transparent marketplace



Prices are dynamically calculated by an algorithm based on supply, demand, etc.

Blackbox algorithm

Blackbox algorithm



HELP

WHAT IS SURGE PRICING?

Blackbox algorithm



WHAT IS SURGE PRICING?

Is Uber Surge Pricing Fair?

© November 3, 2014 News

How to never get slammed with Uber surge pricing again

Detest Uber's surge pricing? Some drivers don't like it either

The practice of tripling, quadrupling and quintupling ride fares in times of high demand may face limits from New York City officials. Many drivers might be OK with that.

Blackbox algorithm



WHAT IS SURGE PRICING?

Is Uber Surge Pricing Fair?

© November 3, 2014 News

How to never get slammed with Uber surge pricing again

Detest Uber's surge pricing? Some drivers don't like it either

The practice of tripling, quadrupling and quintupling ride fares in times of high demand may face limits from New York City officials. Many drivers might be OK with that.

TECH UBER

Uber's surge pricing is good for you, Uber study says

Blackbox algorithm



WHAT IS SURGE PRICING?

Is Uber Surge Pricing Fair?

© November 3, 2014 News

How to never get slammed with Uber surge pricing again

Detest Uber's surge pricing? Some drivers don't like it either

The practice of tripling, quadrupling and quintupling ride fares in times of high demand may face limits from New York City officials. Many drivers might be OK with that.

TECH UBER

Uber's surge pricing is good for you, **Uber study says**

Blackbox algorithm



WHAT IS SURGE PRICING?

Is Uber Surge Pricing Fair?

November 3, 2014 News

How to never
pricing ag

Detes
driven

Popular, but not transparent

The practice of tripling... and may face
limits from New York City

TECH UBER

Uber's surge pricing is good for
you, **Uber study says**

Goal of this study

Goal of this study

- How does the surge pricing algorithm work?

Goal of this study

- How does the surge pricing algorithm work?
 - Verify if it works as Uber claimed

Goal of this study

- How does the surge pricing algorithm work?
 - Verify if it works as Uber claimed
 - Responsiveness of the algorithm

Goal of this study

- How does the surge pricing algorithm work?
 - Verify if it works as Uber claimed
 - Responsiveness of the algorithm
- Can the surge values be predicted?

Goal of this study

- How does the surge pricing algorithm work?
 - Verify if it works as Uber claimed
 - Responsiveness of the algorithm
- Can the surge values be predicted?
- Impact on passengers and drivers

Outline of Uber study

- Motivation
- Data collection
- Surge pricing
- Summary

Data collection

Data collection

- Uber official patent: supply, demand, etc.

Data collection

- Uber official patent: supply, demand, etc.
- Option 1: Uber API

Data collection

- Uber official patent: supply, demand, etc.
- Option 1: Uber API
 - Pros: easy to use; has surge values and EWT

Data collection

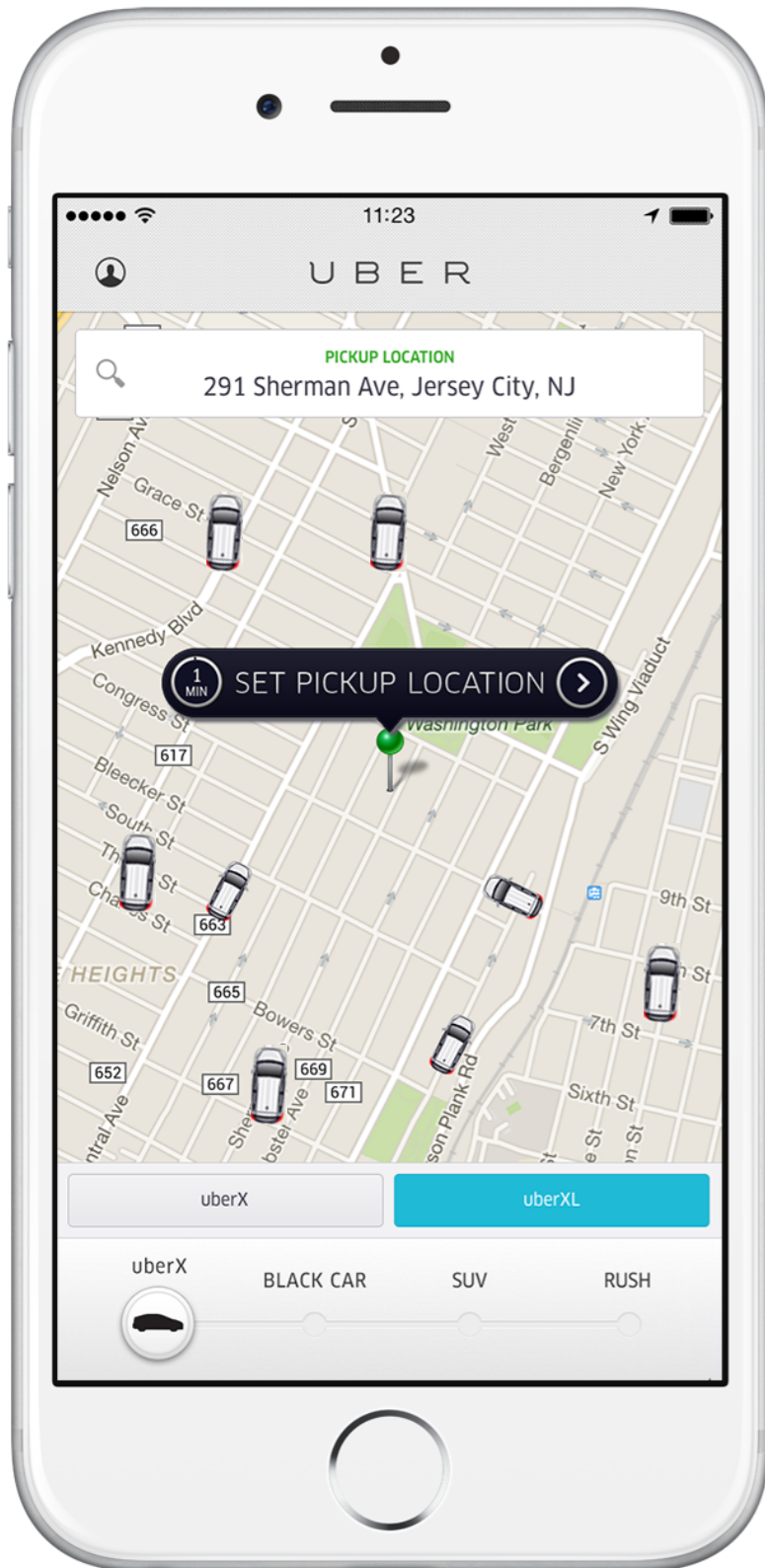
- Uber official patent: supply, demand, etc.
- Option 1: Uber API
 - Pros: easy to use; has surge values and EWT
 - Cons: no car/demand/supply information

Data collection

- Uber official patent: supply, demand, etc.
- Option 1: Uber API
 - Pros: easy to use; has surge values and EWT
 - Cons: no car/demand/supply information
- Option 2: Uber client app

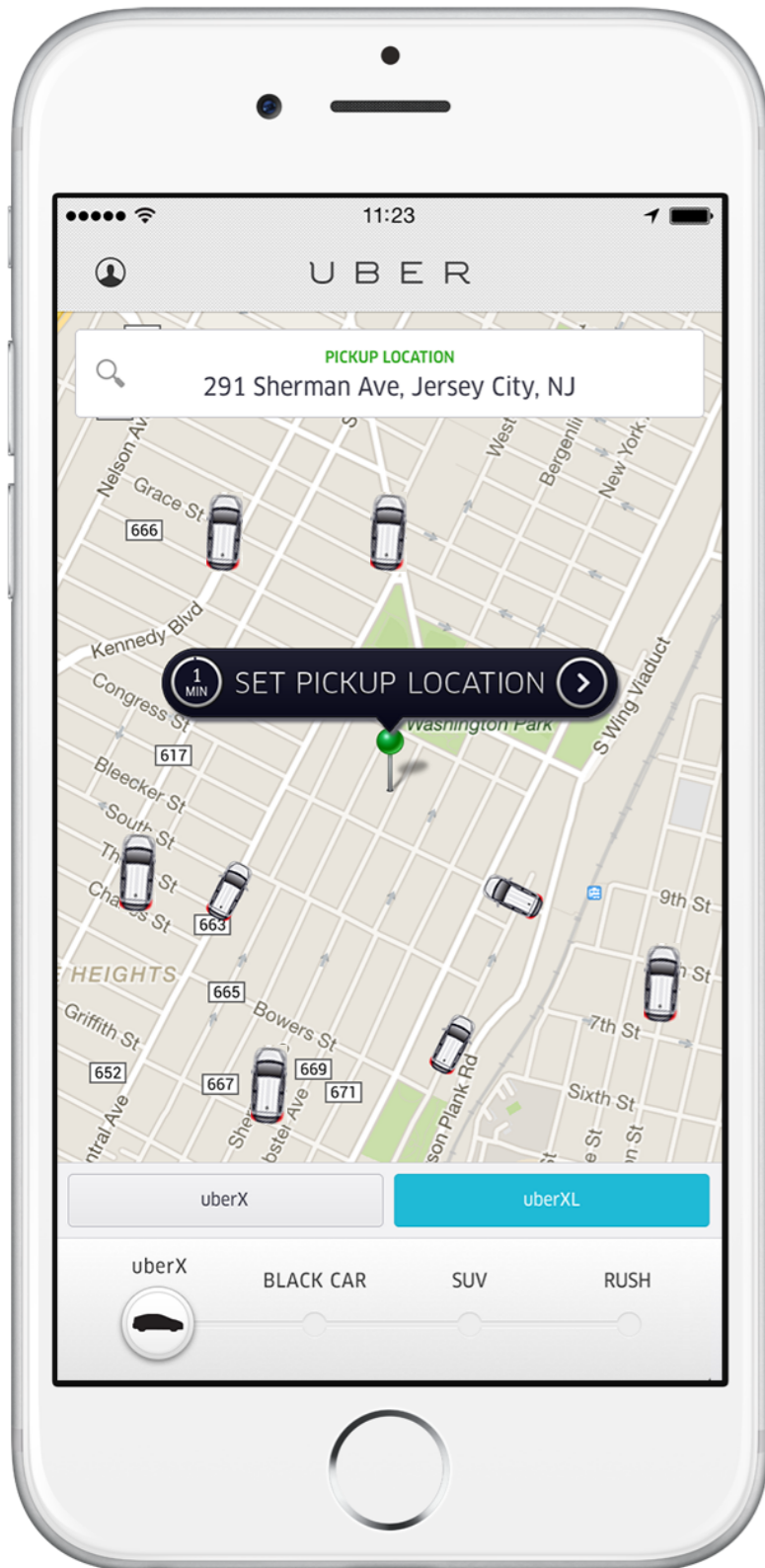
Uber client app

Uber client app



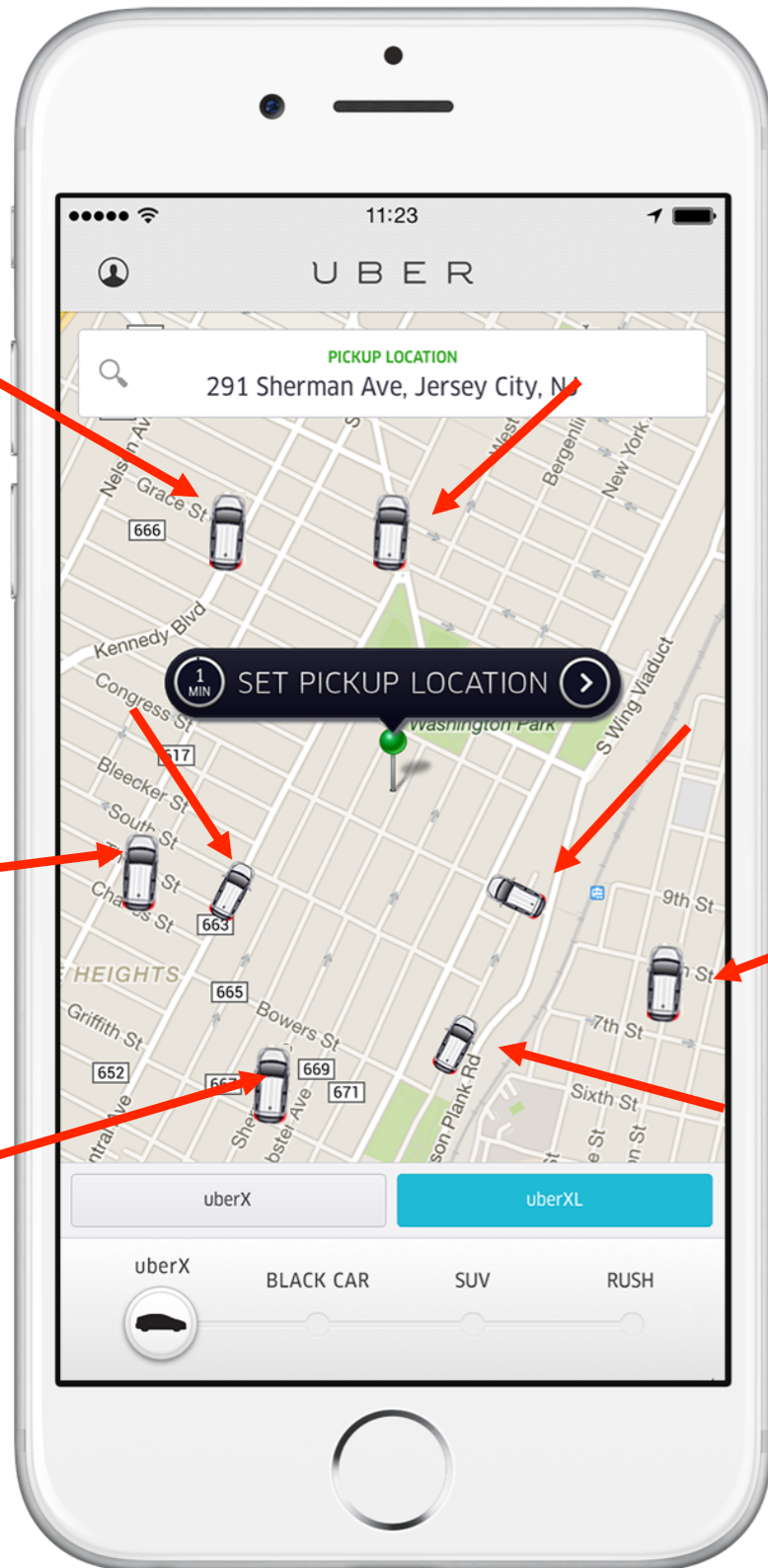
Uber client app

- Pings the server every 5 seconds

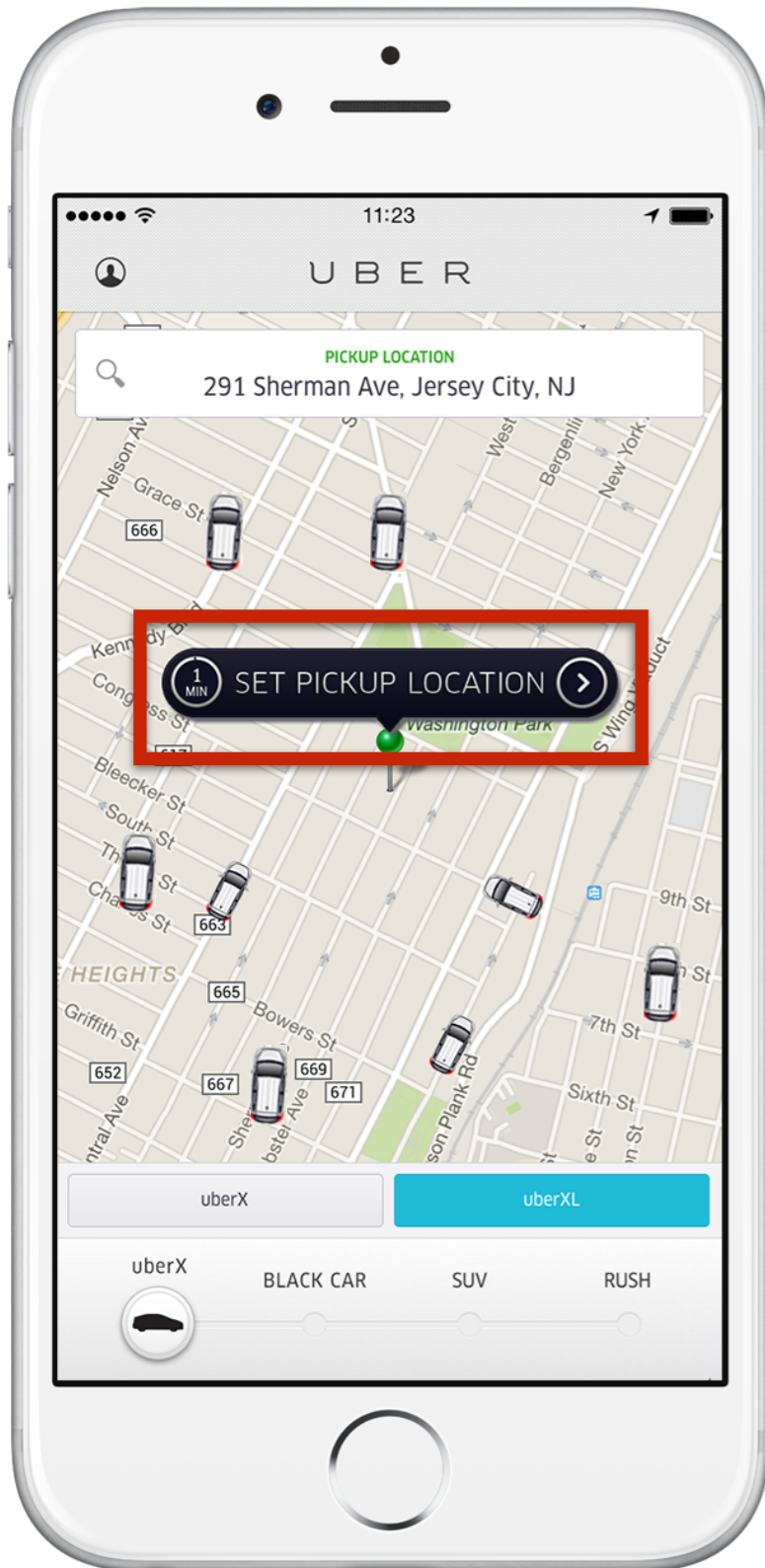


Uber client app

- Pings the server every 5 seconds
- 8 nearest cars

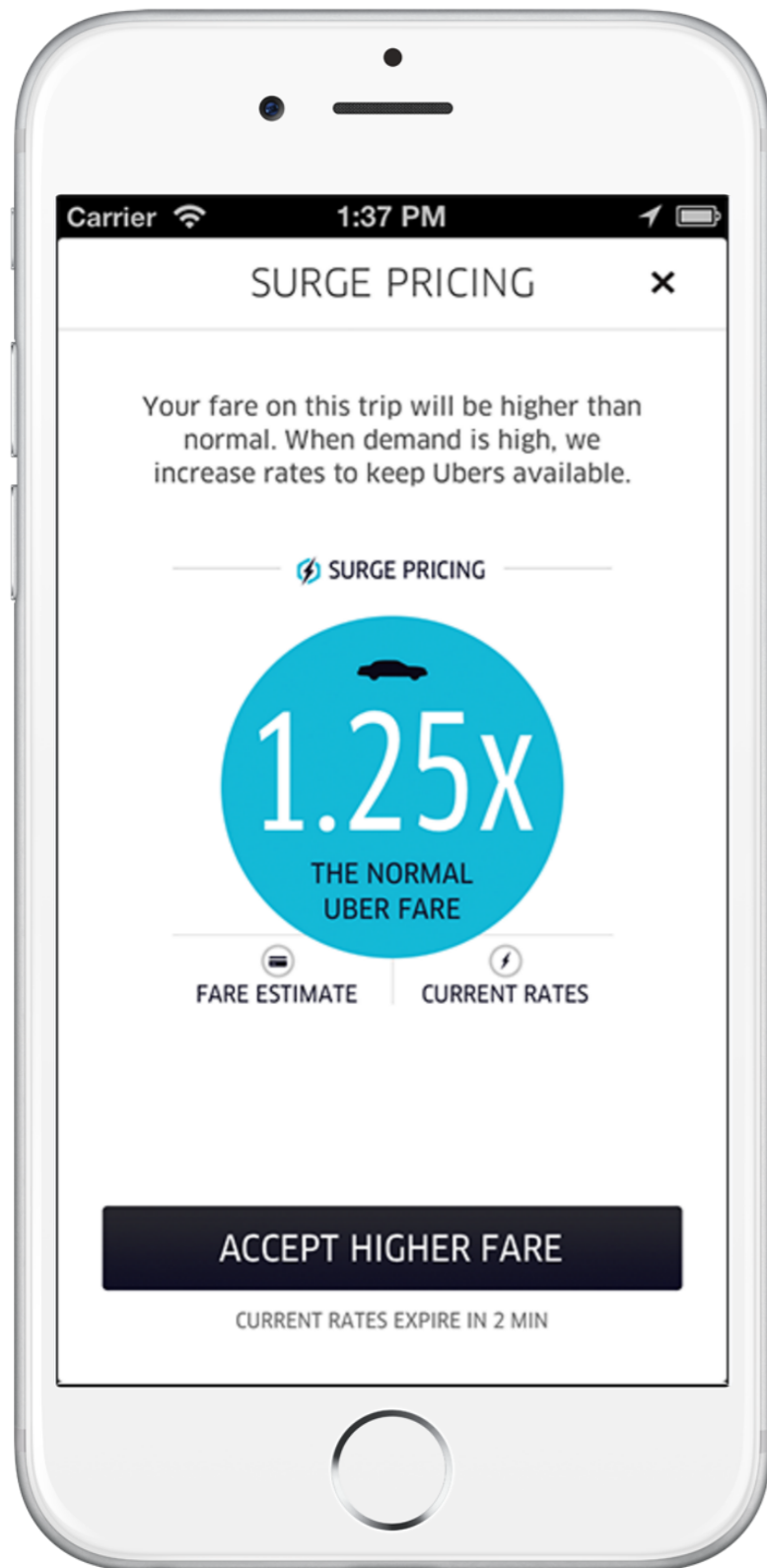


Uber client app



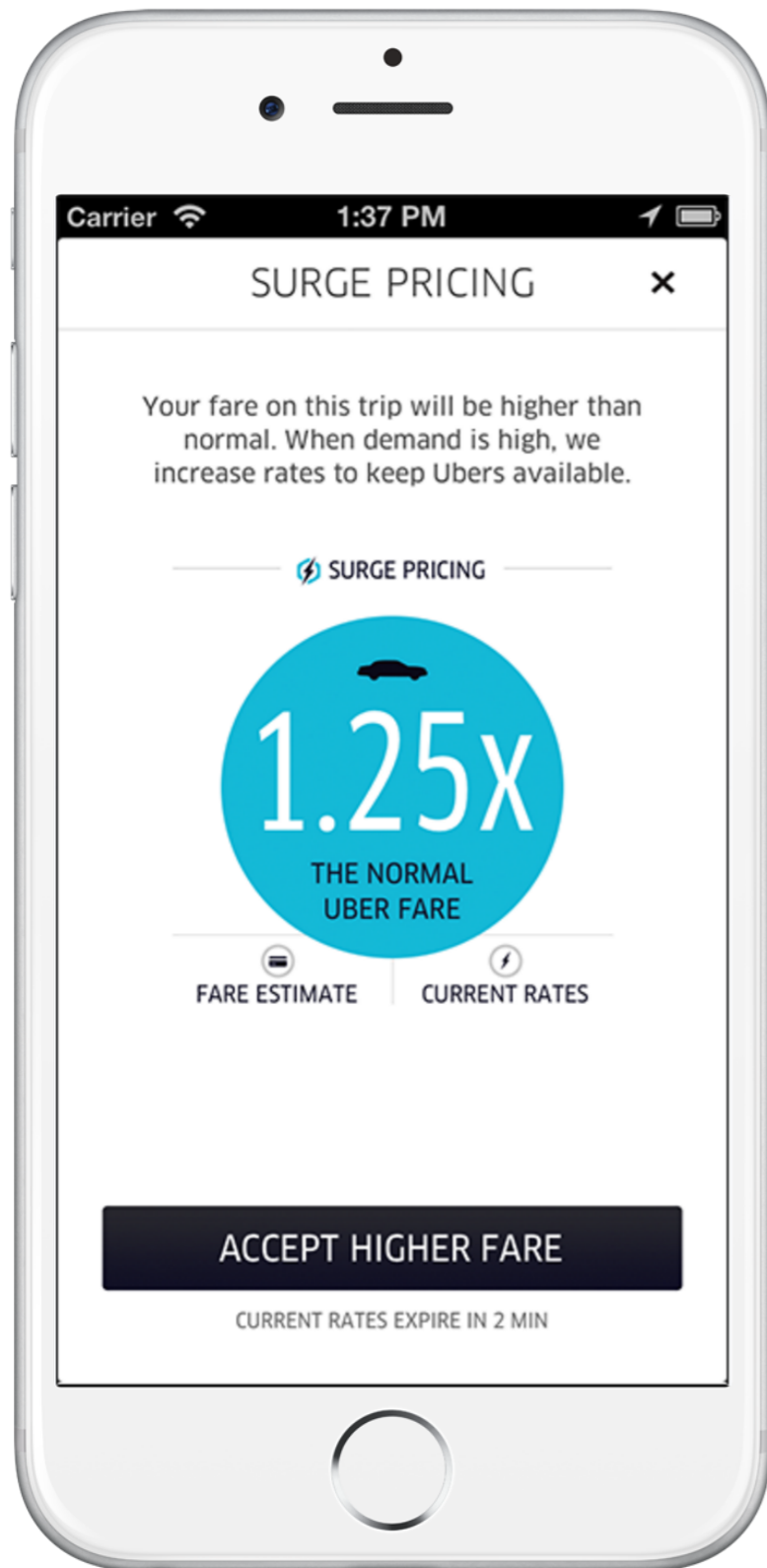
- Pings the server every 5 seconds
- 8 nearest cars
- Estimated Wait Time (EWT)

Uber client app



- Pings the server every 5 seconds
- 8 nearest cars
- Estimated Wait Time (EWT)
- Surge multiplier

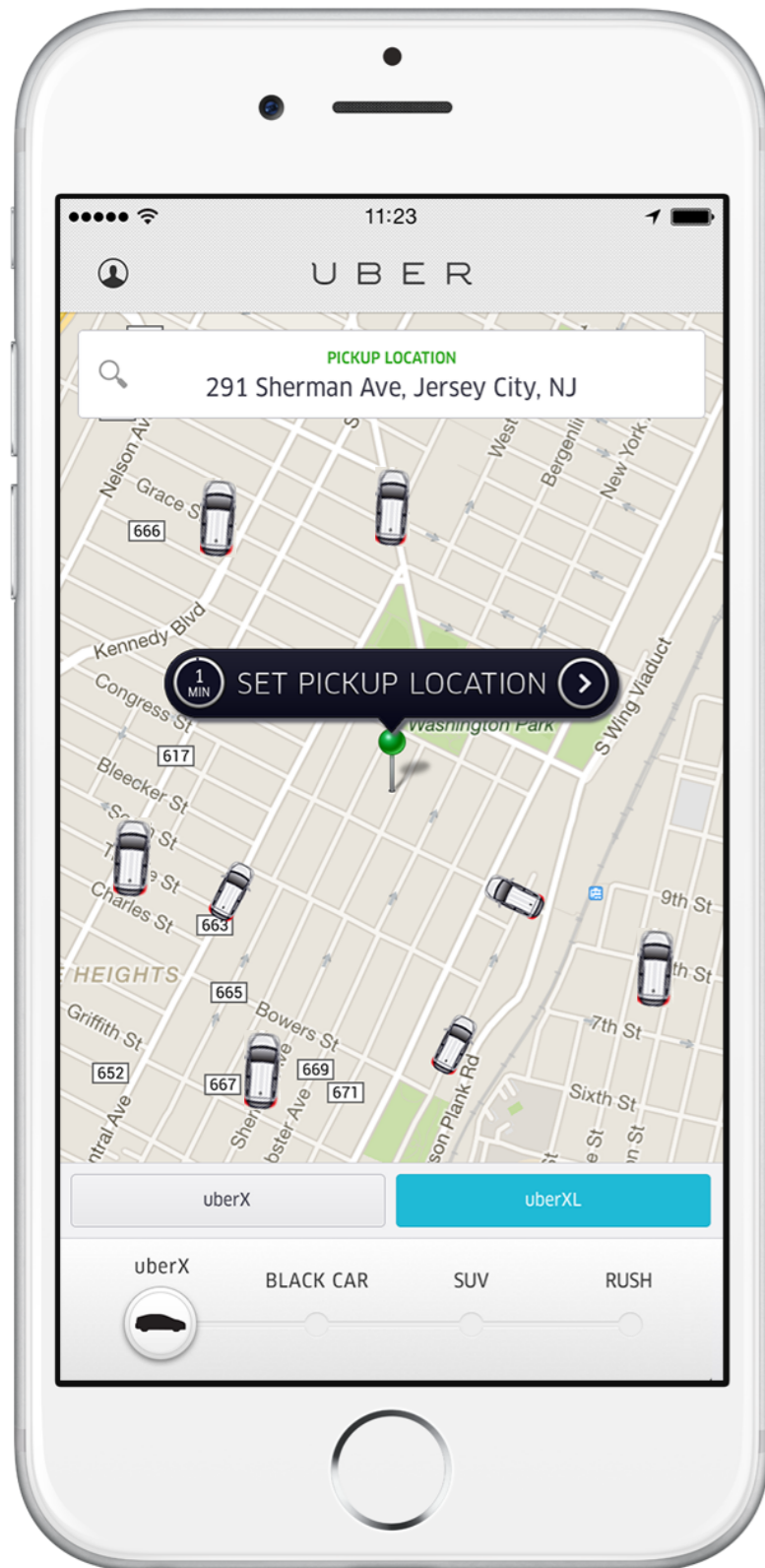
Uber client app



- Pings the server every 5 seconds
- 8 nearest cars
- Estimated Wait Time (EWT)
- Surge multiplier
- More information (supply/demand)

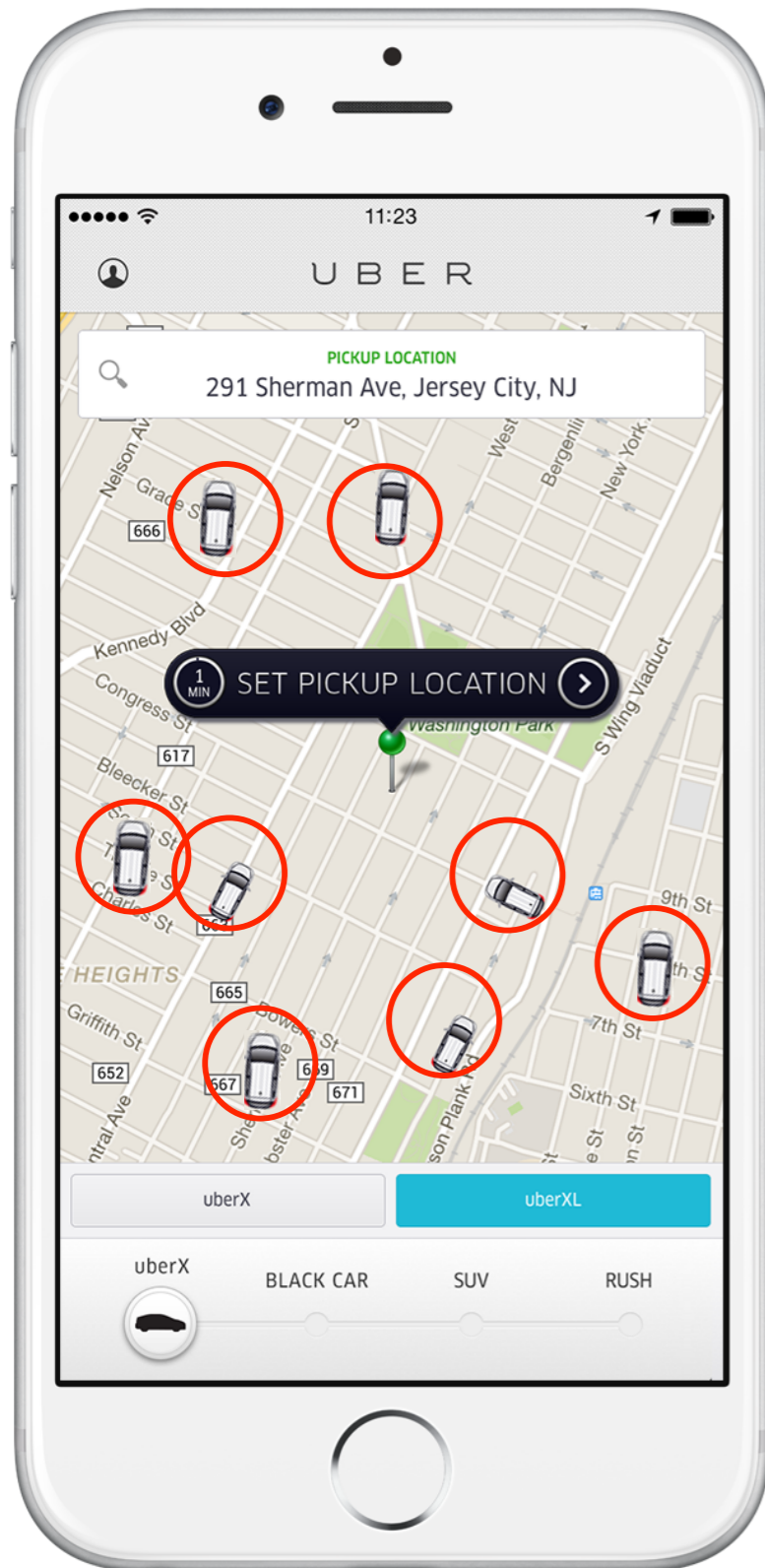
Uber client app limitations

Uber client app limitations



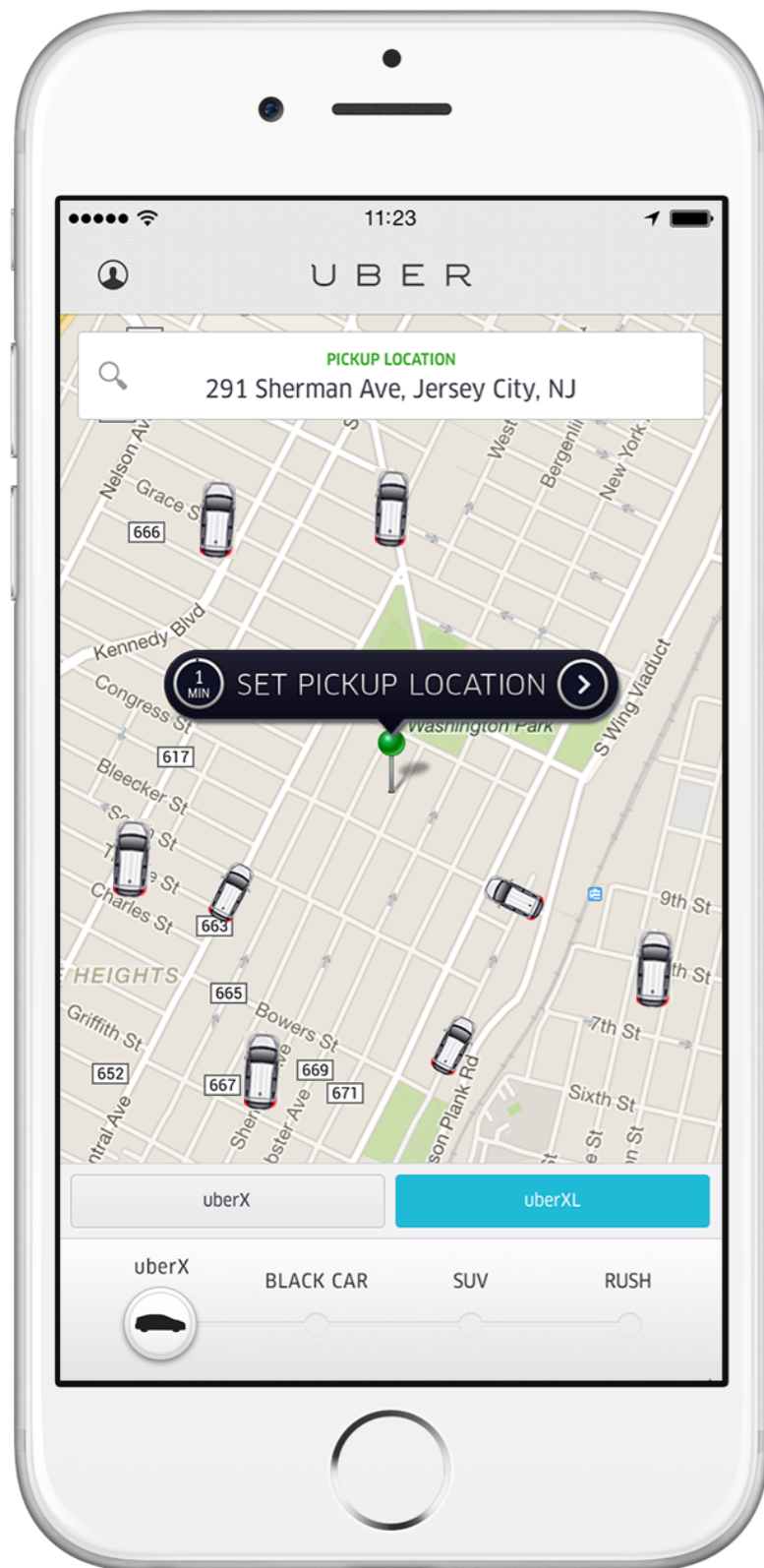
- Supply and demand are estimations

Uber client app limitations



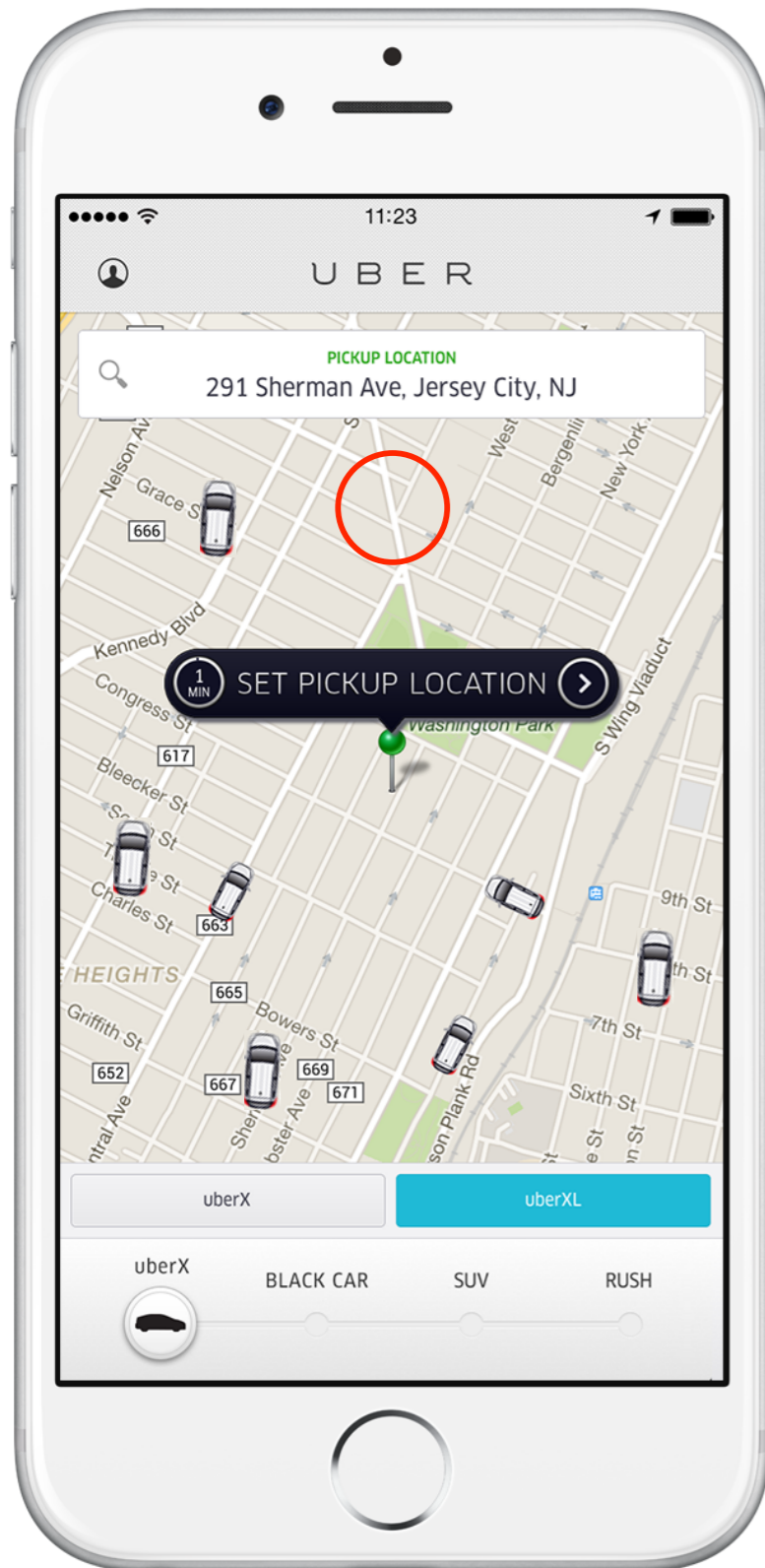
- Supply and demand are estimations
 - Supply = number of cars on the road

Uber client app limitations



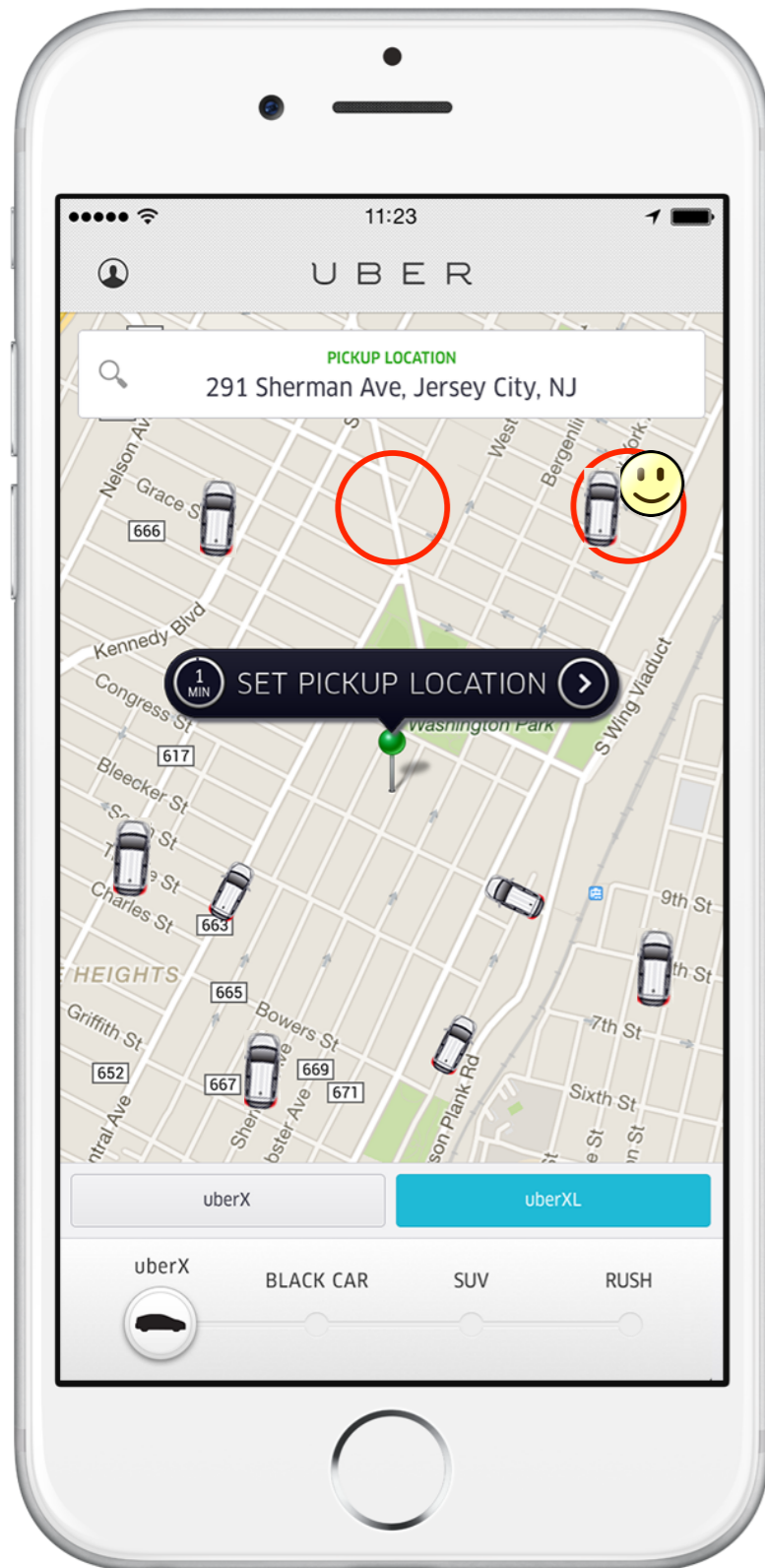
- Supply and demand are estimations
 - Supply = number of cars on the road
 - Fulfilled demand = number of cars
go offline

Uber client app limitations



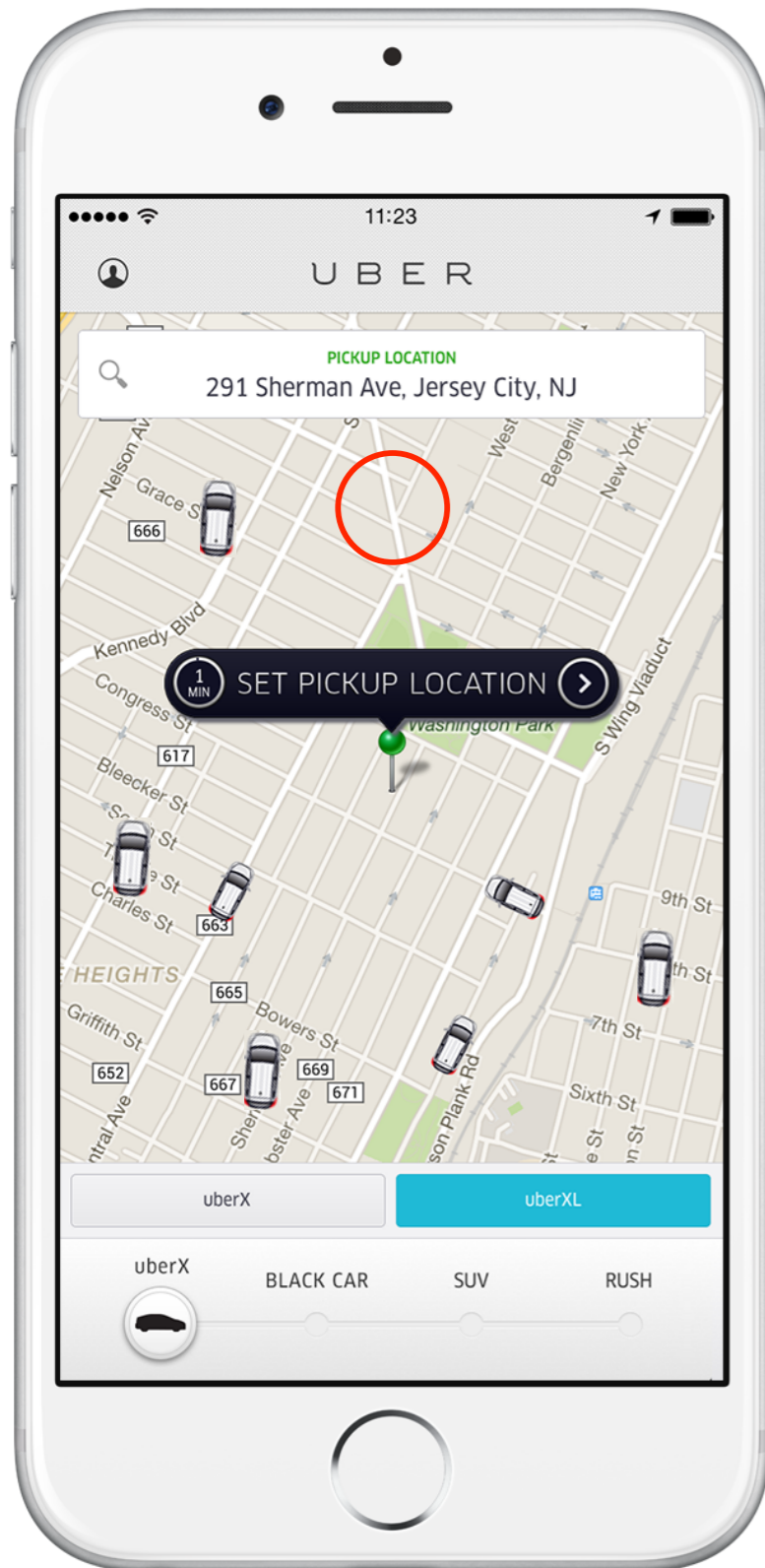
- Supply and demand are estimations
 - Supply = number of cars on the road
 - Fulfilled demand = number of cars
go offline

Uber client app limitations



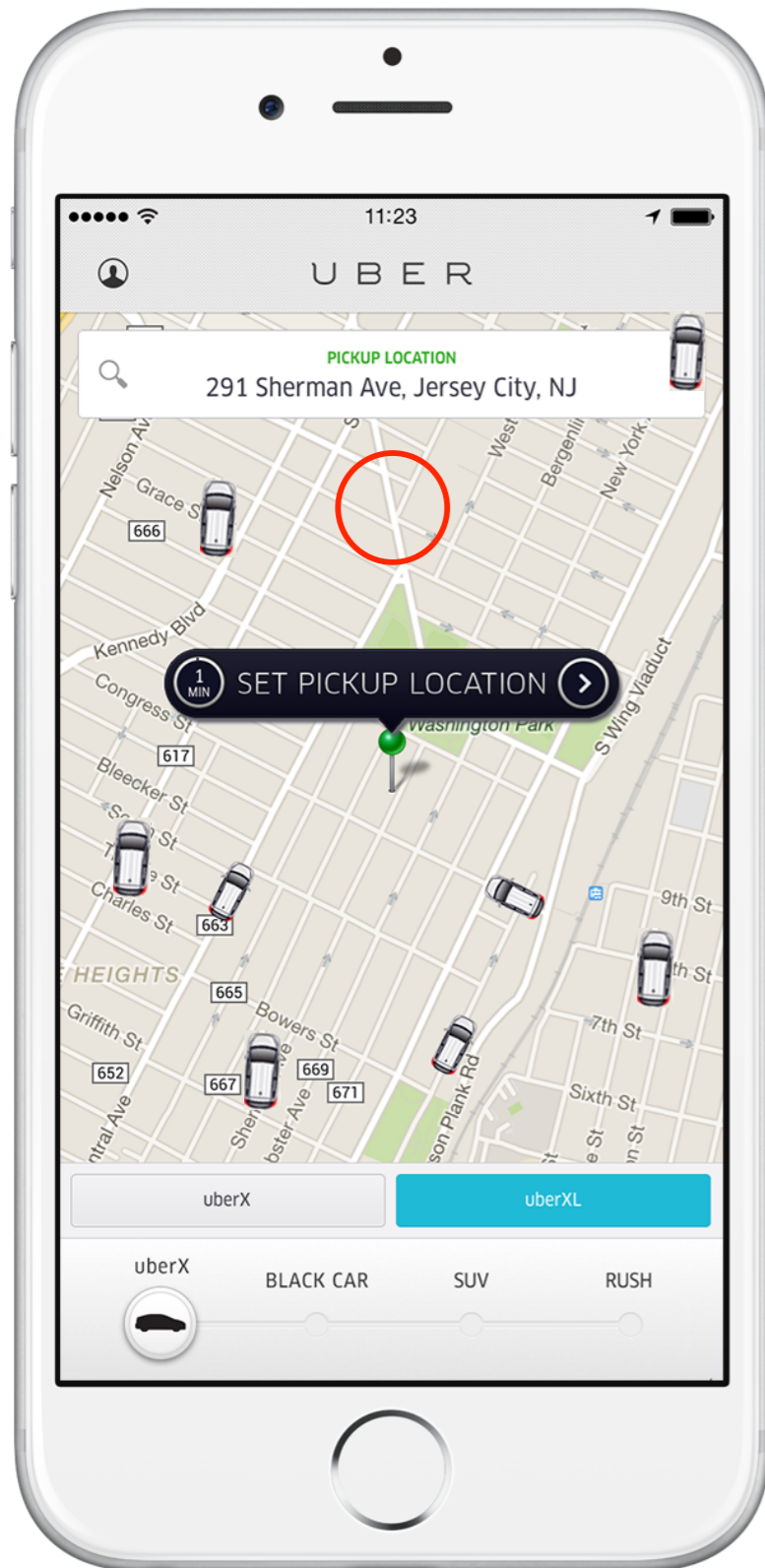
- Supply and demand are estimations
 - Supply = number of cars on the road
 - Fulfilled demand = number of cars
go offline
 - ▶ Got booked

Uber client app limitations



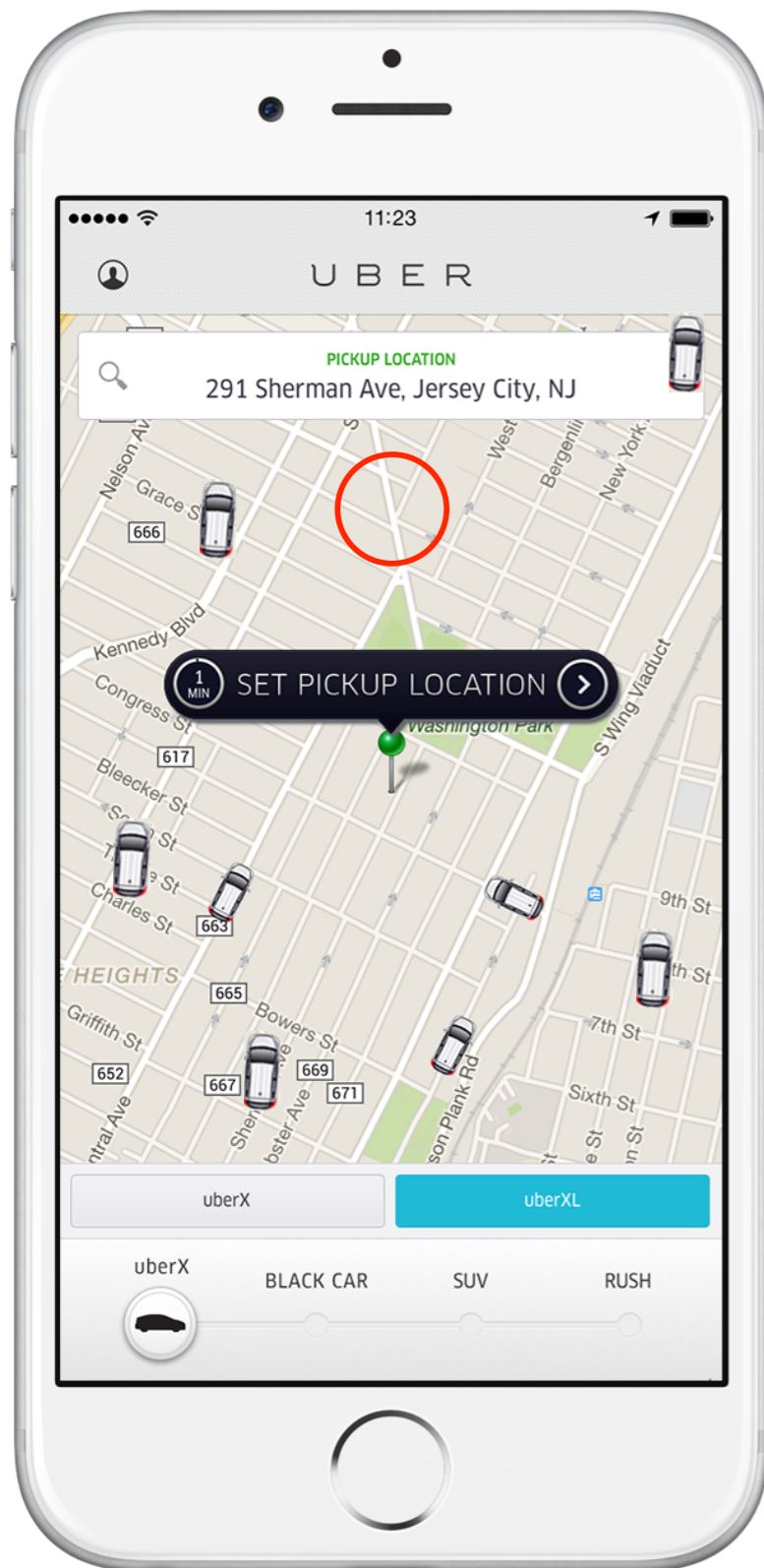
- Supply and demand are estimations
 - Supply = number of cars on the road
 - Fulfilled demand = number of cars
go offline
 - ▶ Got booked
 - ▶ Logged out

Uber client app limitations



- Supply and demand are estimations
 - Supply = number of cars on the road
 - Fulfilled demand = number of cars
go offline
 - ▶ Got booked
 - ▶ Logged out
 - ▶ Driving out of boundary

Uber client app limitations



- Supply and demand are estimations
 - Supply = number of cars on the road
 - Fulfilled demand = number of cars go offline
 - ▶ Got booked
 - ▶ Logged out
 - ▶ Driving out of boundary
- Upper bound of fulfilled demand!

Uber client app limitations

Uber client app limitations

- Limited visibility

Uber client app limitations

- Limited visibility
 - Only 8 nearest cars

Uber client app limitations

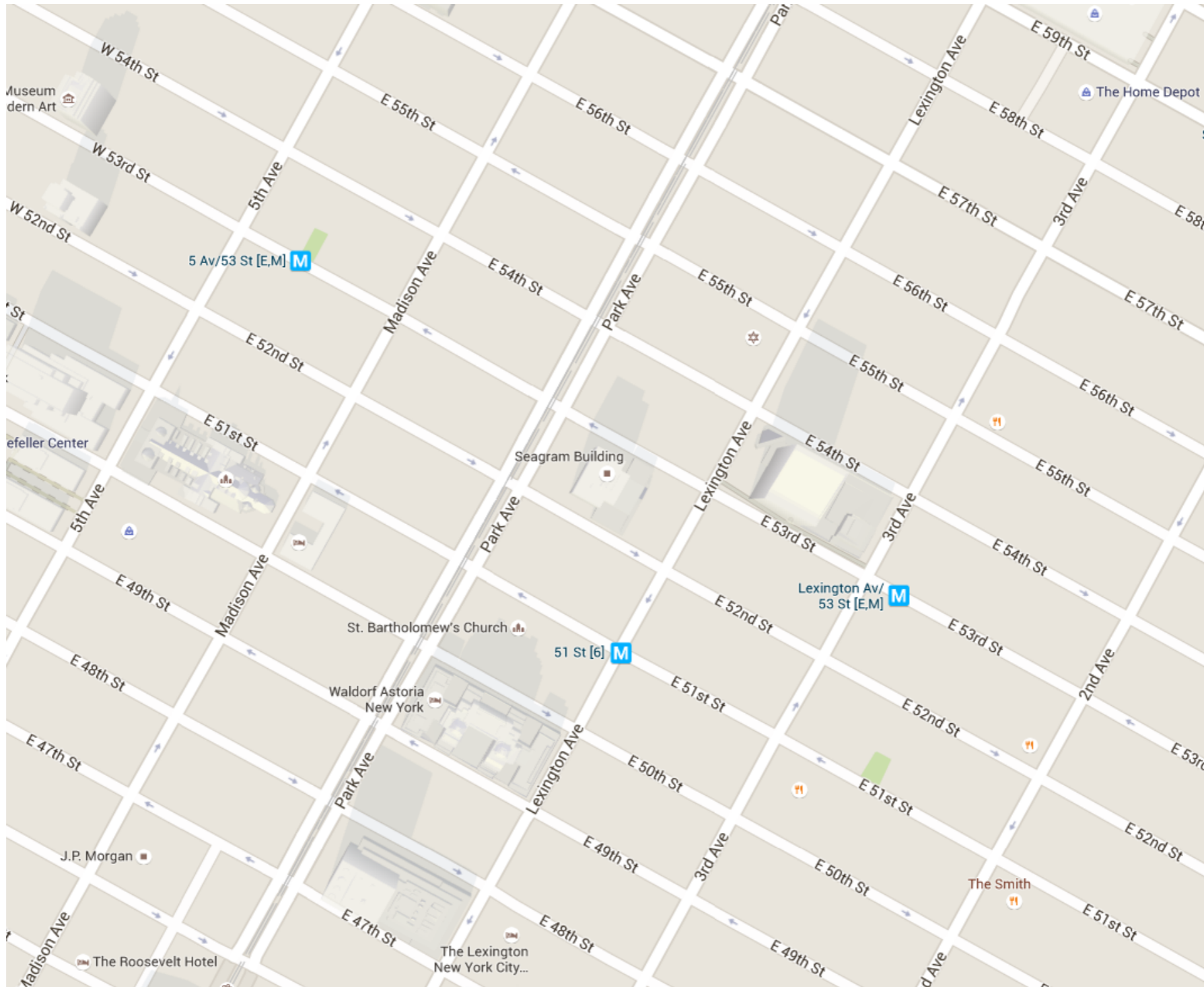
- Limited visibility
 - Only 8 nearest cars
 - Limited number of measuring points

Uber client app limitations

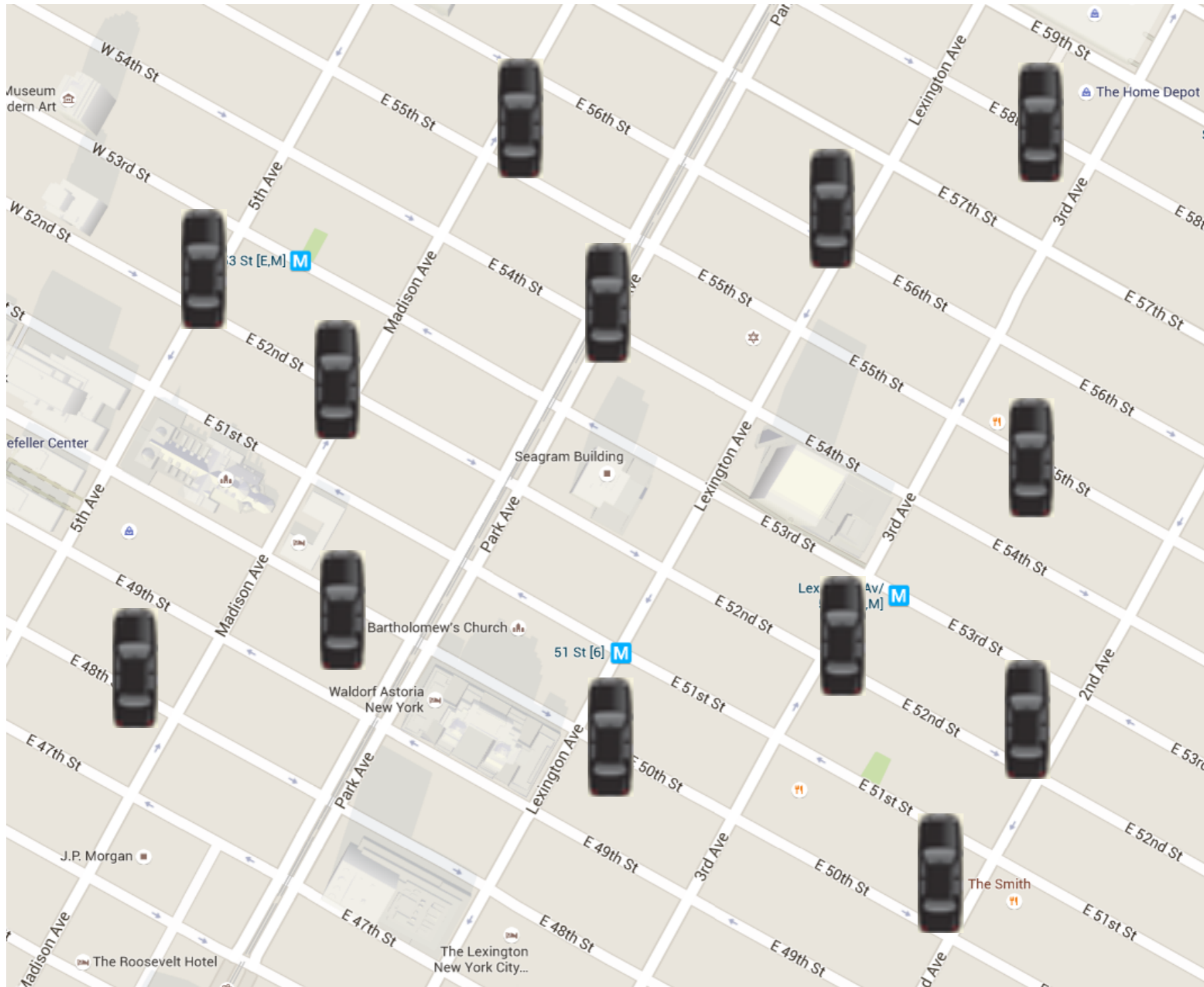
- Limited visibility
 - Only 8 nearest cars
 - Limited number of measuring points
- How far away between points?

Limited visibility

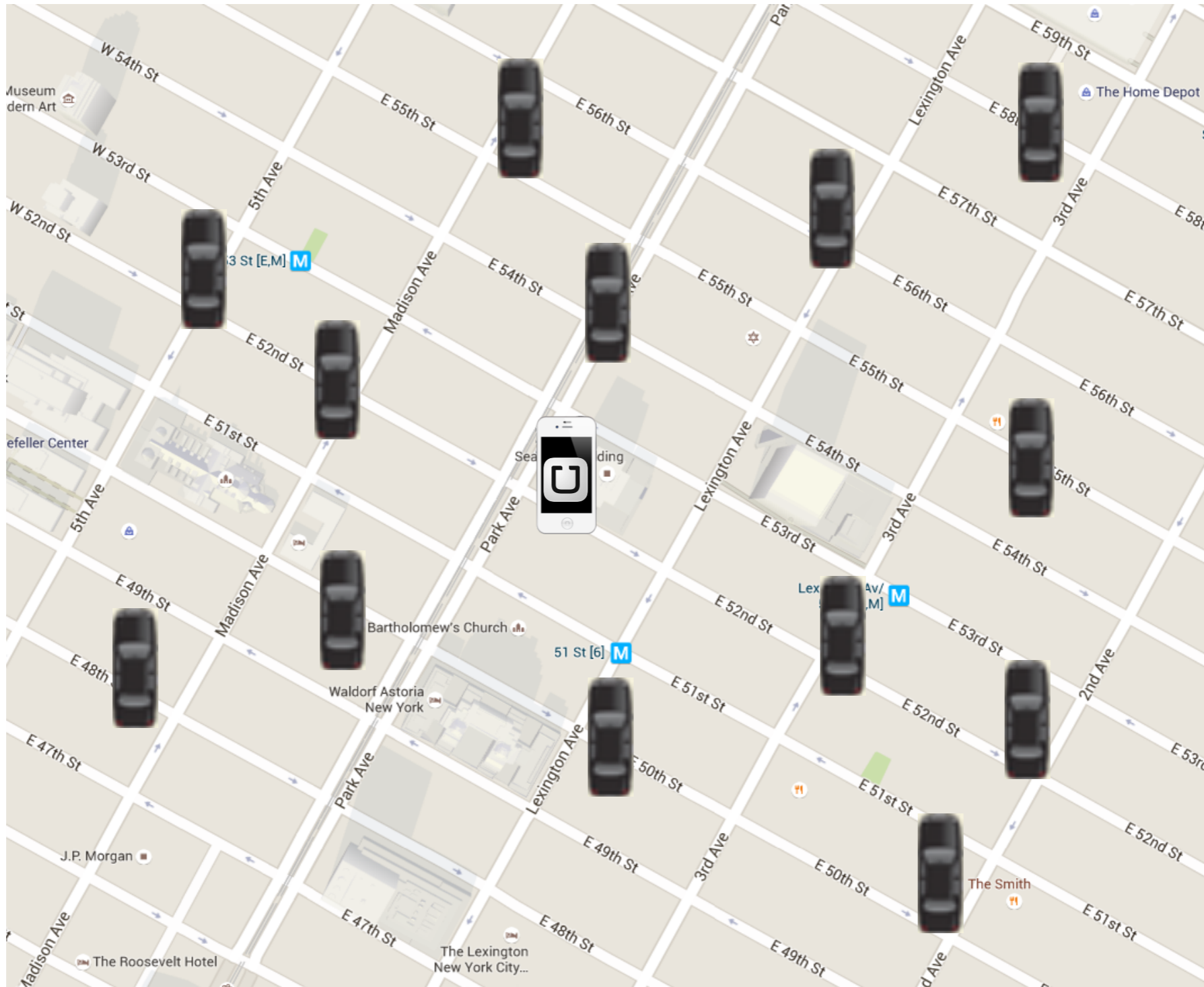
Limited visibility



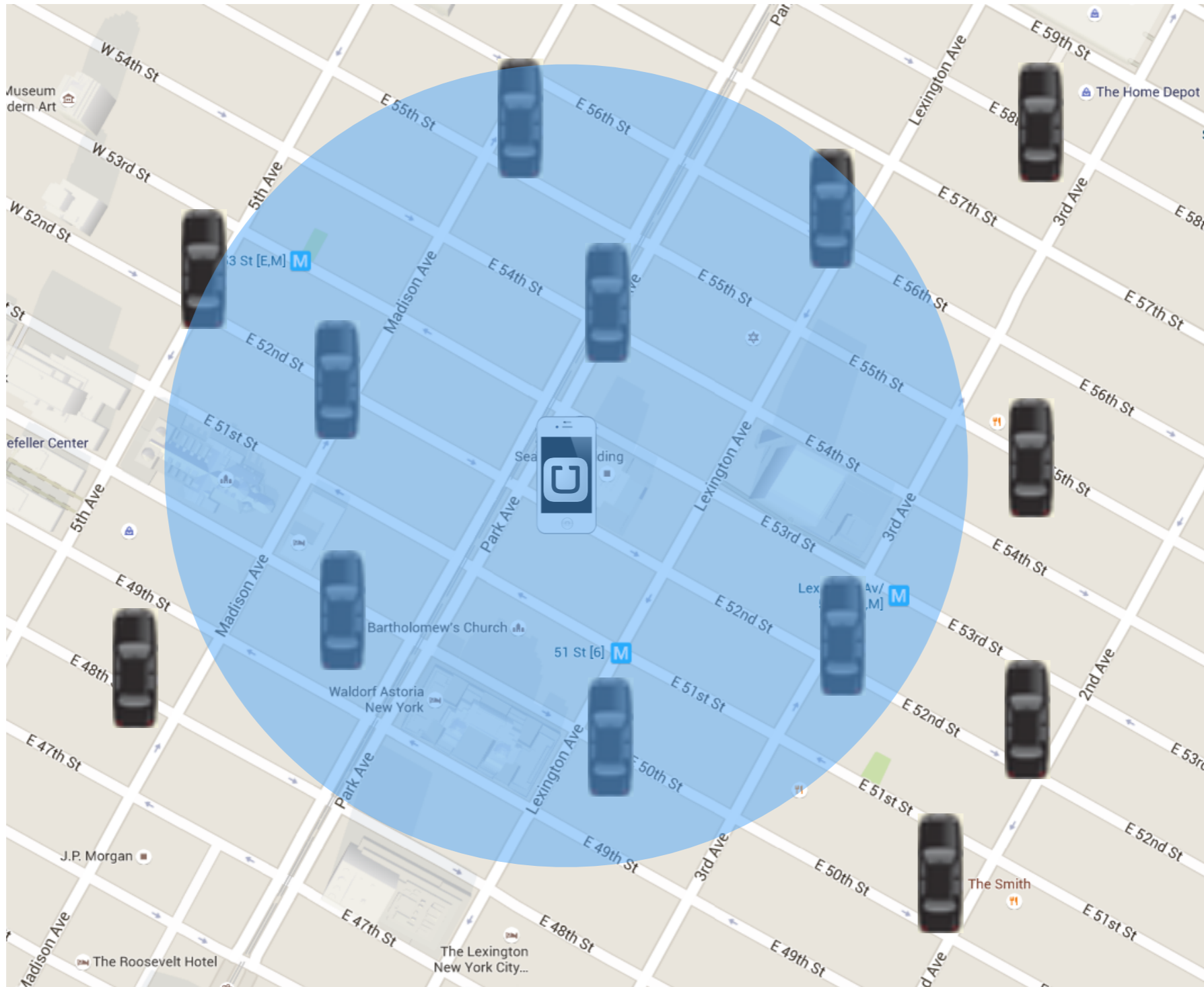
Limited visibility



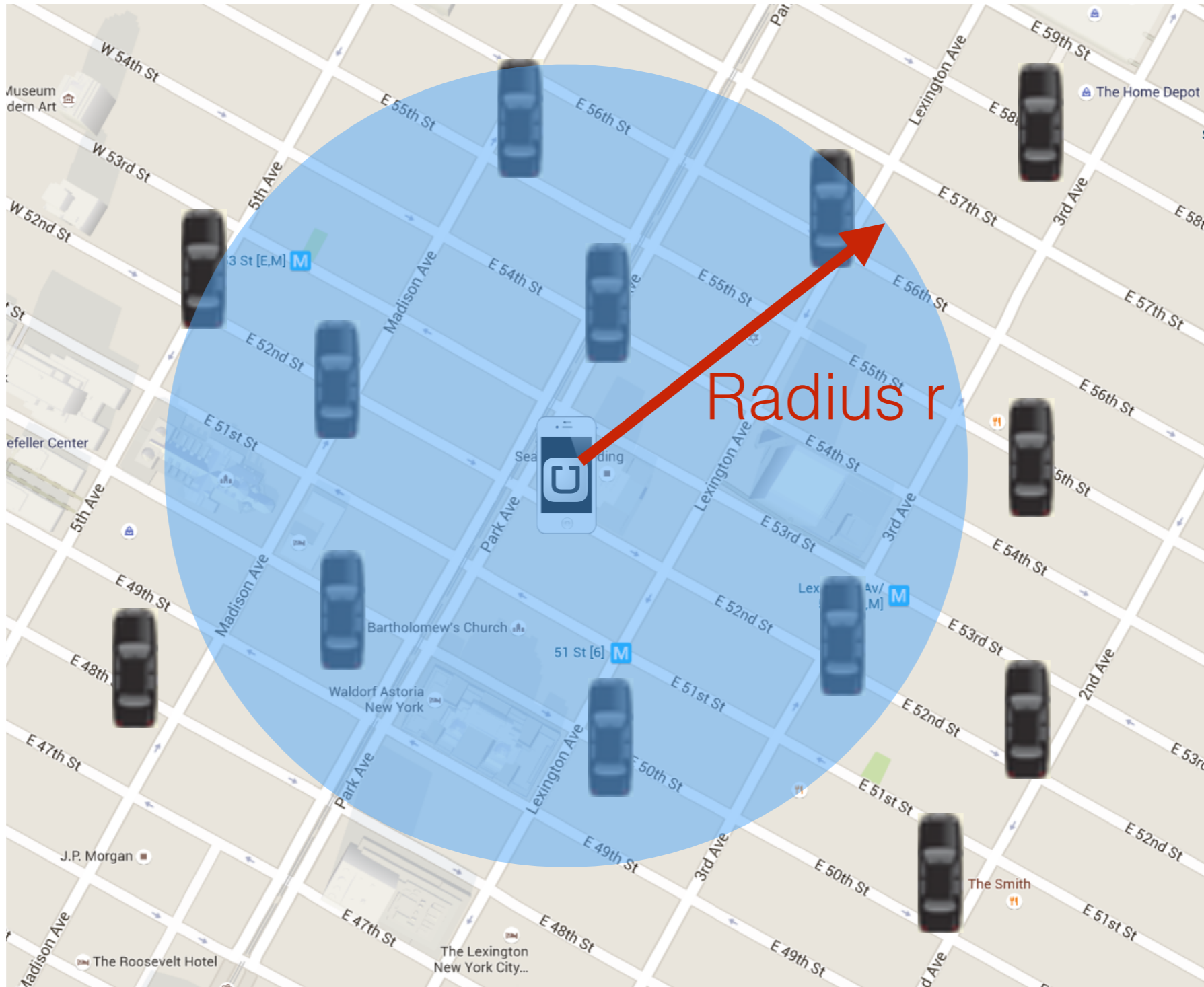
Limited visibility



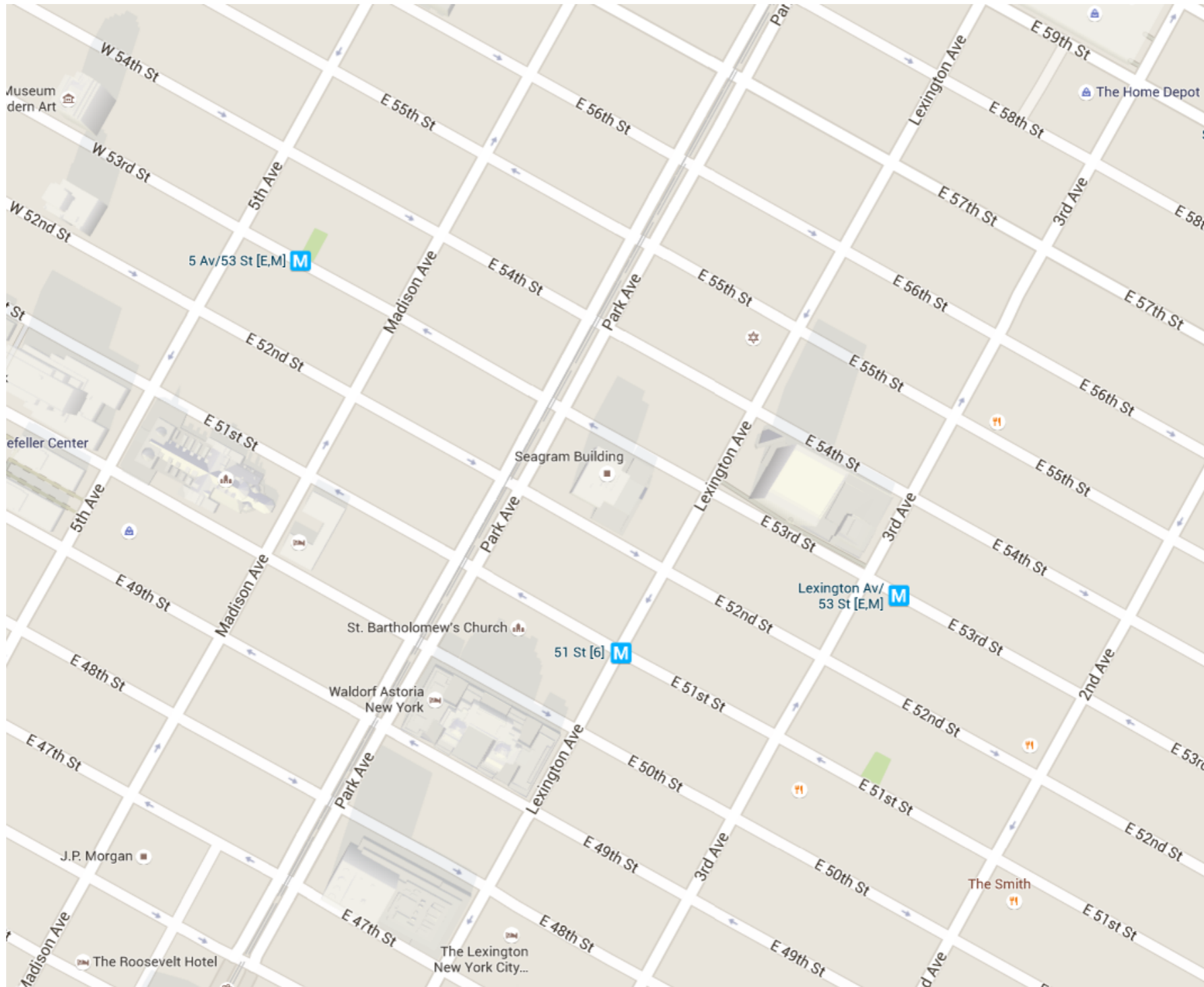
Limited visibility



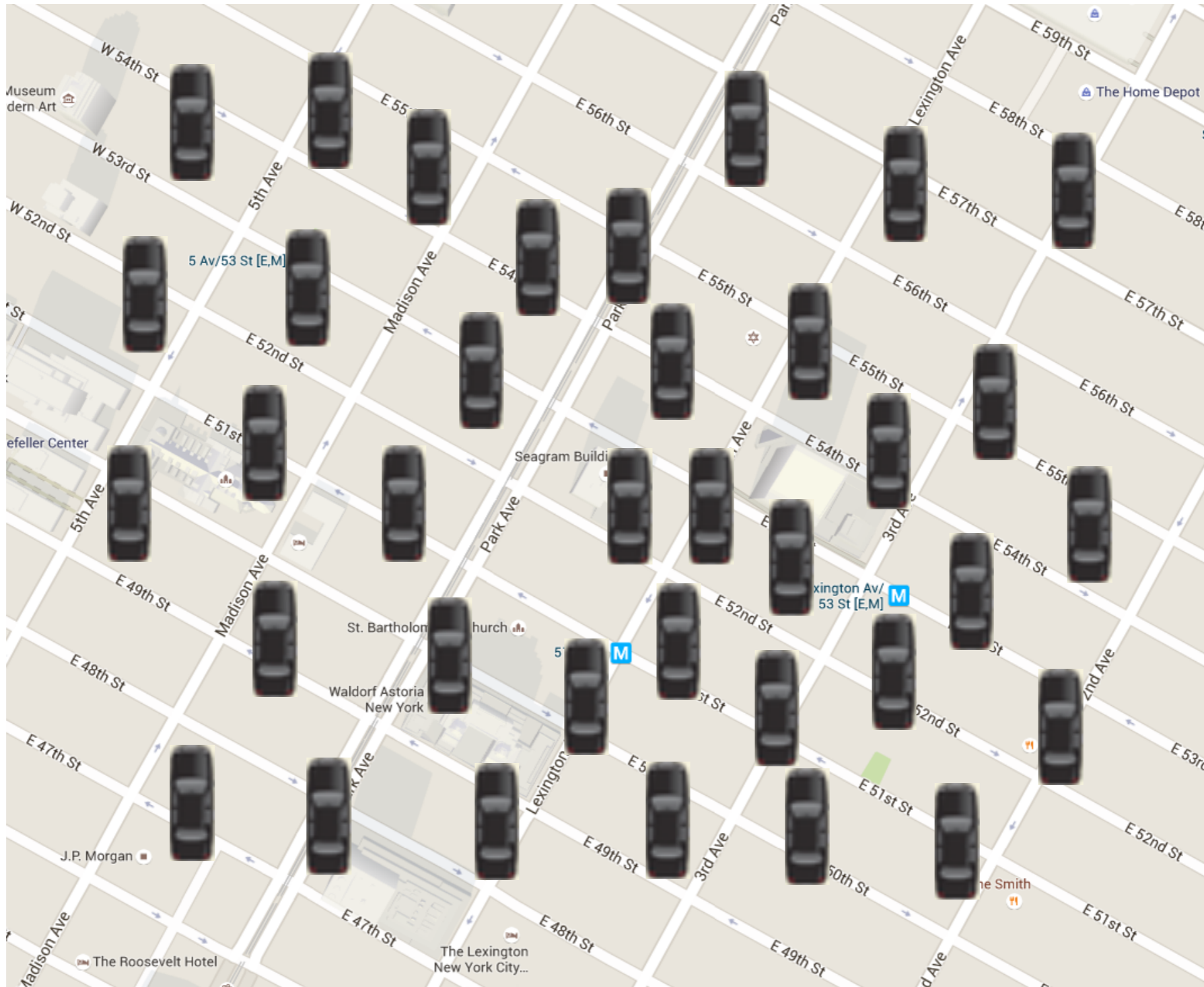
Limited visibility



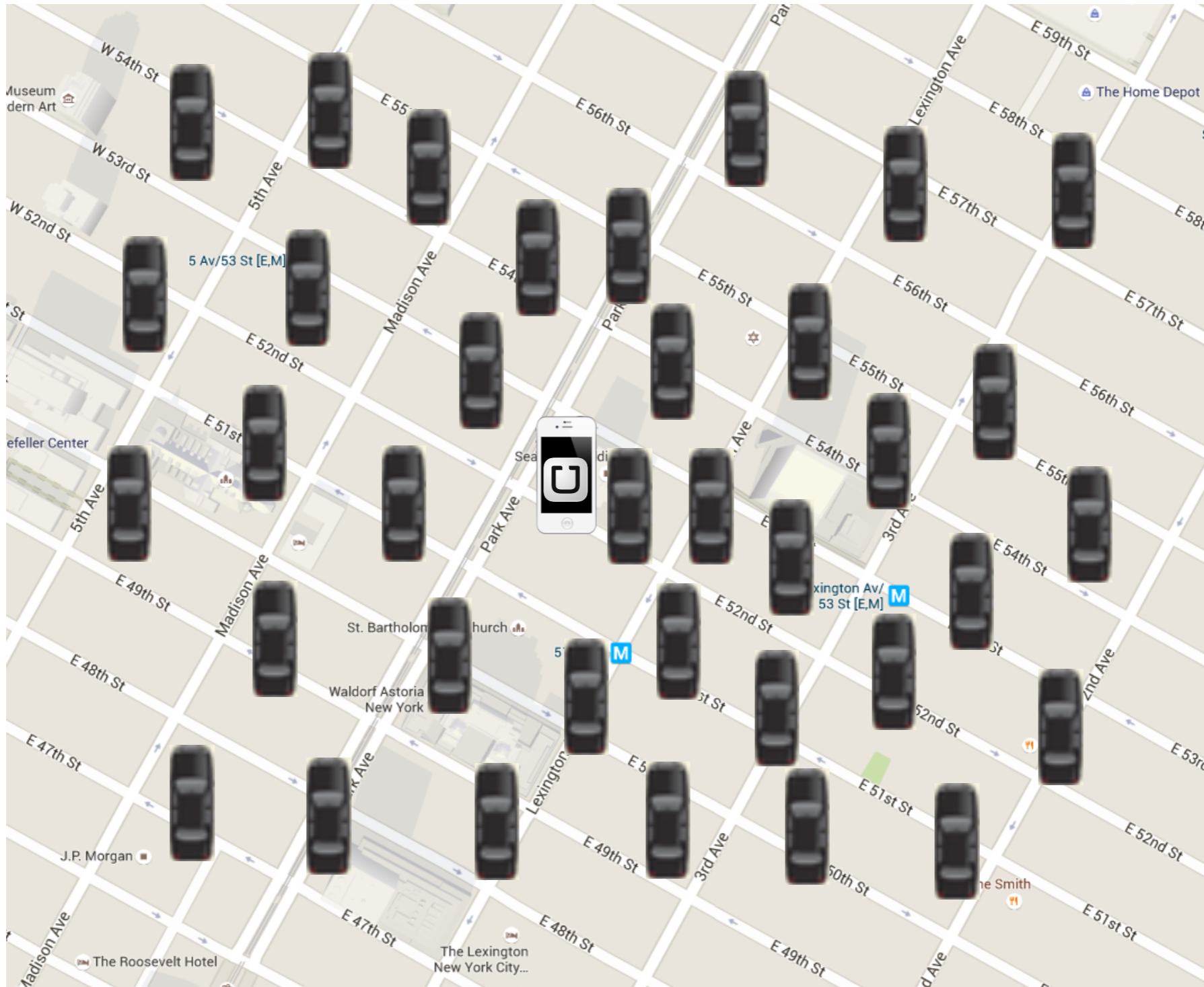
Limited visibility



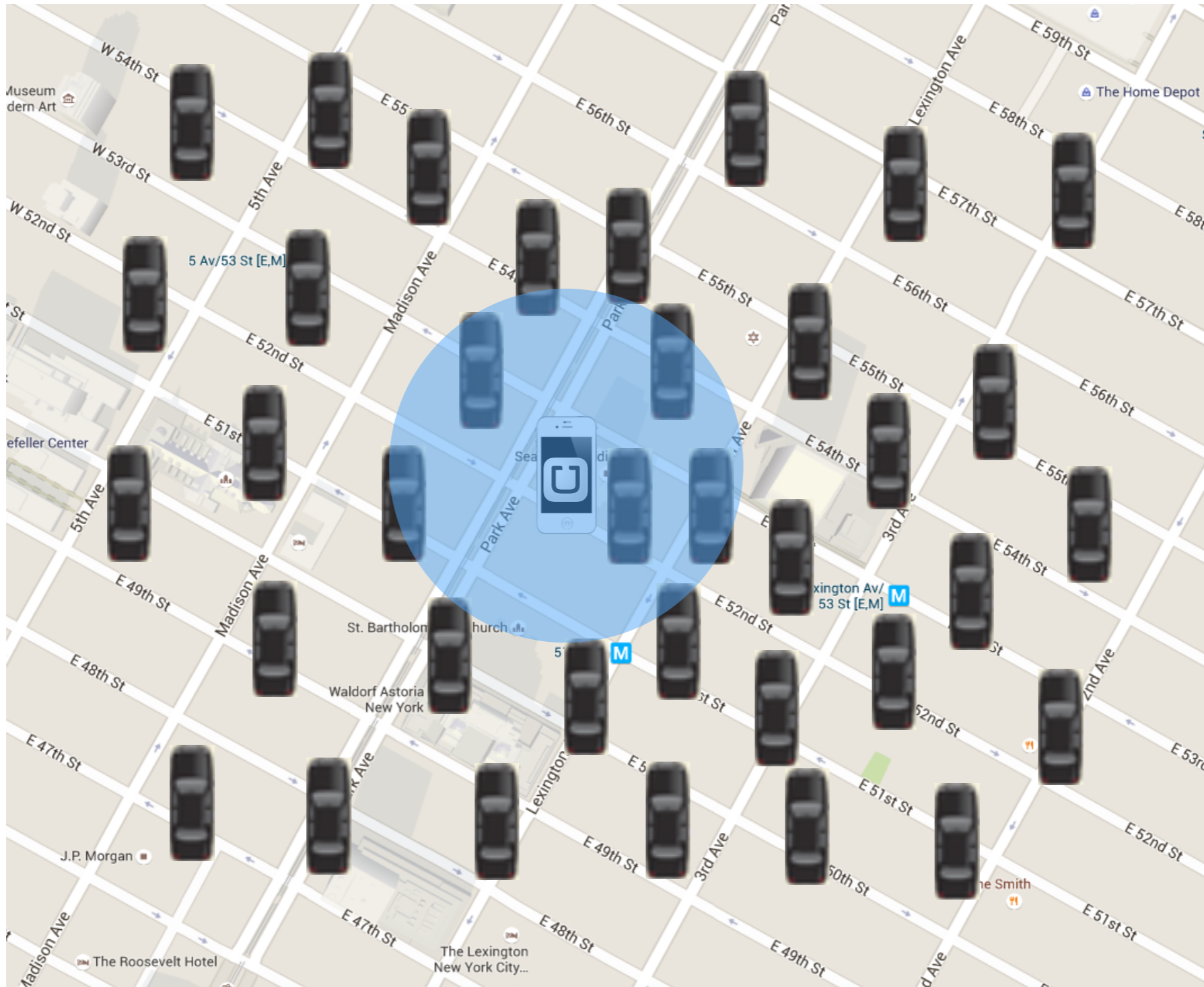
Limited visibility



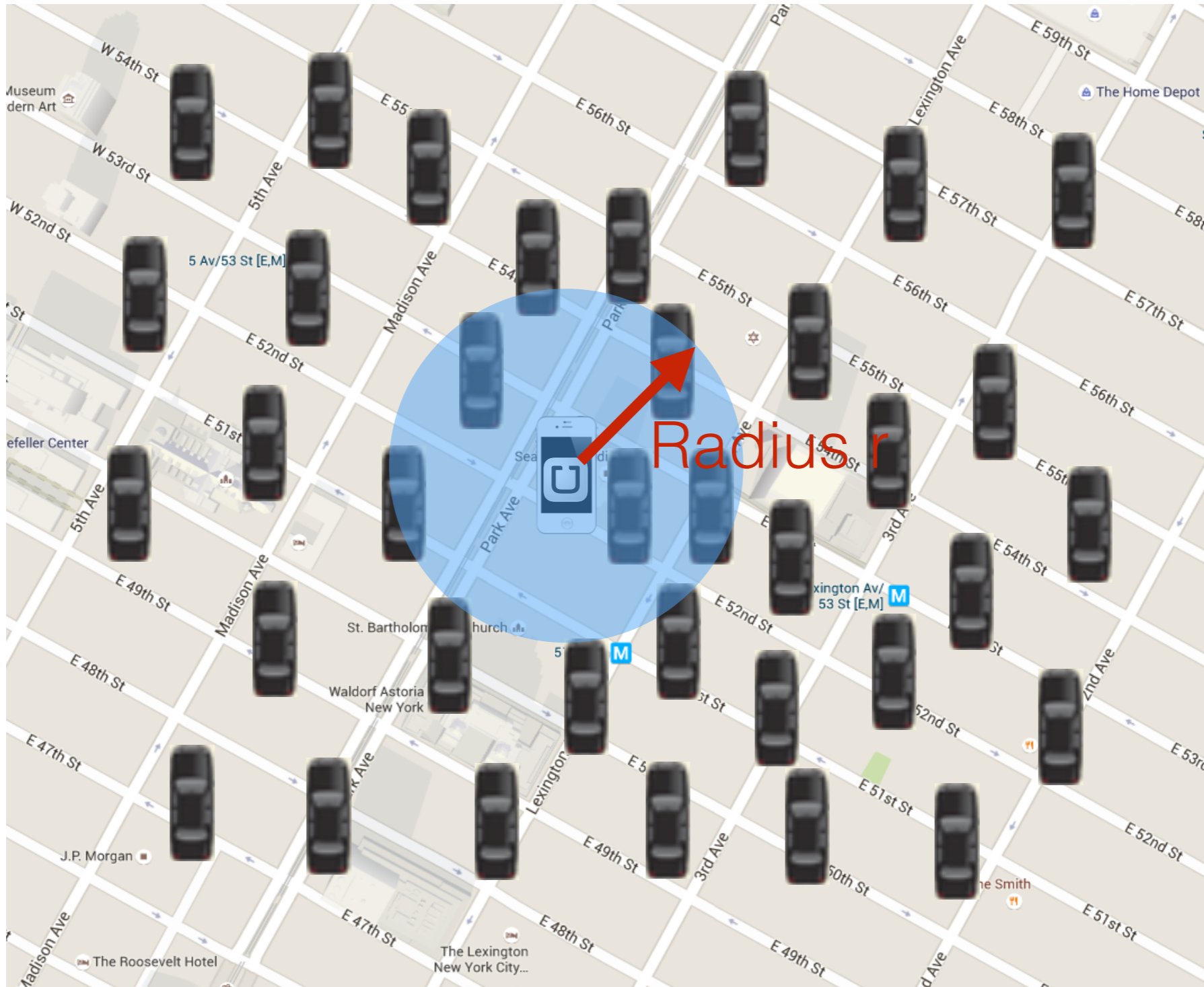
Limited visibility



Limited visibility



Limited visibility



Radius measurement

Radius measurement

- Measuring radius (details in paper)

Radius measurement

- Measuring radius (details in paper)
 - Midtown Manhattan: 247m

Radius measurement

- Measuring radius (details in paper)
 - Midtown Manhattan: 247m
 - Downtown San Francisco: 387m

Radius measurement

- Measuring radius (details in paper)
 - Midtown Manhattan: 247m
 - Downtown San Francisco: 387m
 - Use the NYC taxi data to validate that 97% of supply and 95% of the demand are covered

Radius measurement

- Measuring radius (details in paper)
 - Midtown Manhattan: 247m
 - Downtown San Francisco: 387m
 - Use the NYC taxi data to validate that 97% of supply and 95% of the demand are covered
- NYC and SF are the targets for collecting data

Radius measurement

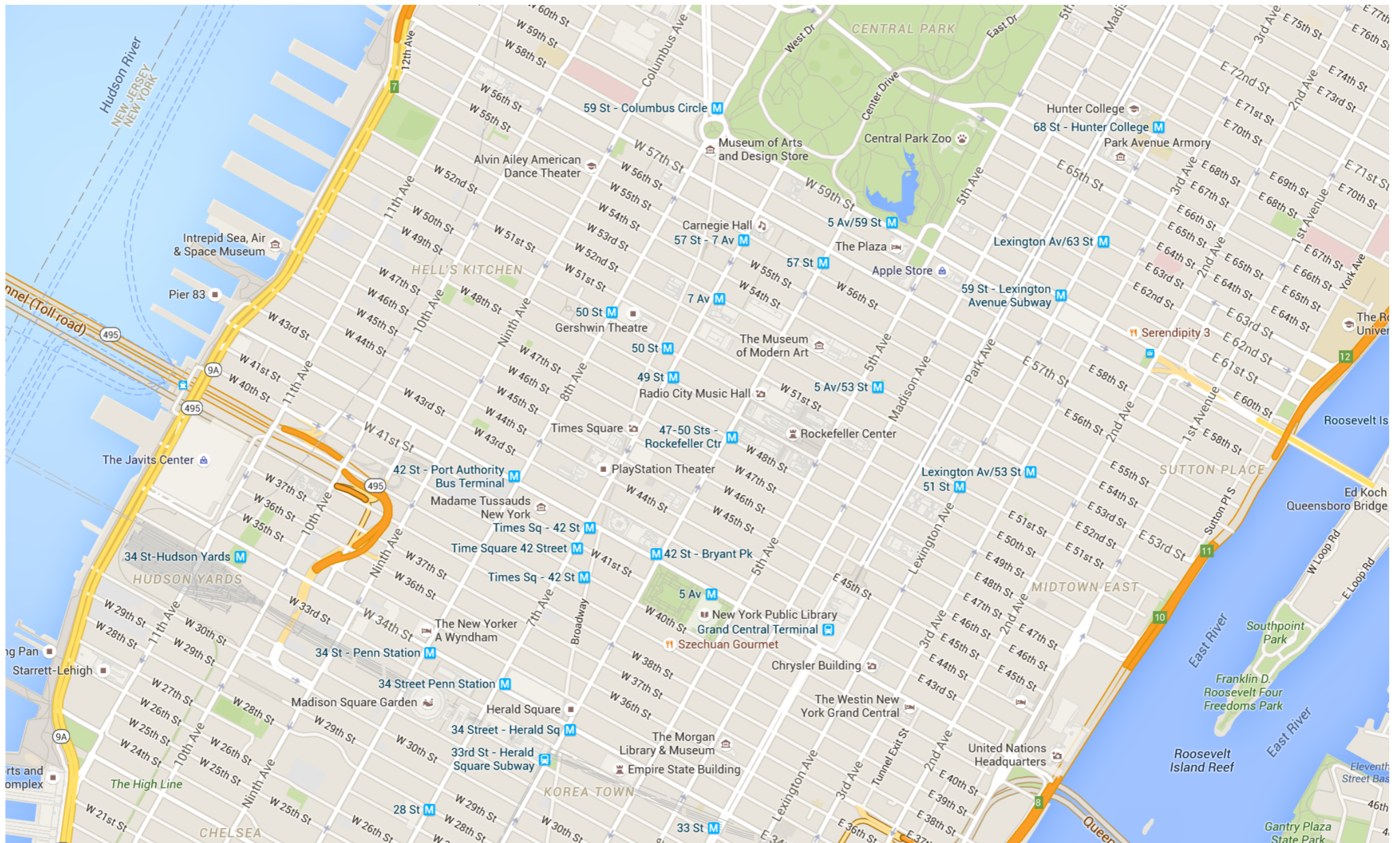
- Measuring radius (details in paper)
 - Midtown Manhattan: 247m
 - Downtown San Francisco: 387m
 - Use the NYC taxi data to validate that 97% of supply and 95% of the demand are covered
- NYC and SF are the targets for collecting data
 - 2nd and 3rd largest Uber market

Radius measurement

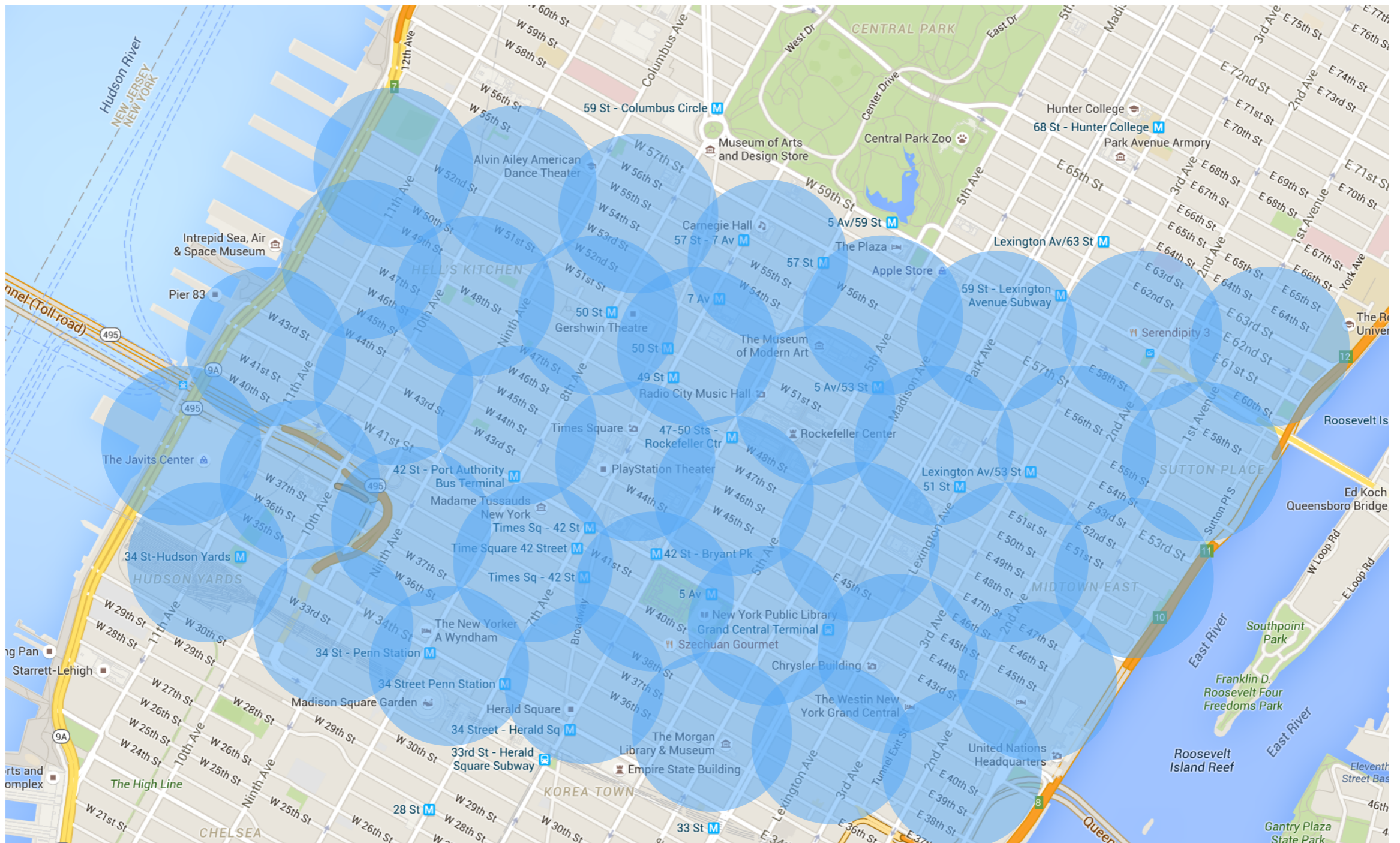
- Measuring radius (details in paper)
 - Midtown Manhattan: 247m
 - Downtown San Francisco: 387m
 - Use the NYC taxi data to validate that 97% of supply and 95% of the demand are covered
- NYC and SF are the targets for collecting data
 - 2nd and 3rd largest Uber market
 - Different access to public transportation

The measurement grid

The measurement grid



The measurement grid



Ethics

Ethics

- We did not collect personal information of passengers and Uber drivers

Ethics

- We did not collect personal information of passengers and Uber drivers
- We never booked Uber rides

Ethics

- We did not collect personal information of passengers and Uber drivers
- We never booked Uber rides
- We did not induce surges

Outline of Uber study

- Motivation
- Data collection
- Surge pricing
- Summary

Research questions

Research questions

- How much and how often does it surge on Uber?

Research questions

- How much and how often does it surge on Uber?
- How long do surge lasts?

Research questions

- How much and how often does it surge on Uber?
- How long do surge lasts?
- How do surge prices vary by location?

Research questions

- How much and how often does it surge on Uber?
- How long do surge lasts?
- How do surge prices vary by location?
- What features does Uber use to calculate surge multipliers?

Research questions

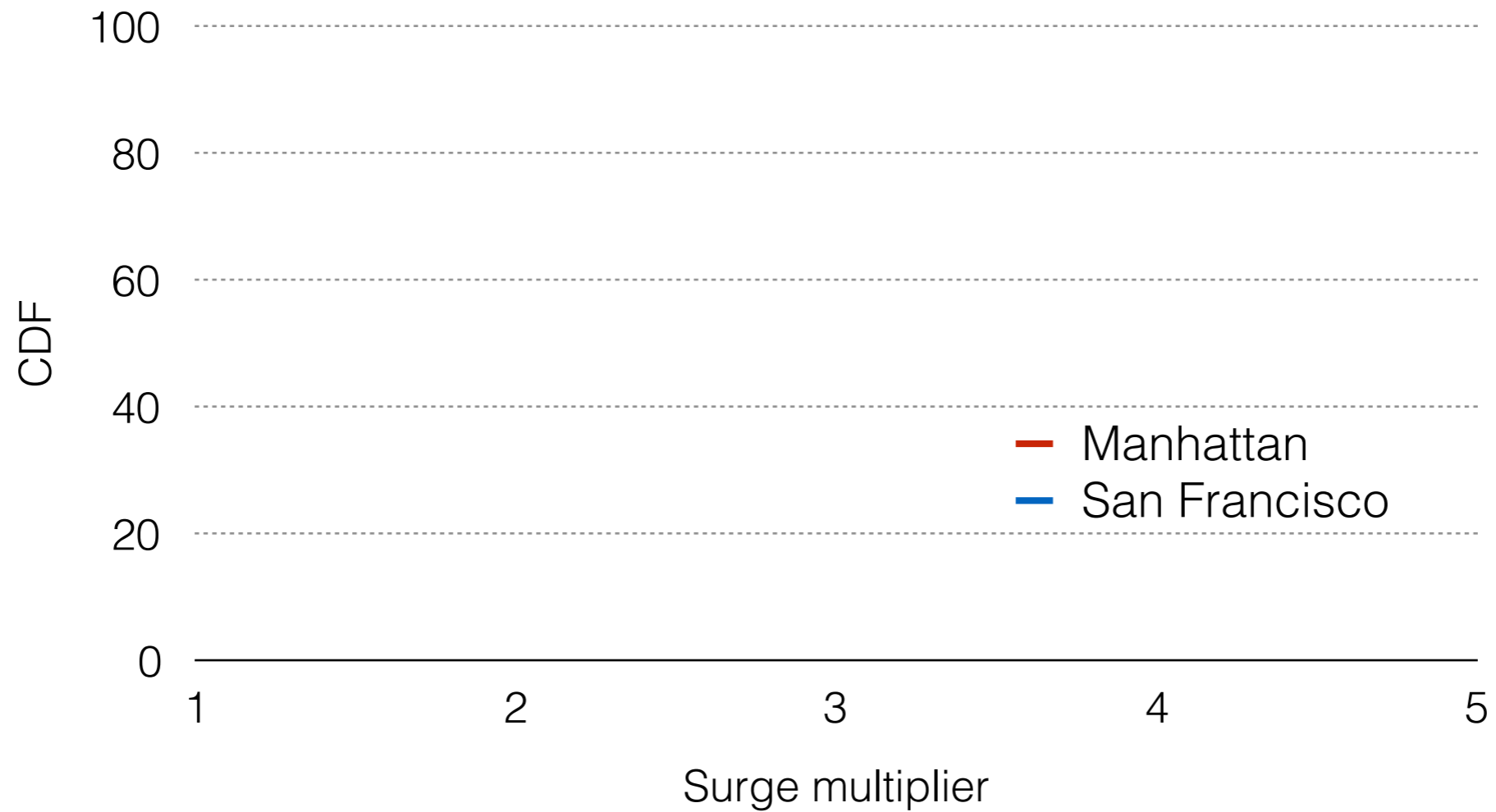
- How much and how often does it surge on Uber?
- How long do surge lasts?
- How do surge prices vary by location?
- What features does Uber use to calculate surge multipliers?
- Can we predict surges? If not, can we avoid surges?

Research questions

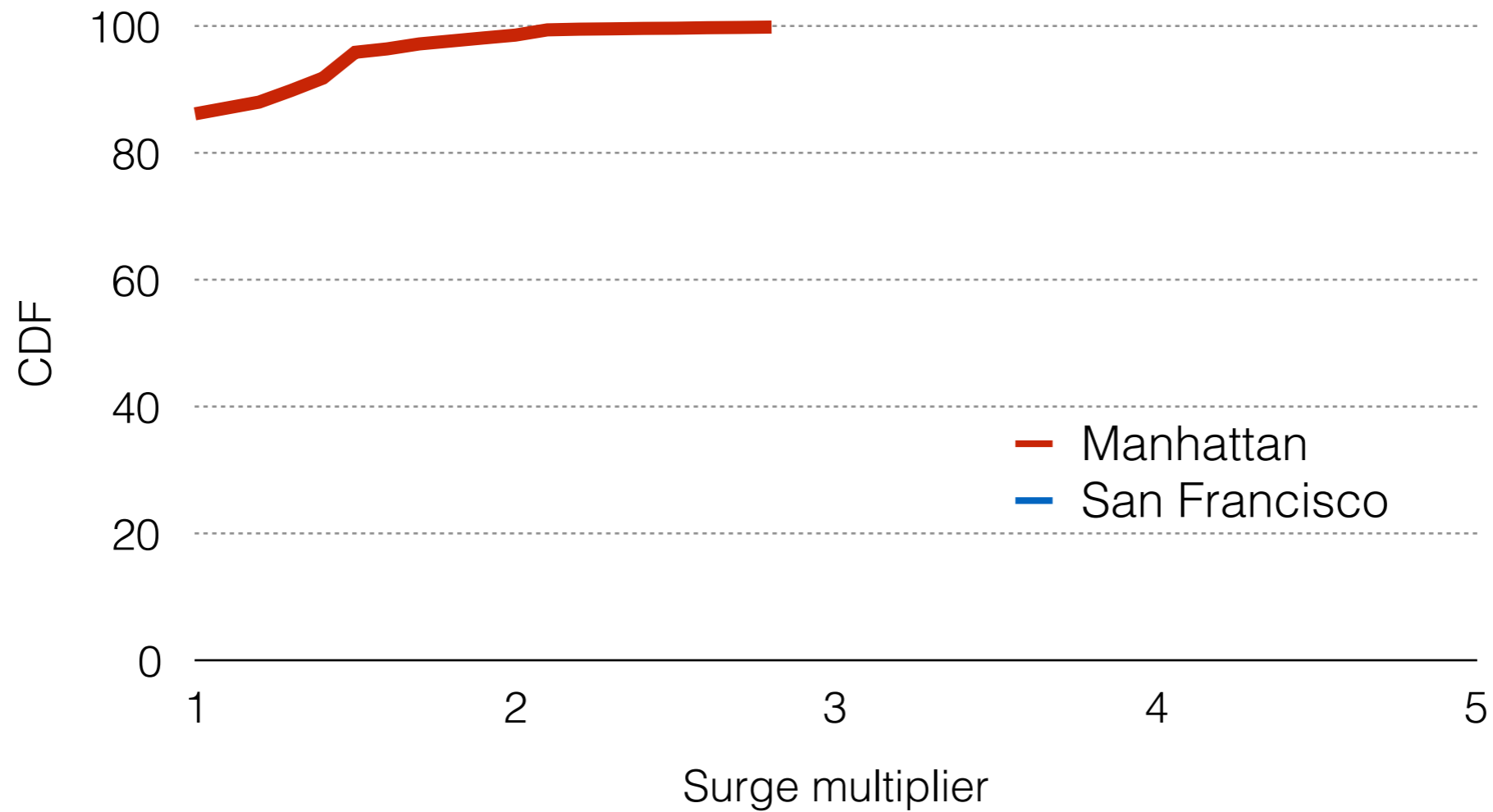
- How much and how often does it surge on Uber?
- How long do surge lasts?
- How do surge prices vary by location?
- What features does Uber use to calculate surge multipliers?
- Can we predict surges? If not, can we avoid surges?
- What is the impact of surge pricing?

How much and how often
does it surge on Uber?

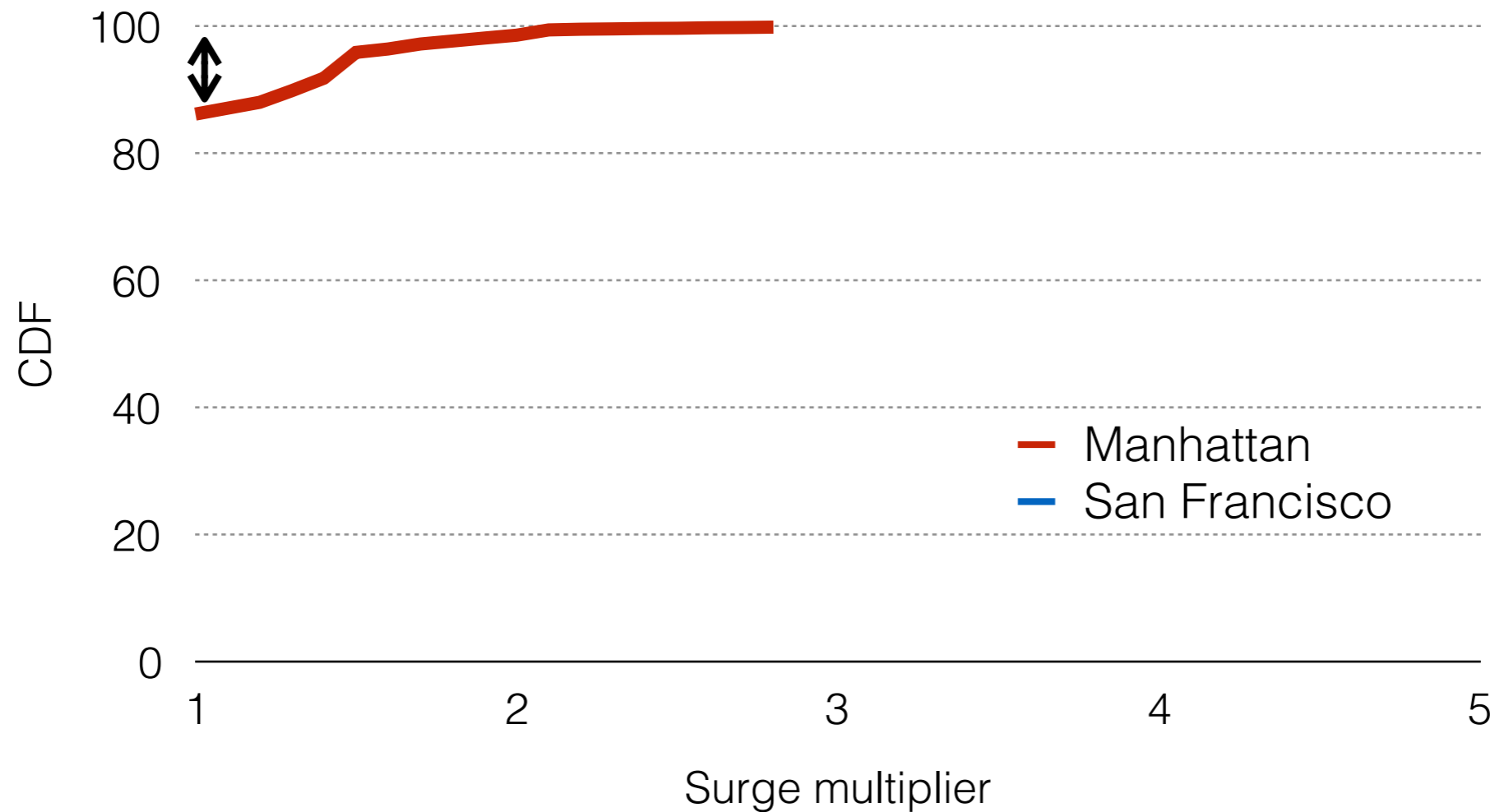
How much and how often does it surge on Uber?



How much and how often does it surge on Uber?

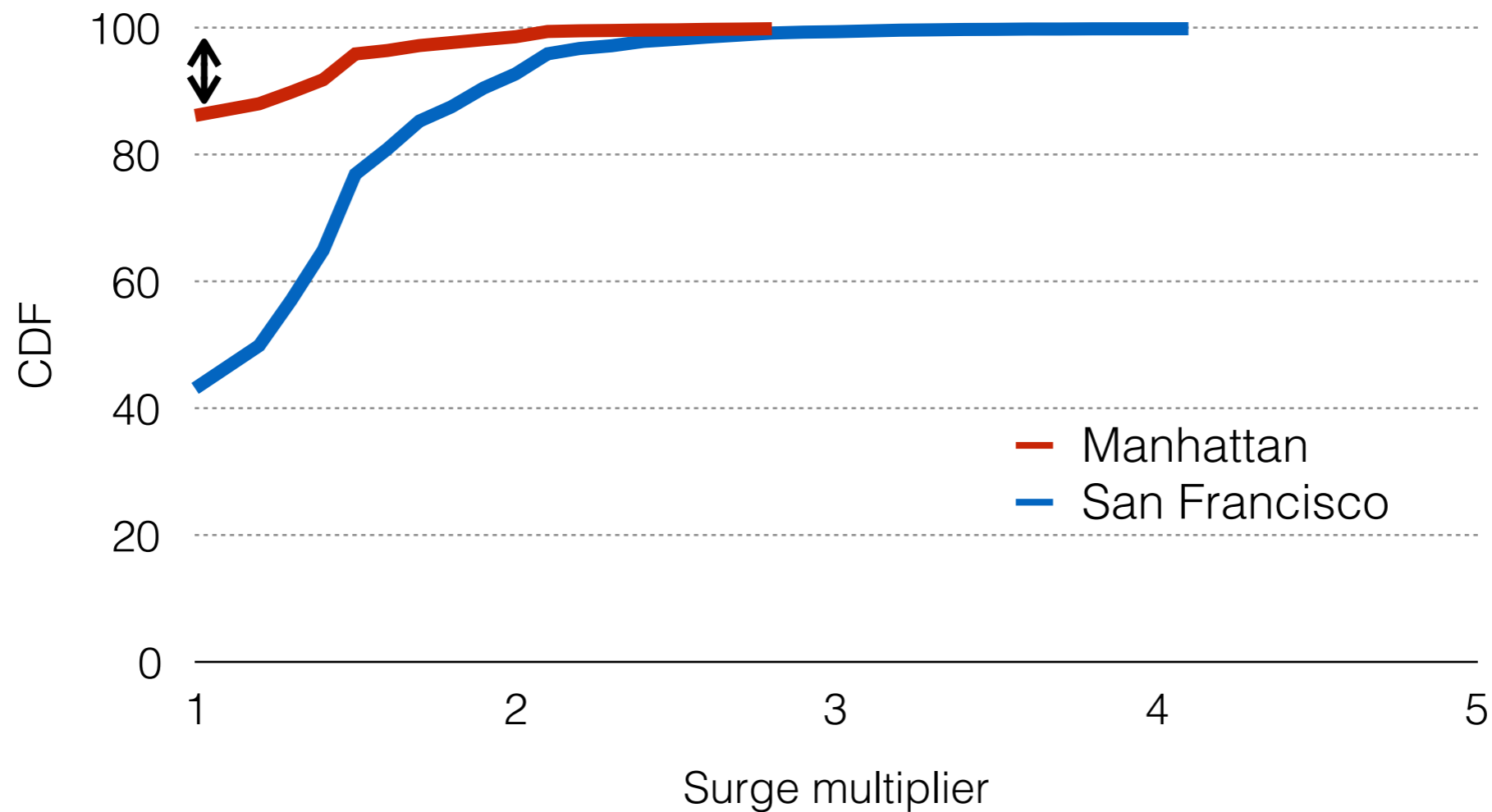


How much and how often does it surge on Uber?



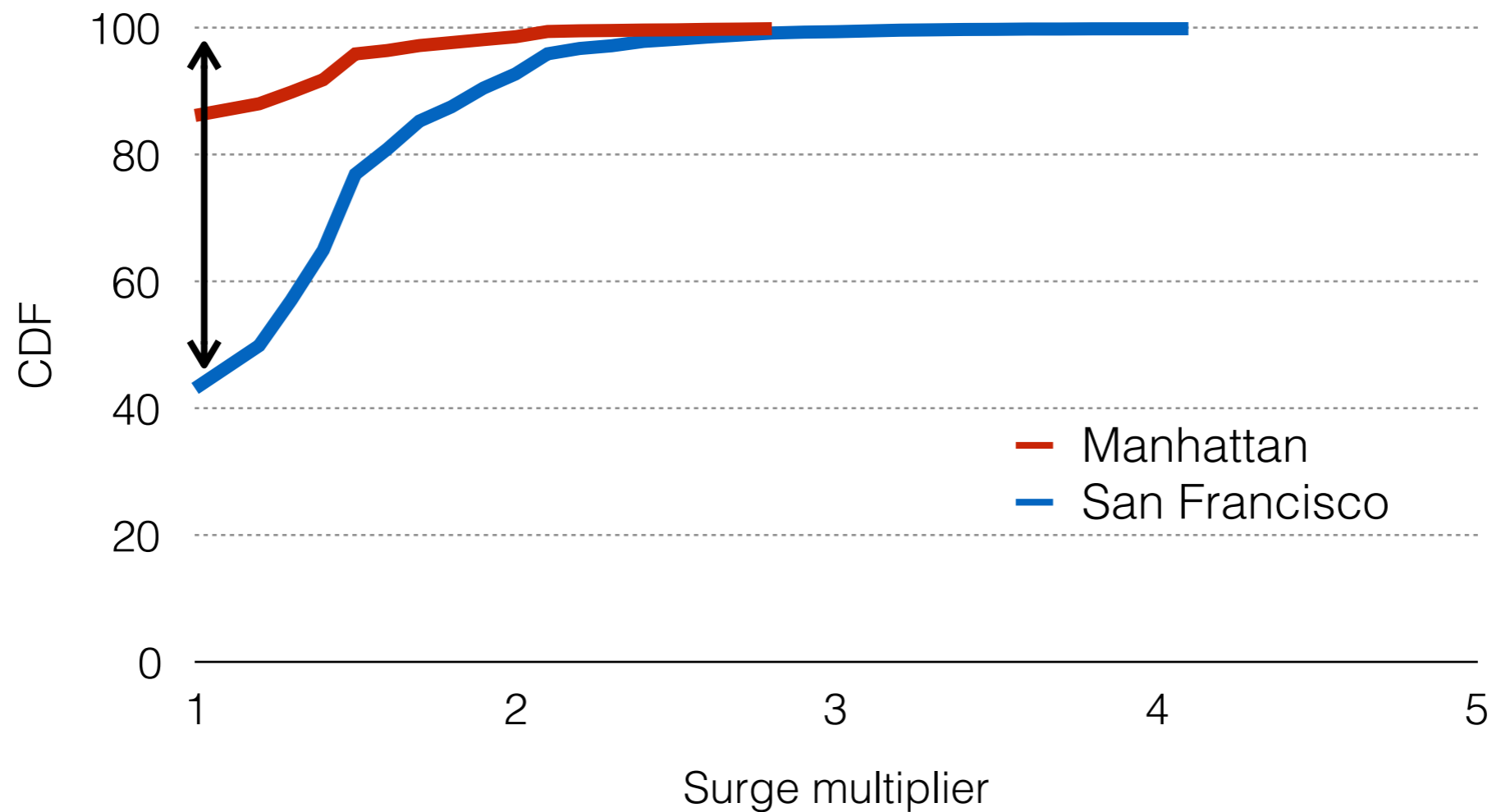
- 14% of time surging in Manhattan

How much and how often does it surge on Uber?



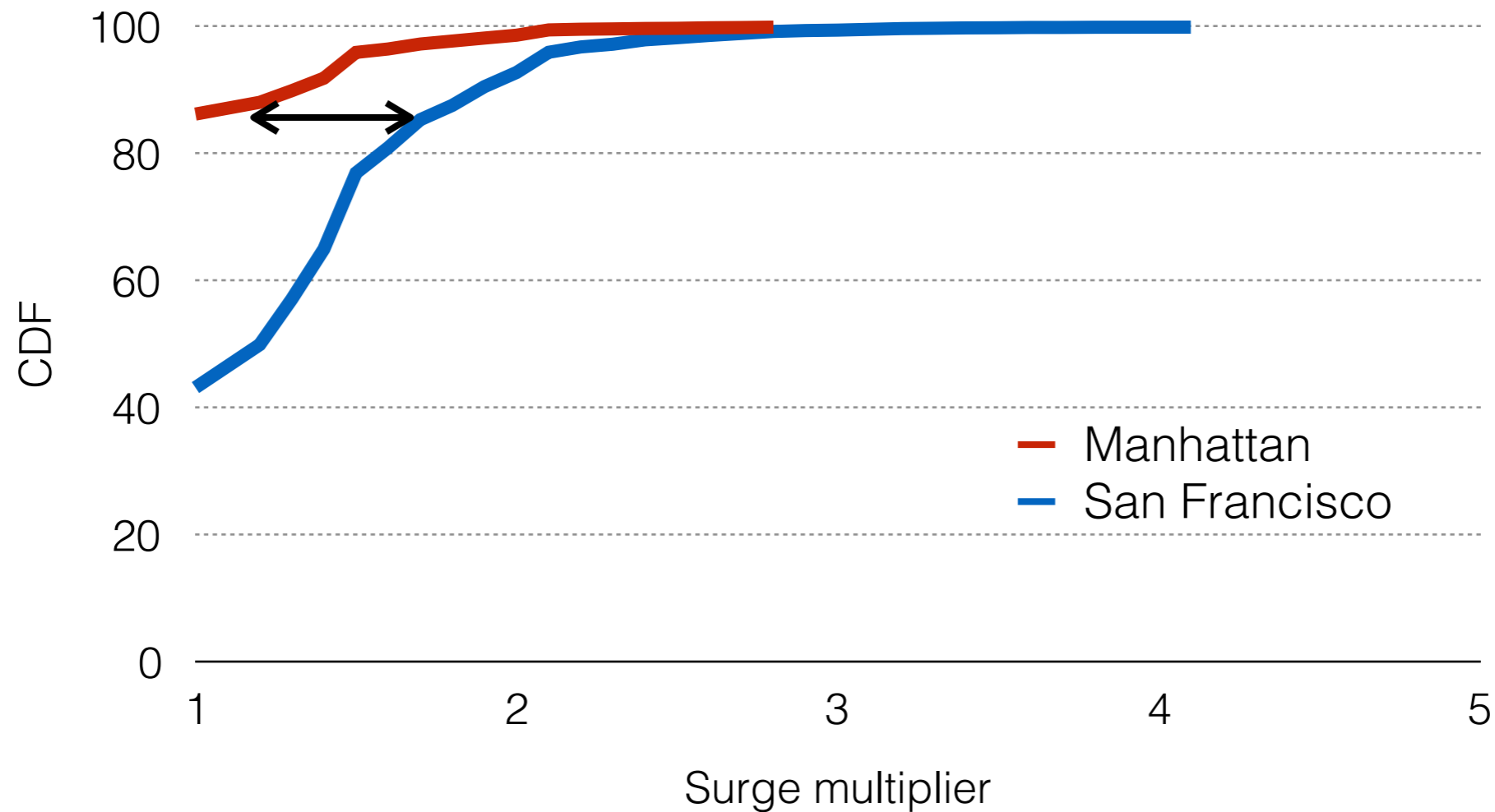
- 14% of time surging in Manhattan

How much and how often does it surge on Uber?



- 14% of time surging in Manhattan
- 57% of time surging in San Francisco

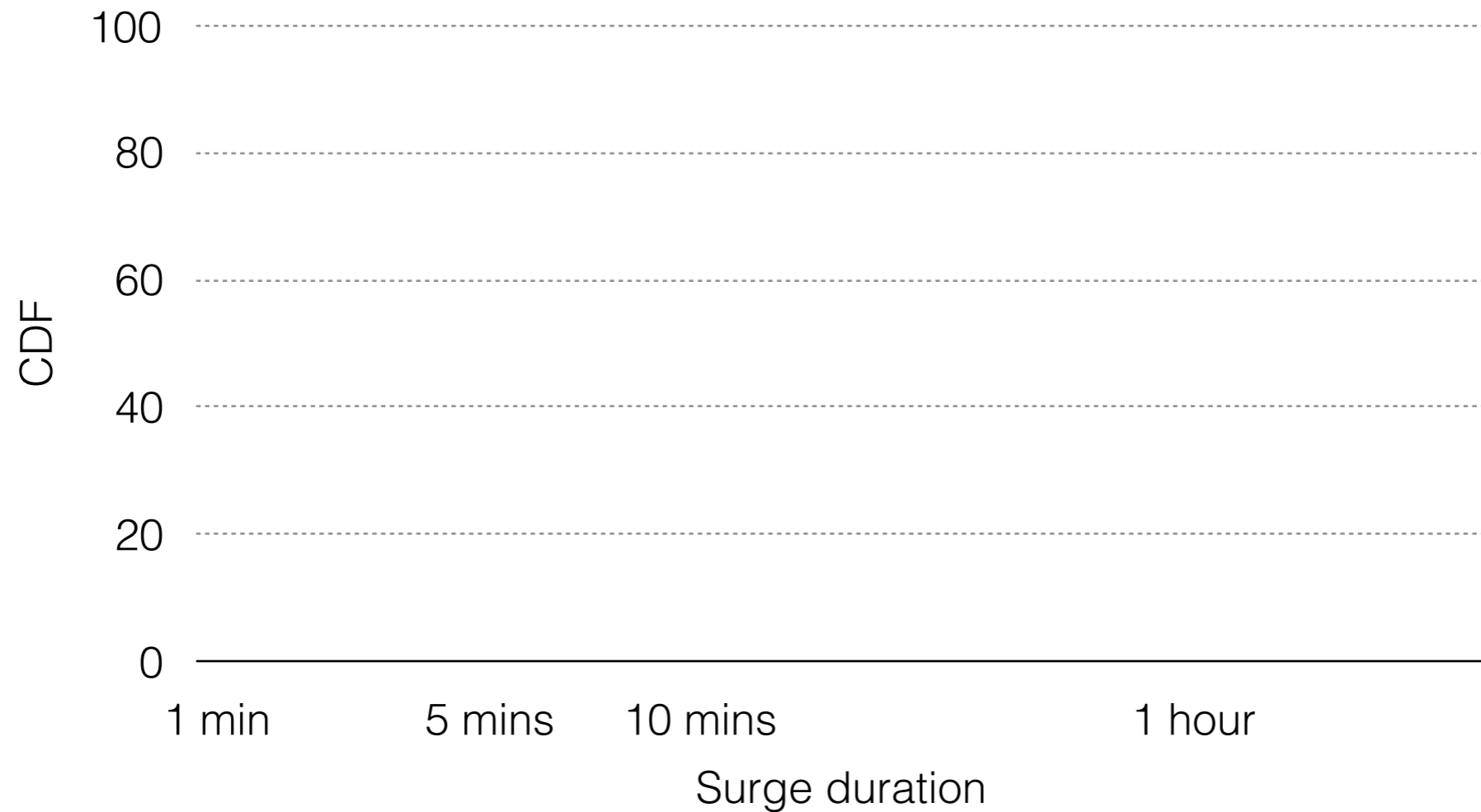
How much and how often does it surge on Uber?



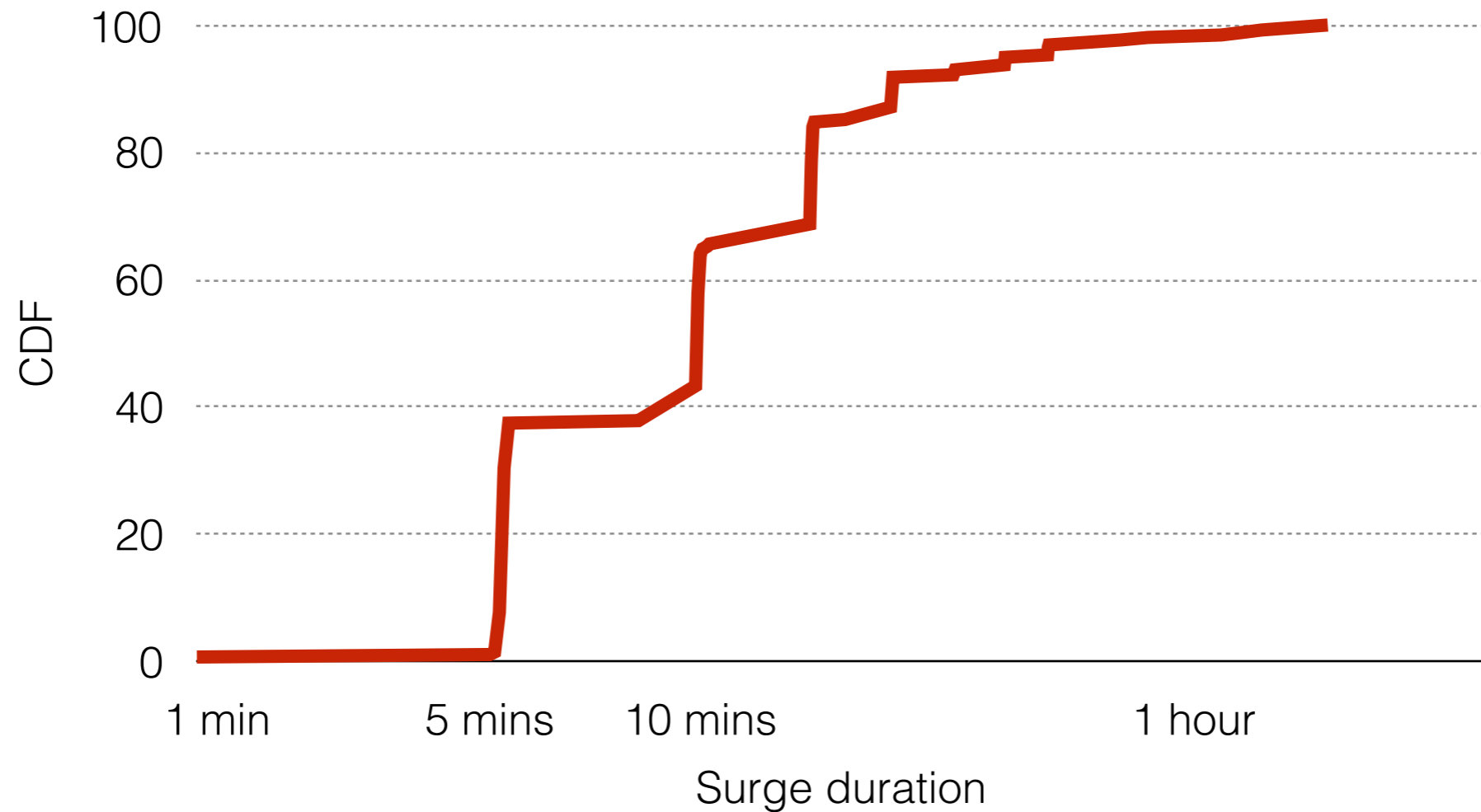
- 14% of time surging in Manhattan
- 57% of time surging in San Francisco
- SF has higher surge values

How long do surges last?

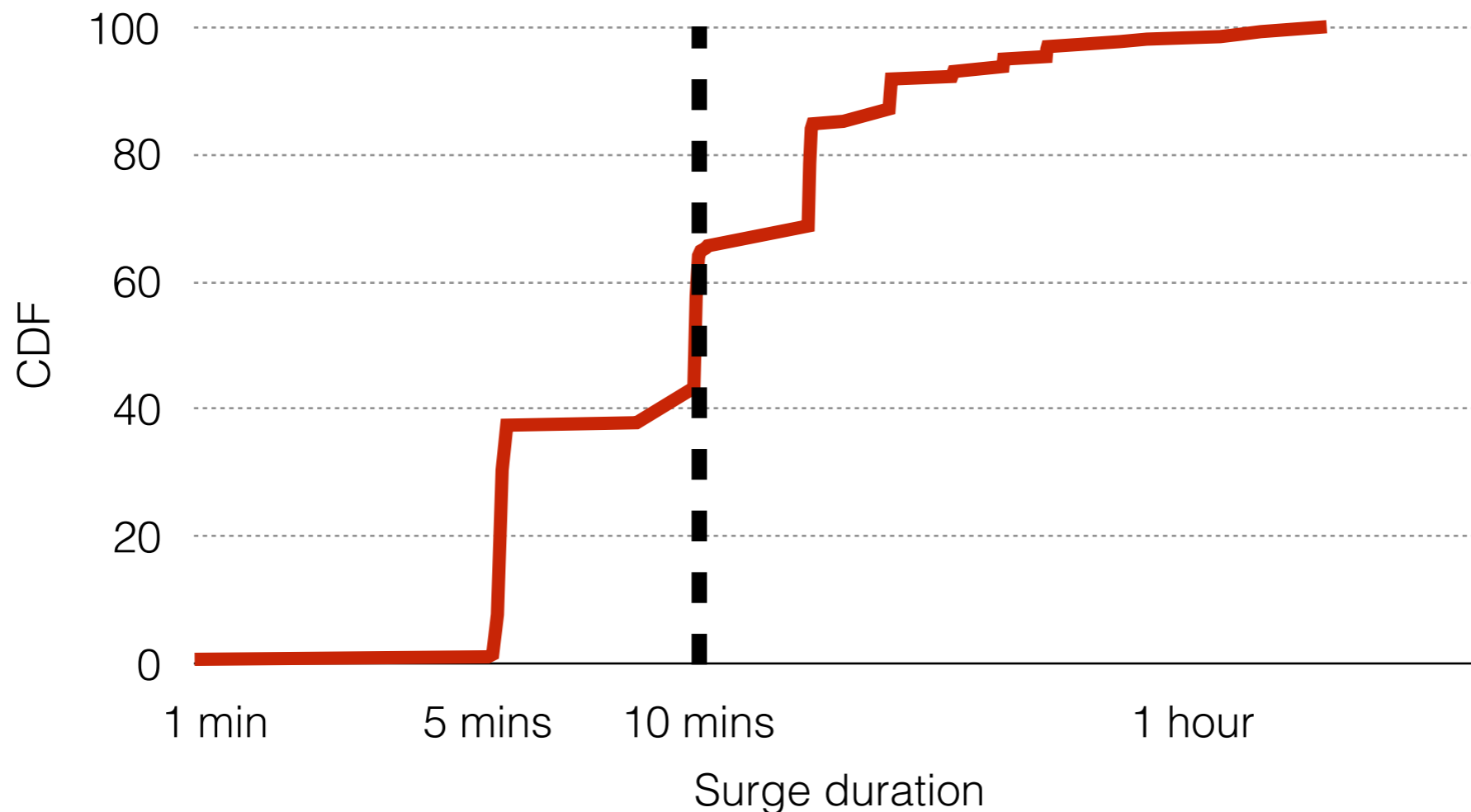
How long do surges last?



How long do surges last?

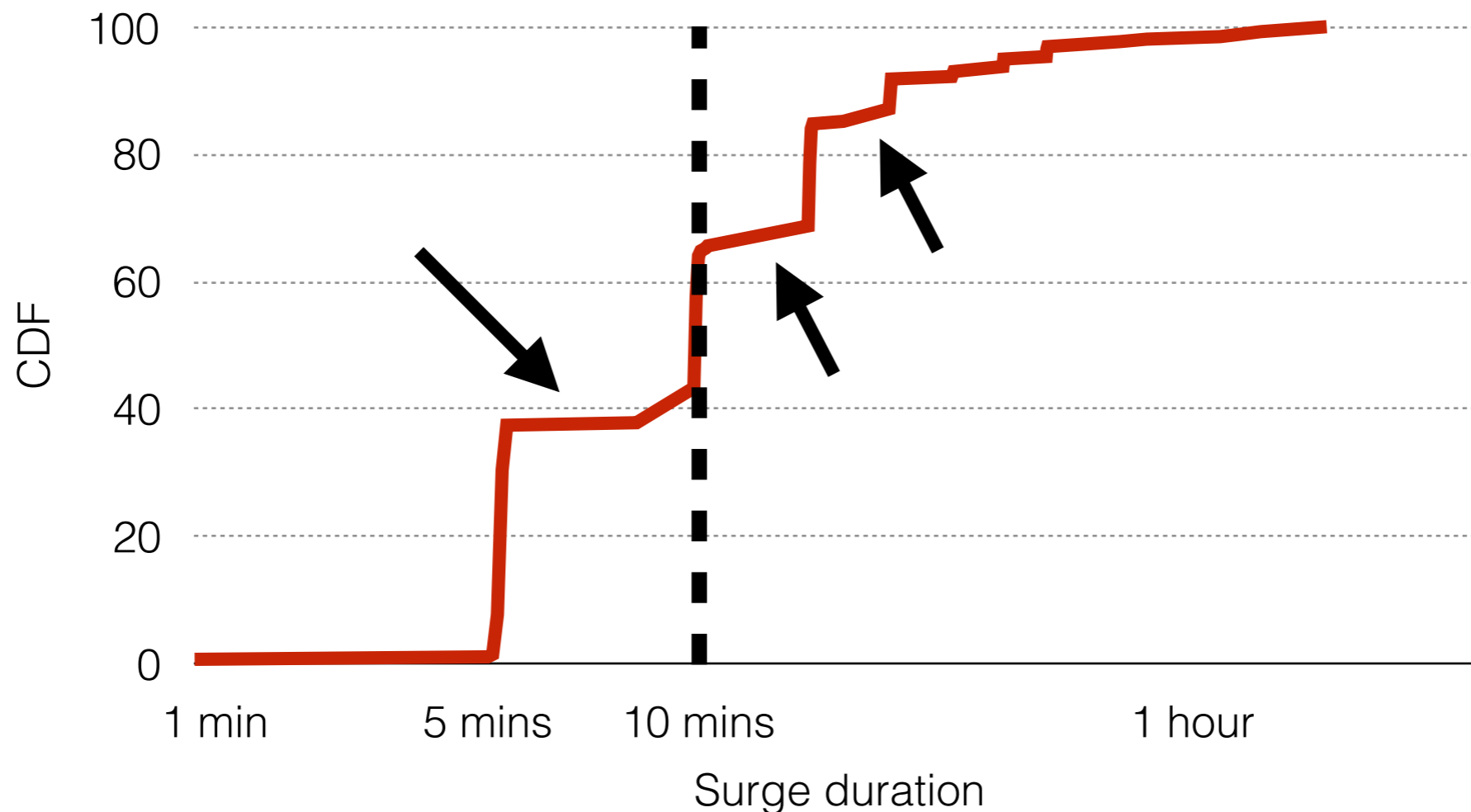


How long do surges last?



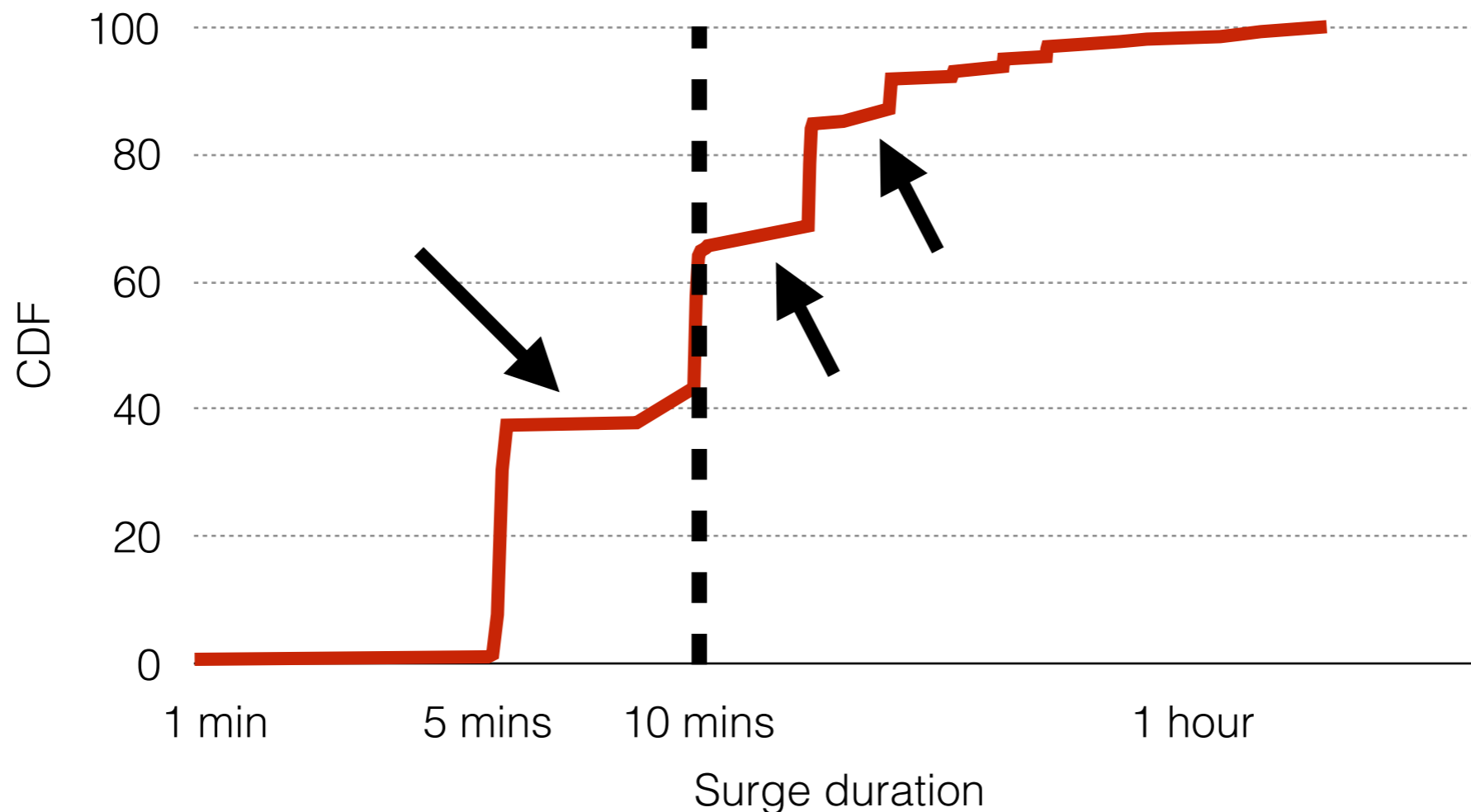
- Noisiness: 70% of surges last less than 10 mins

How long do surges last?



- Noisiness: 70% of surges last less than 10 mins
- Staircase CDF: multiple times of 5 minutes

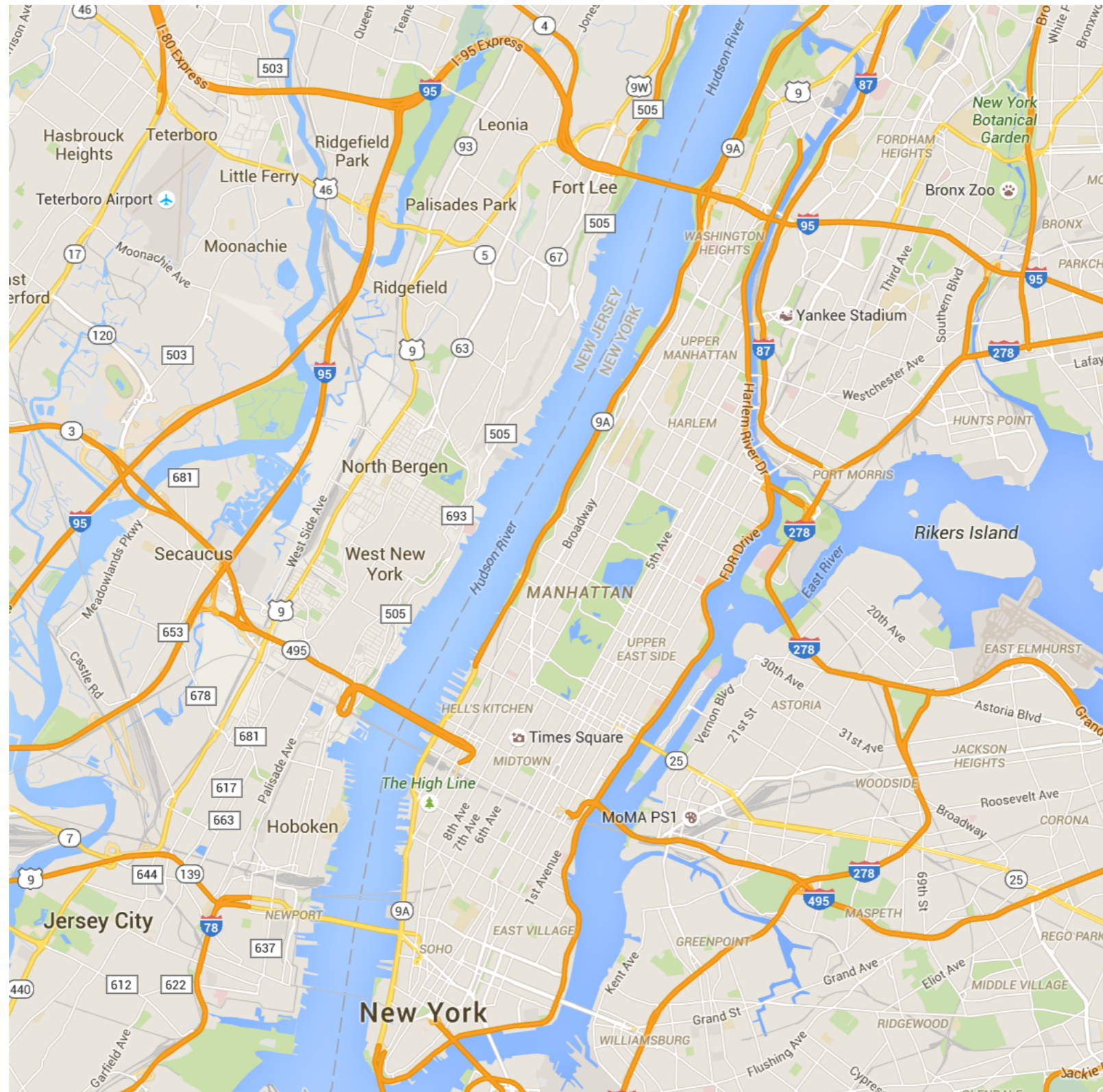
How long do surges last?



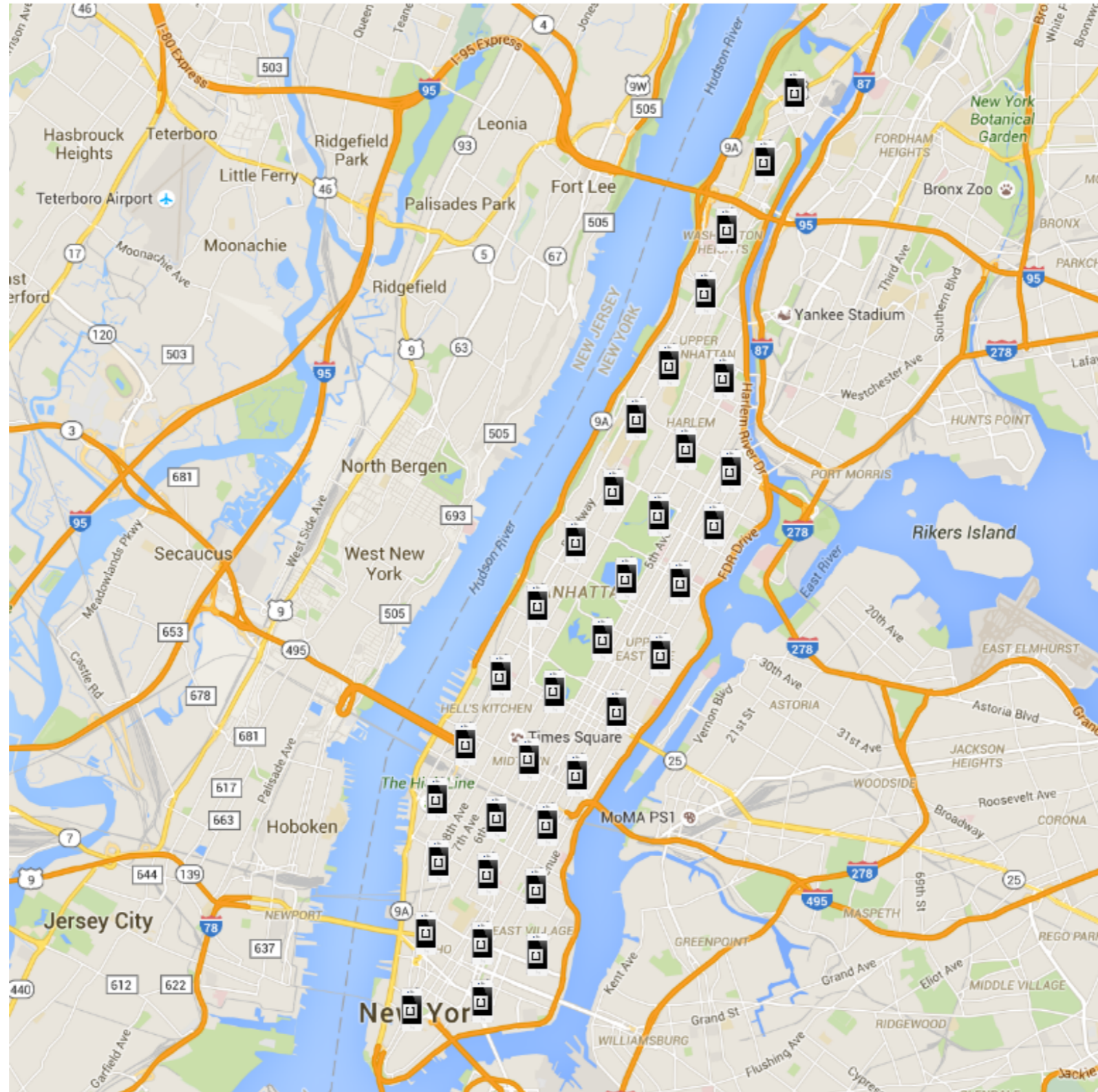
- Noisiness: 70% of surges last less than 10 mins
- Staircase CDF: multiple times of 5 minutes
- Uber updates surge values every 5 minutes

How do surge prices vary by location?

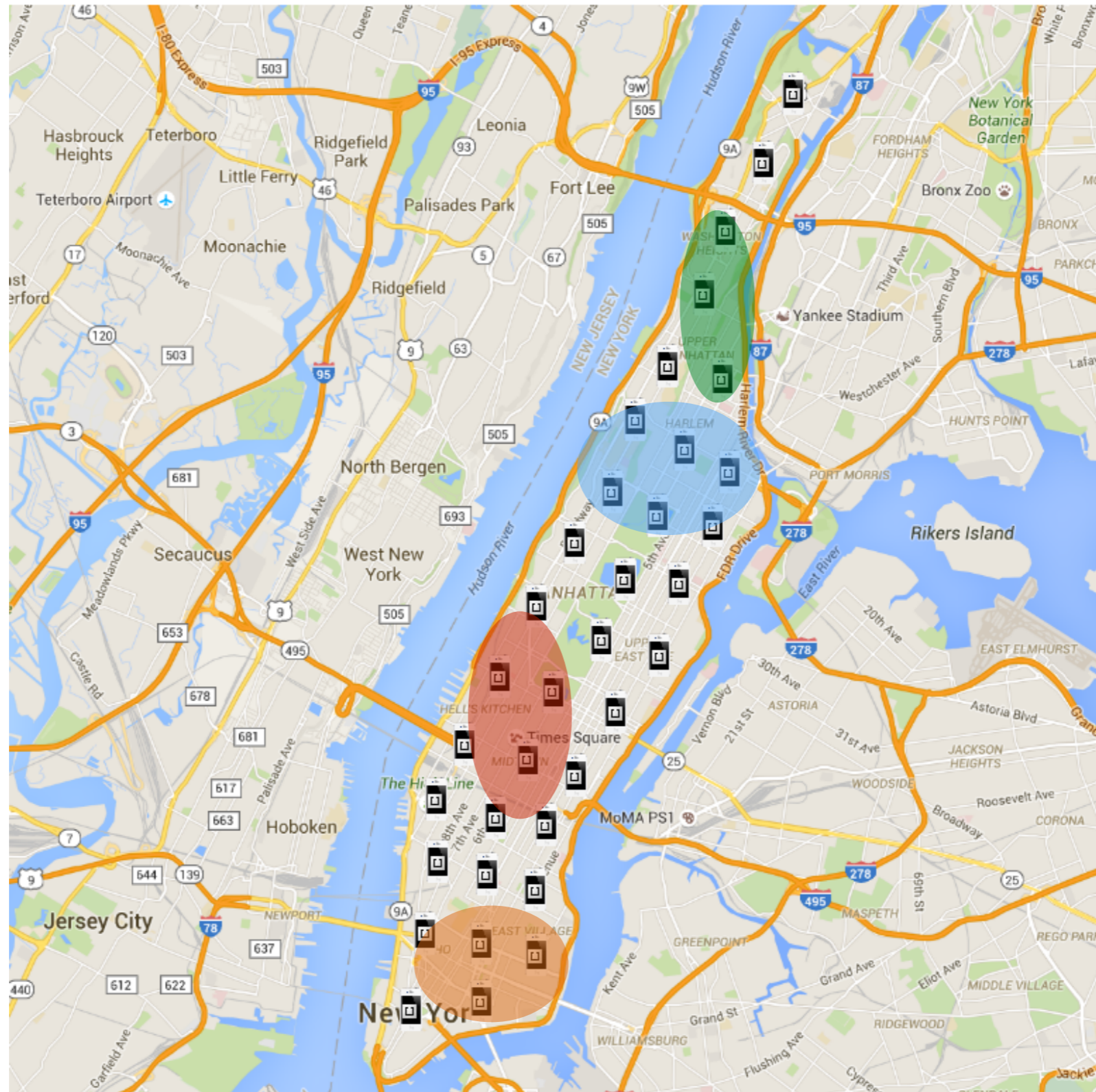
How do surge prices vary by location?



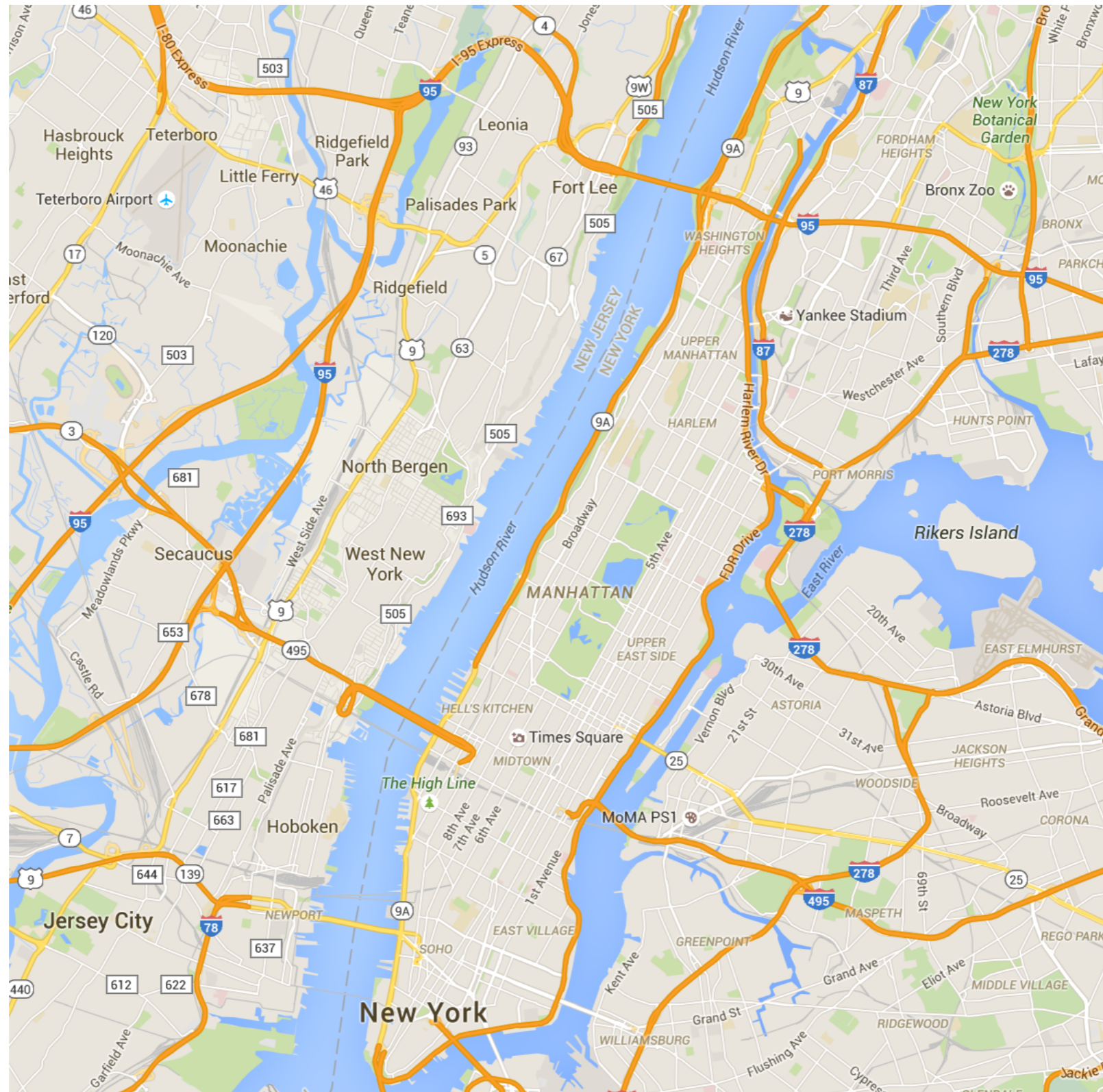
How do surge prices vary by location?



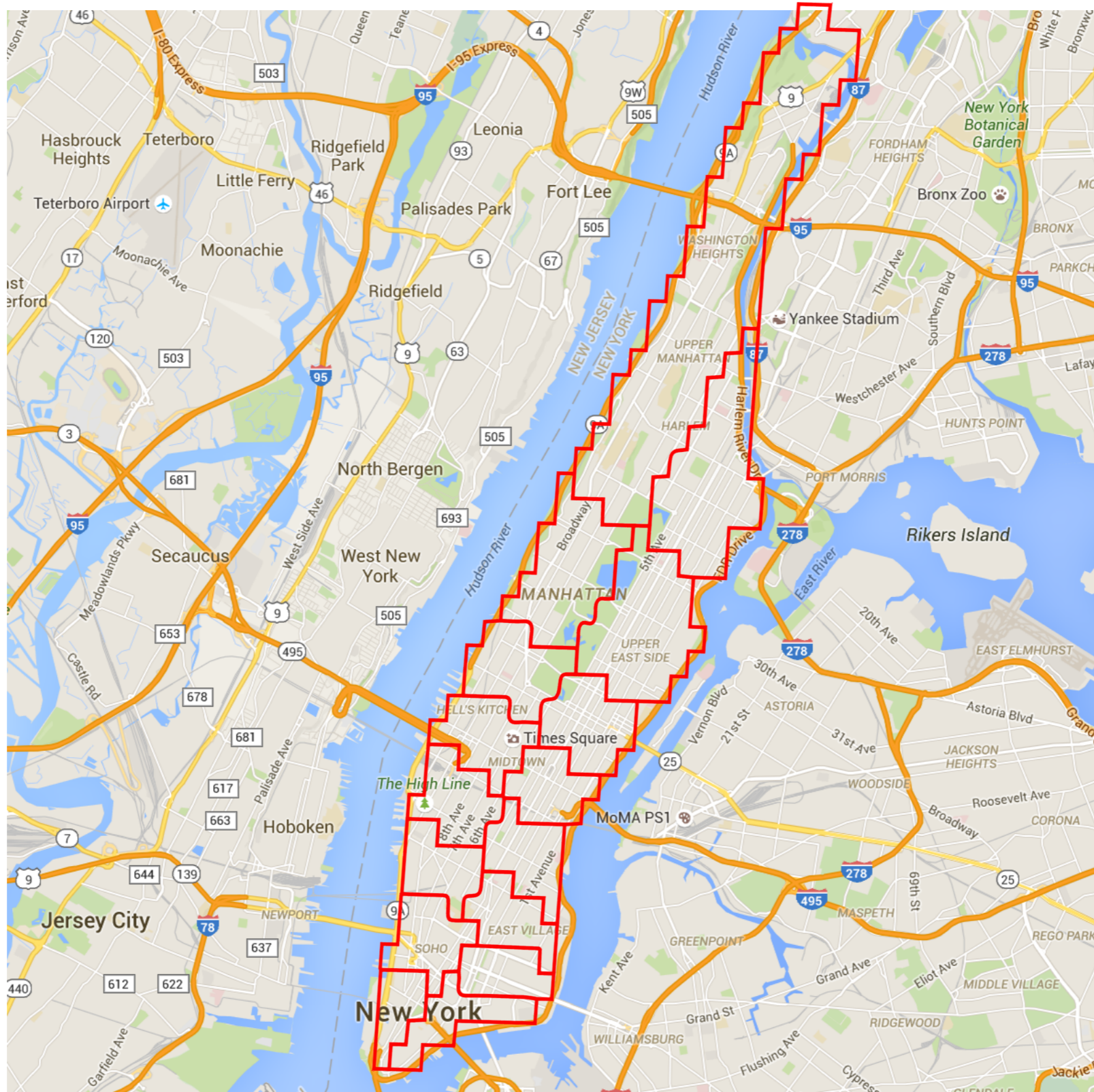
How do surge prices vary by location?



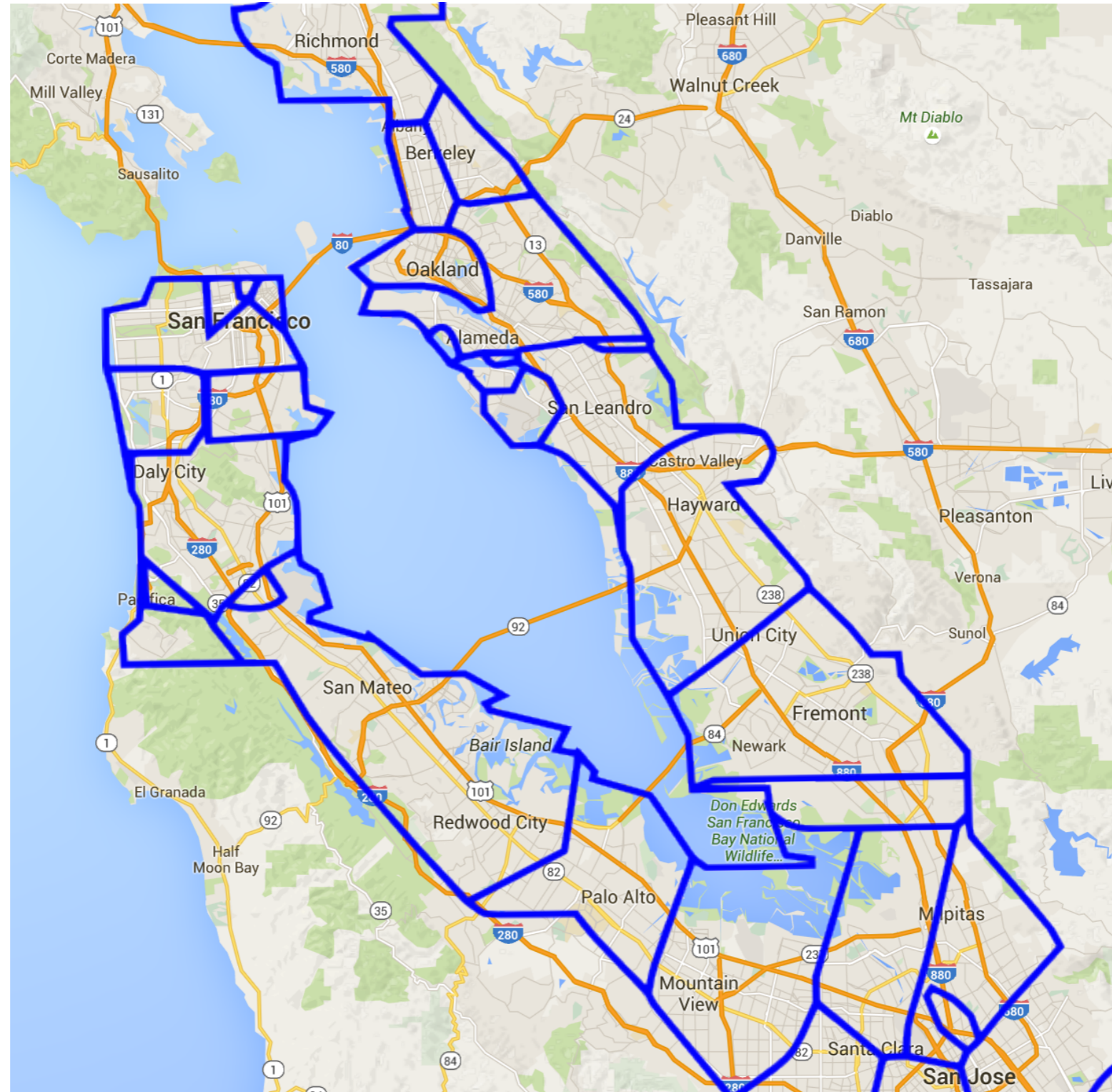
How do surge prices vary by location?



How do surge prices vary by location?



How do surge prices vary by location?



Surge areas

Surge areas

- Chicago
- Los Angeles
- Miami
- Seattle
- NYC, including all 5 boroughs
- Boston
- DC
- London
- Paris

What features does Uber use to calculate surge multipliers?

What features does Uber use to calculate surge multipliers?

- Find variables that correlate with surge pricing

What features does Uber use to calculate surge multipliers?


- Find variables that correlate with surge pricing
 - Supply, demand, EWT, etc.

What features does Uber use to calculate surge multipliers?

- Find variables that correlate with surge pricing
 - Supply, demand, EWT, etc.
 - Cross-correlation

What features does Uber use to calculate surge multipliers?

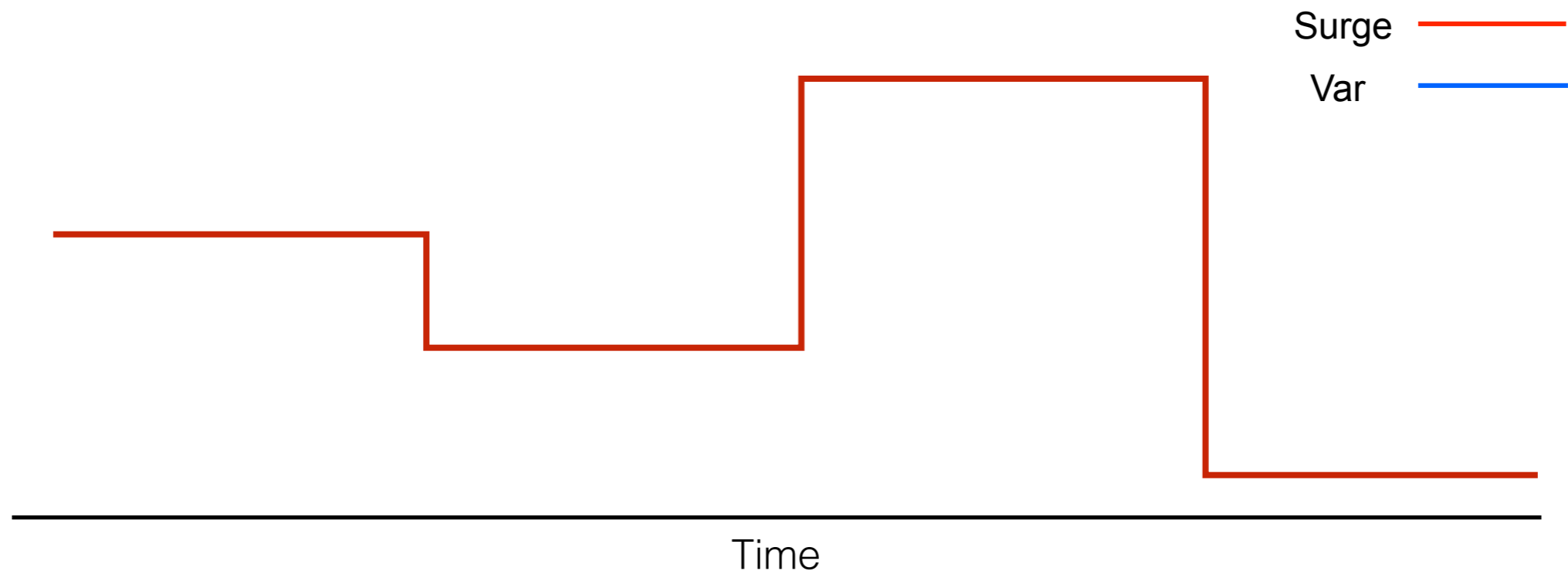
- Find variables that correlate with surge pricing
 - Supply, demand, EWT, etc.
 - Cross-correlation

Surge 
Var 

Time

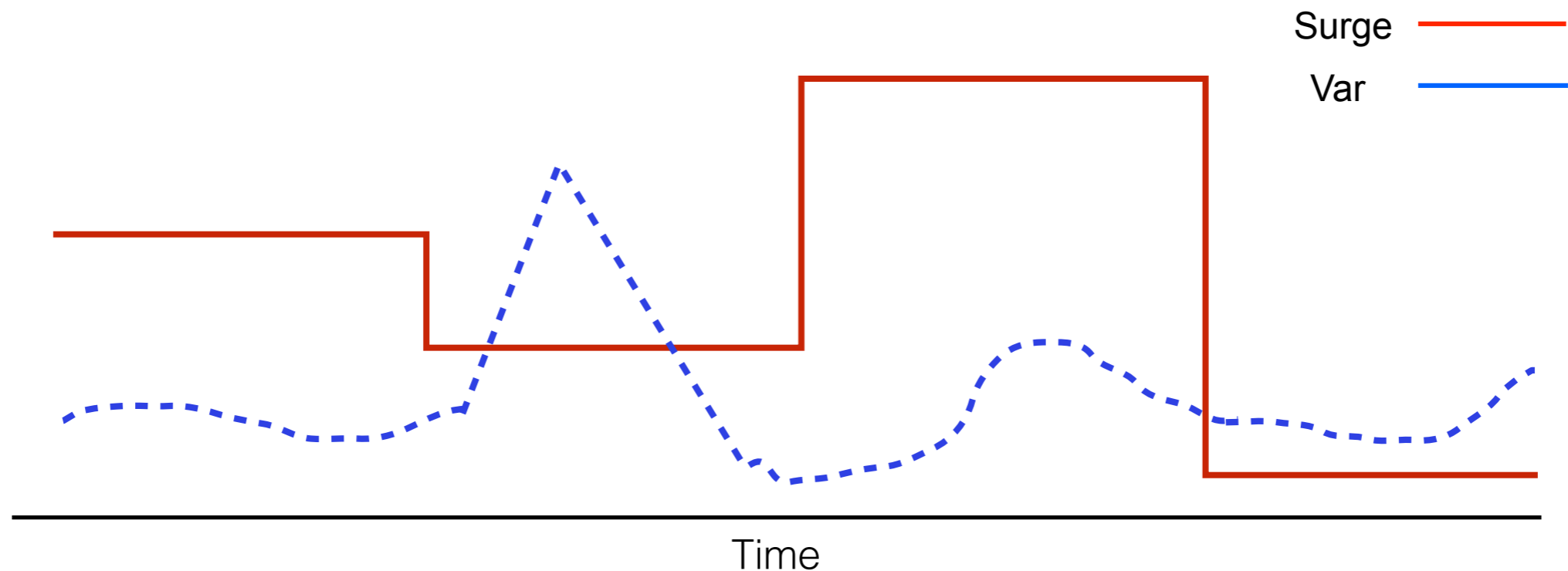
What features does Uber use to calculate surge multipliers?

- Find variables that correlate with surge pricing
 - Supply, demand, EWT, etc.
 - Cross-correlation



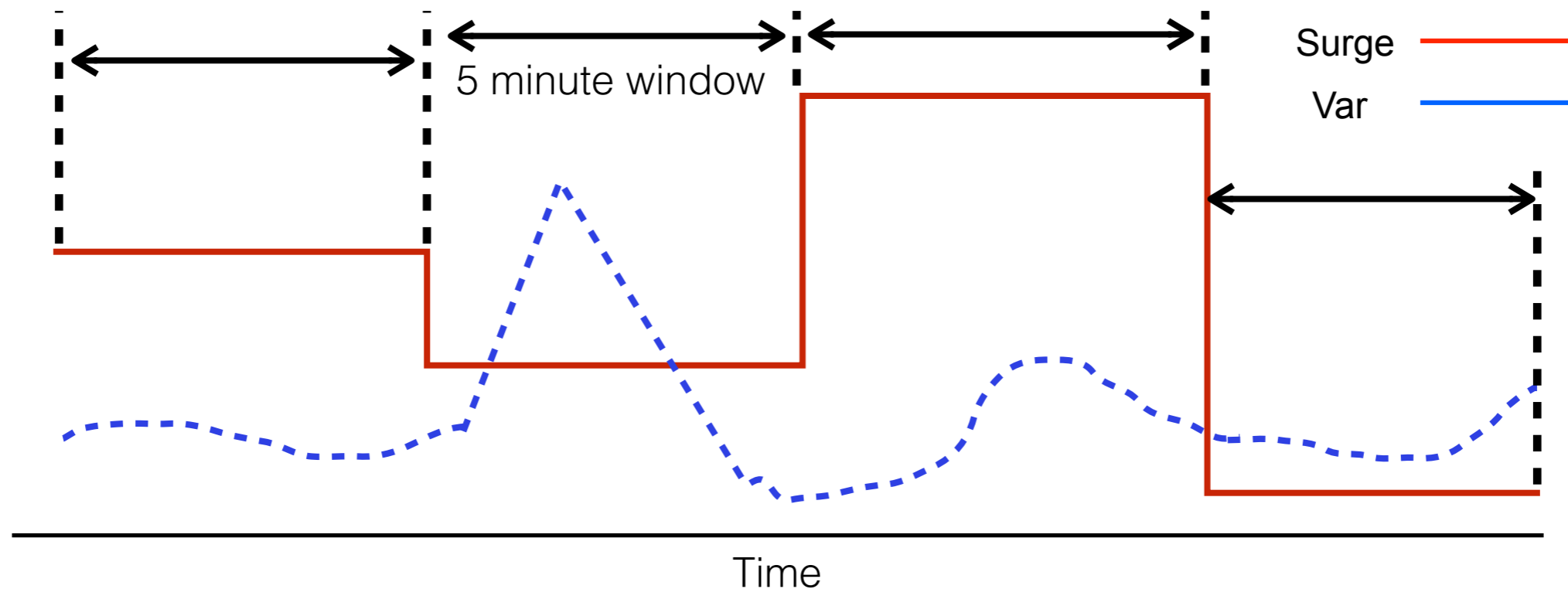
What features does Uber use to calculate surge multipliers?

- Find variables that correlate with surge pricing
 - Supply, demand, EWT, etc.
 - Cross-correlation



What features does Uber use to calculate surge multipliers?

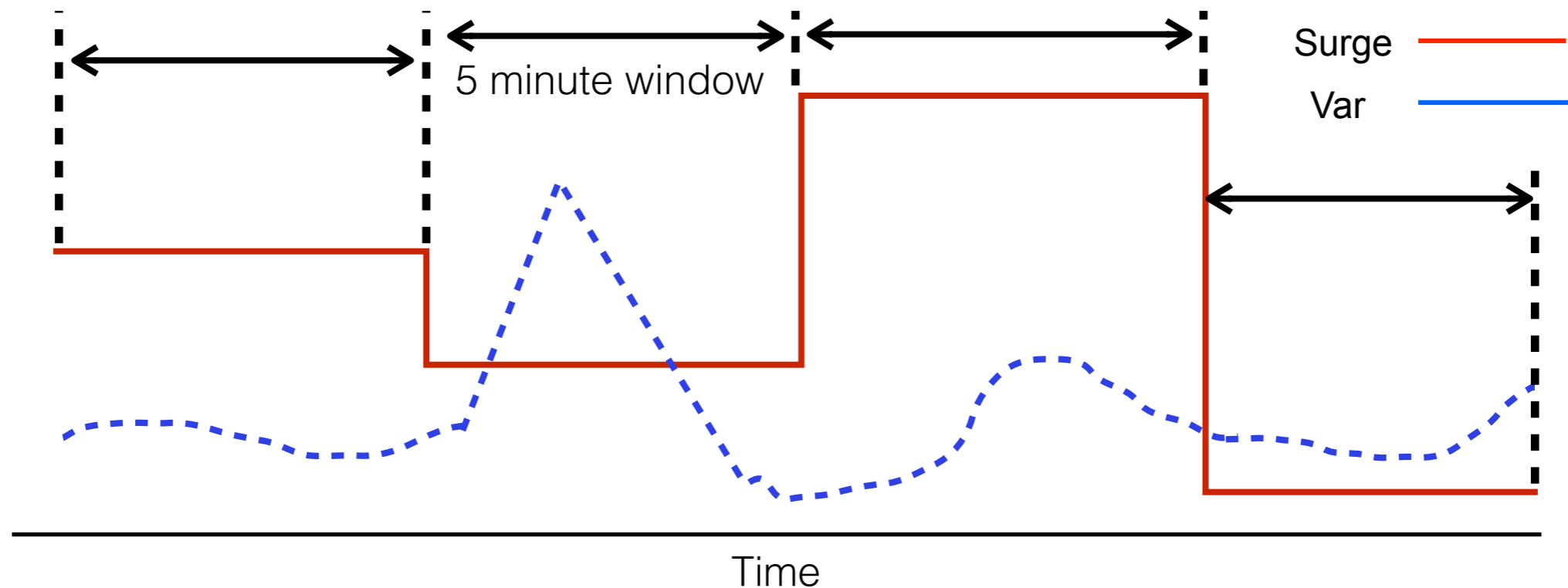
- Find variables that correlate with surge pricing
 - Supply, demand, EWT, etc.
 - Cross-correlation



What features does Uber use to calculate surge multipliers?

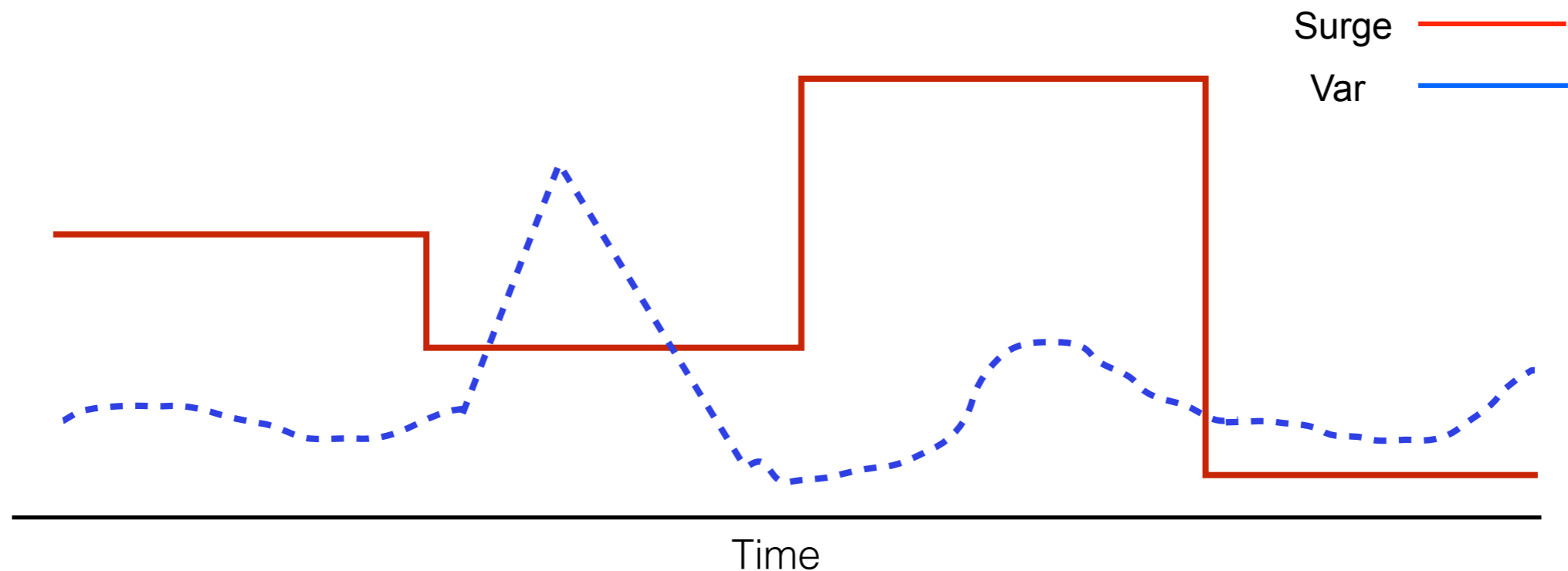
- Find variables that correlate with surge pricing
 - Supply, demand, EWT, etc.

- Cross-correlation **Time Difference = 0**



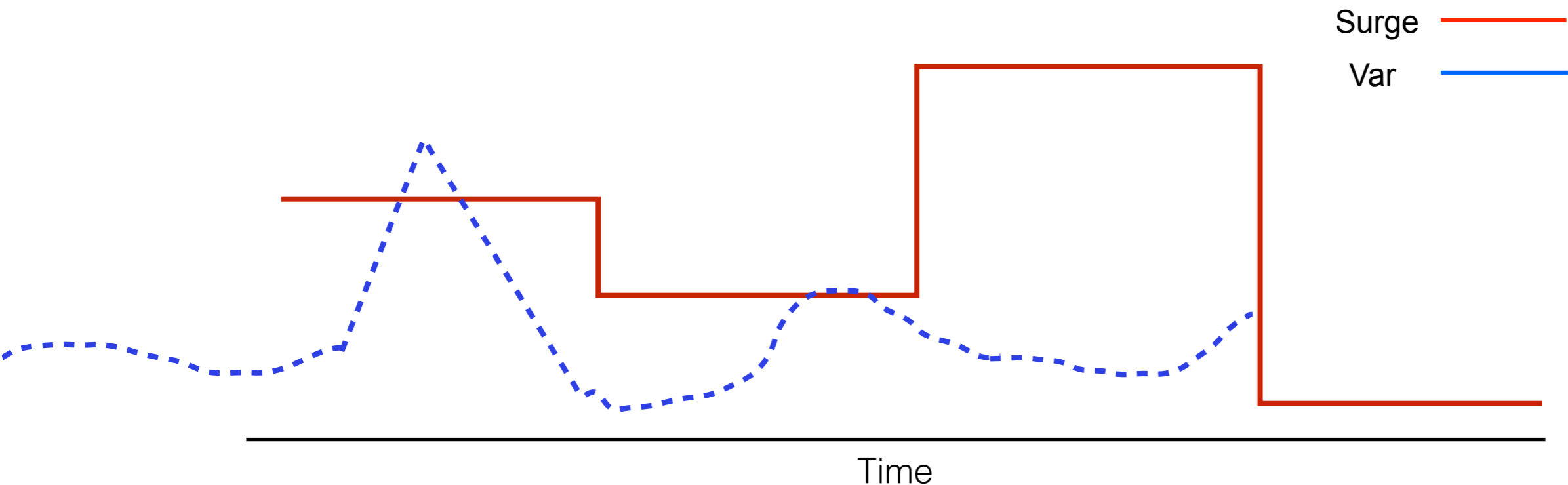
What features does Uber use to calculate surge multipliers?

- Find variables that correlate with surge pricing
 - Supply, demand, EWT, etc.
 - Cross-correlation **Time Difference = 0**



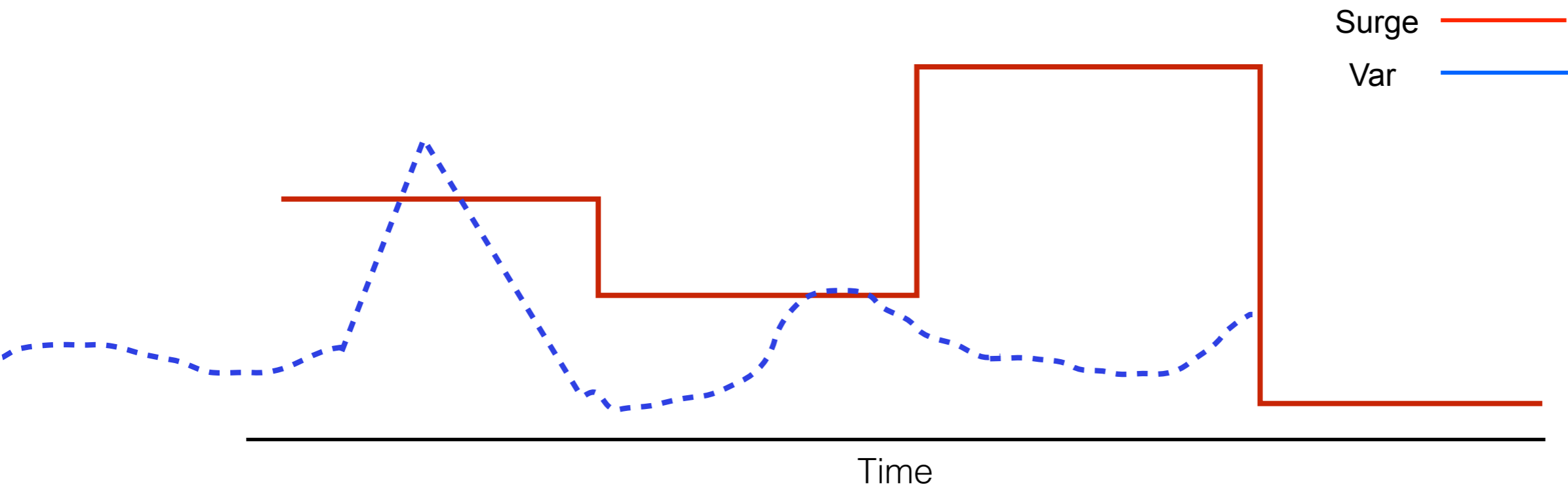
What features does Uber use to calculate surge multipliers?

- Find variables that correlate with surge pricing
 - Supply, demand, EWT, etc.
 - Cross-correlation **Time Difference = 0**



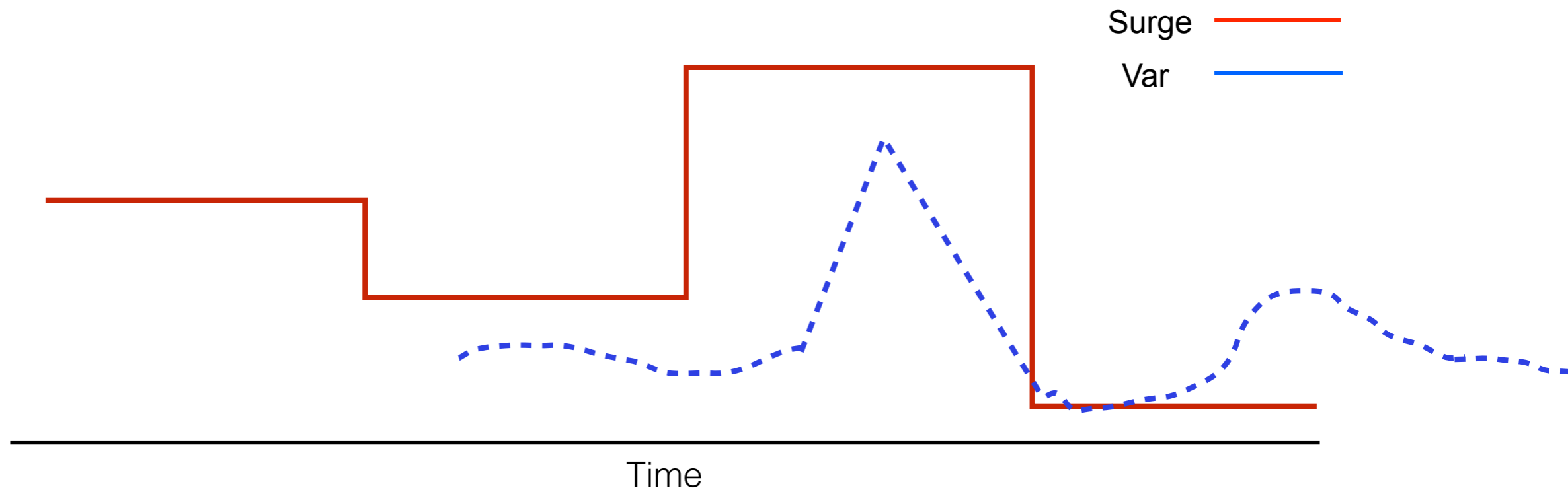
What features does Uber use to calculate surge multipliers?

- Find variables that correlate with surge pricing
 - Supply, demand, EWT, etc.
 - Cross-correlation **Time Difference = -5**



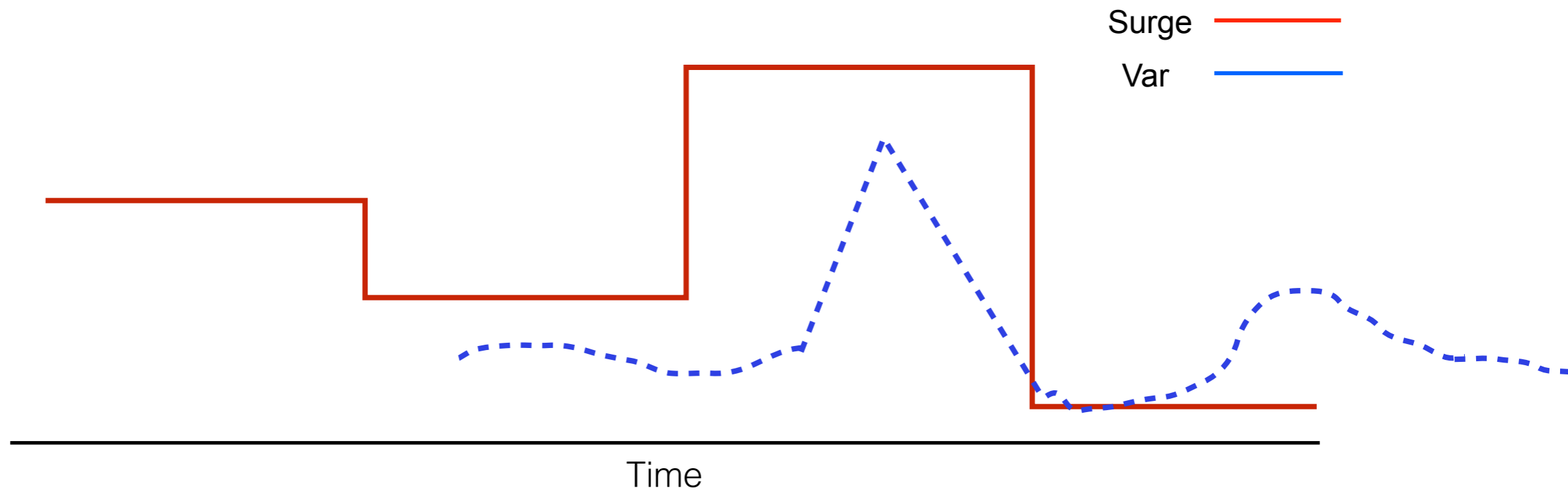
What features does Uber use to calculate surge multipliers?

- Find variables that correlate with surge pricing
 - Supply, demand, EWT, etc.
 - Cross-correlation **Time Difference = -5**



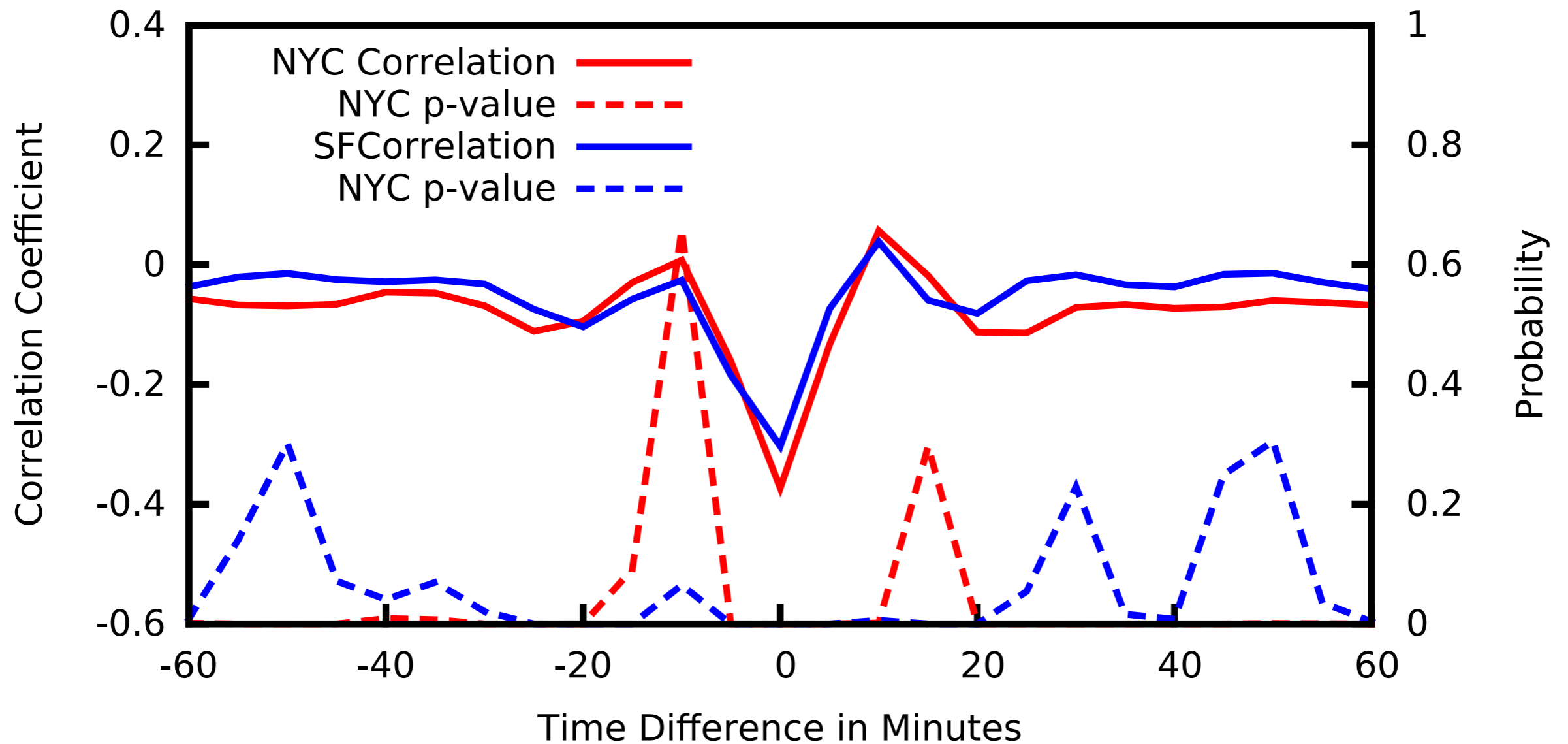
What features does Uber use to calculate surge multipliers?

- Find variables that correlate with surge pricing
 - Supply, demand, EWT, etc.
 - Cross-correlation **Time Difference = +5**

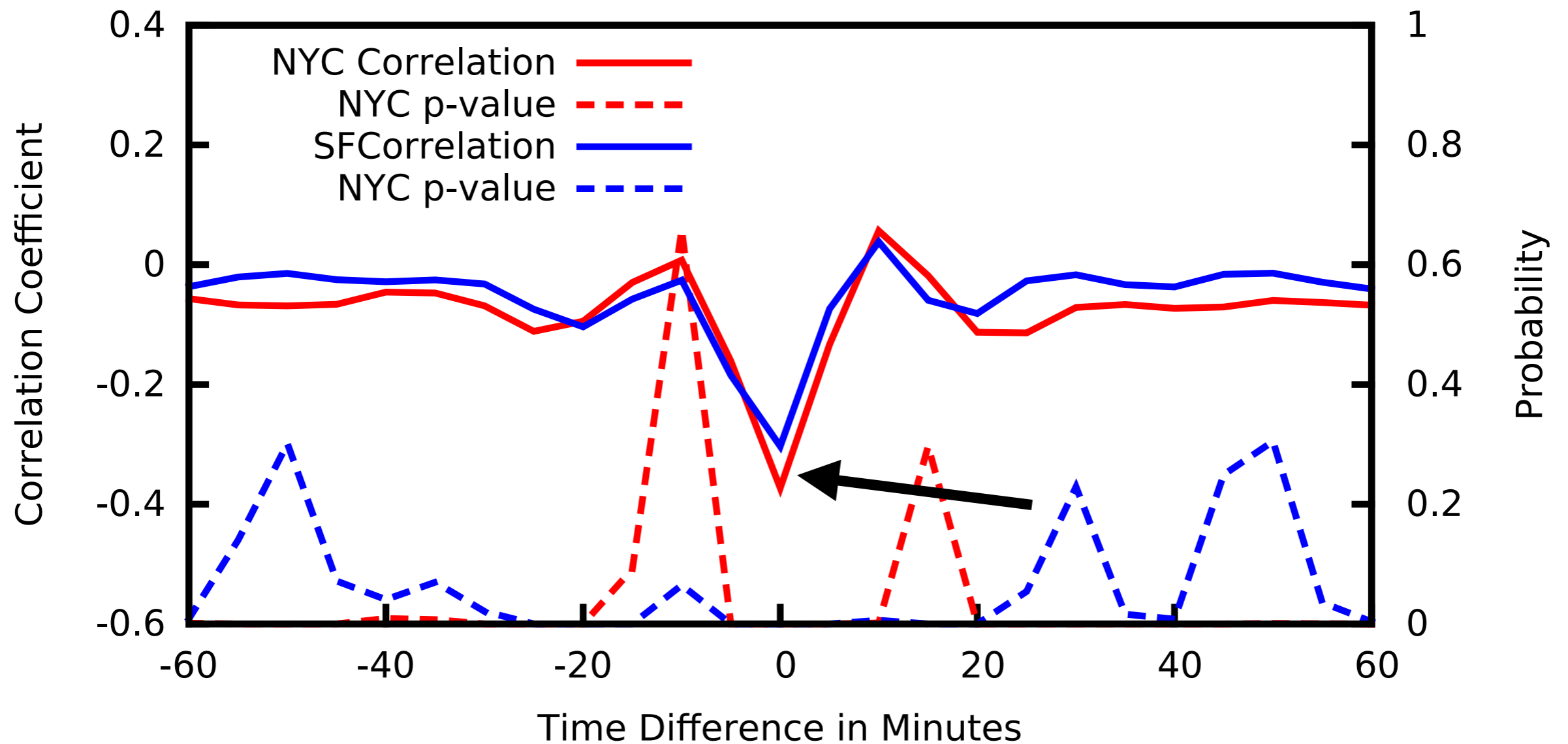


(Supply - demand) vs
surge multipliers

(Supply - demand) vs surge multipliers

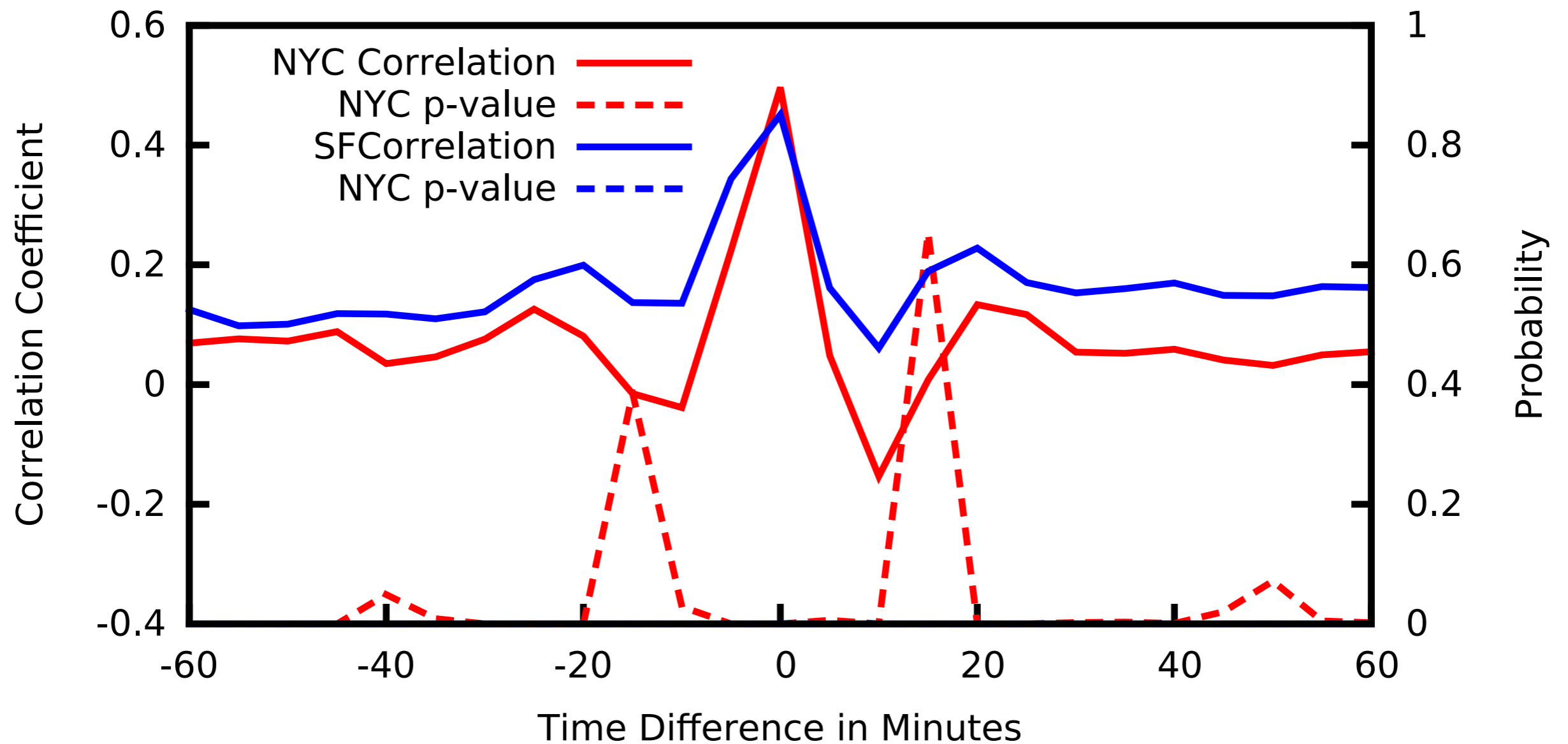


(Supply - demand) vs surge multipliers

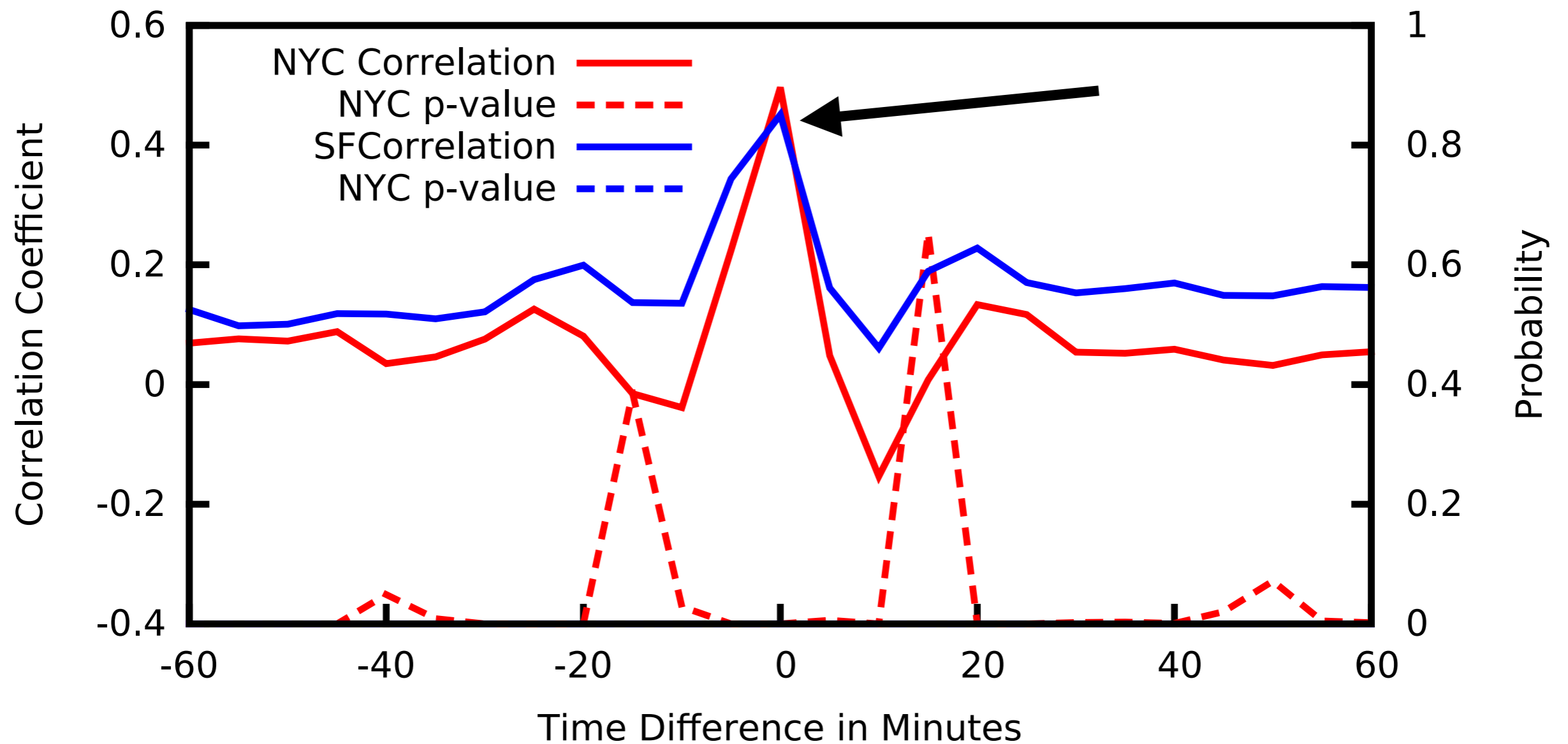


EWT vs surge multipliers

EWT vs surge multipliers

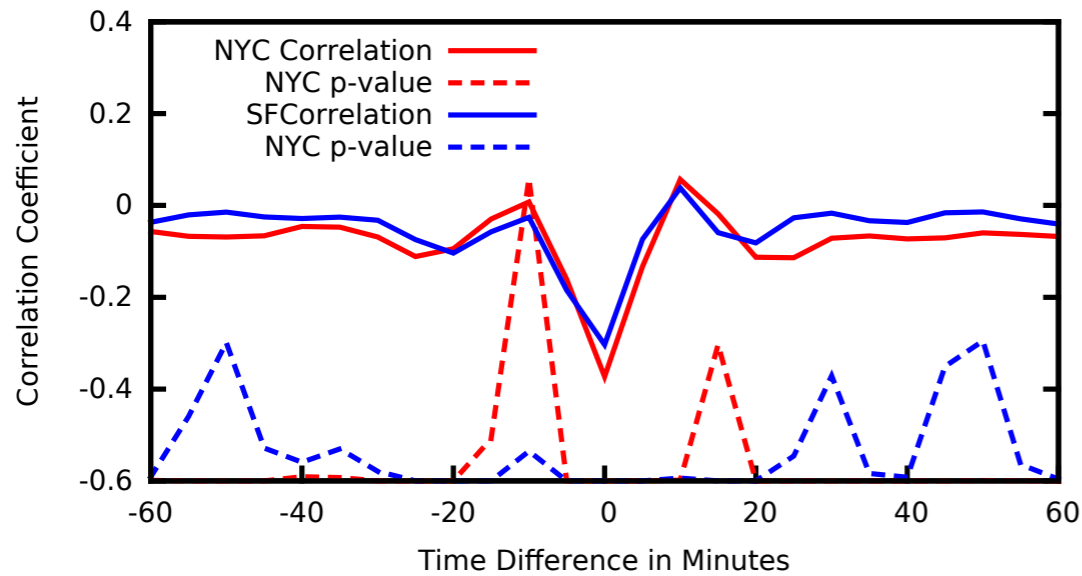


EWT vs surge multipliers

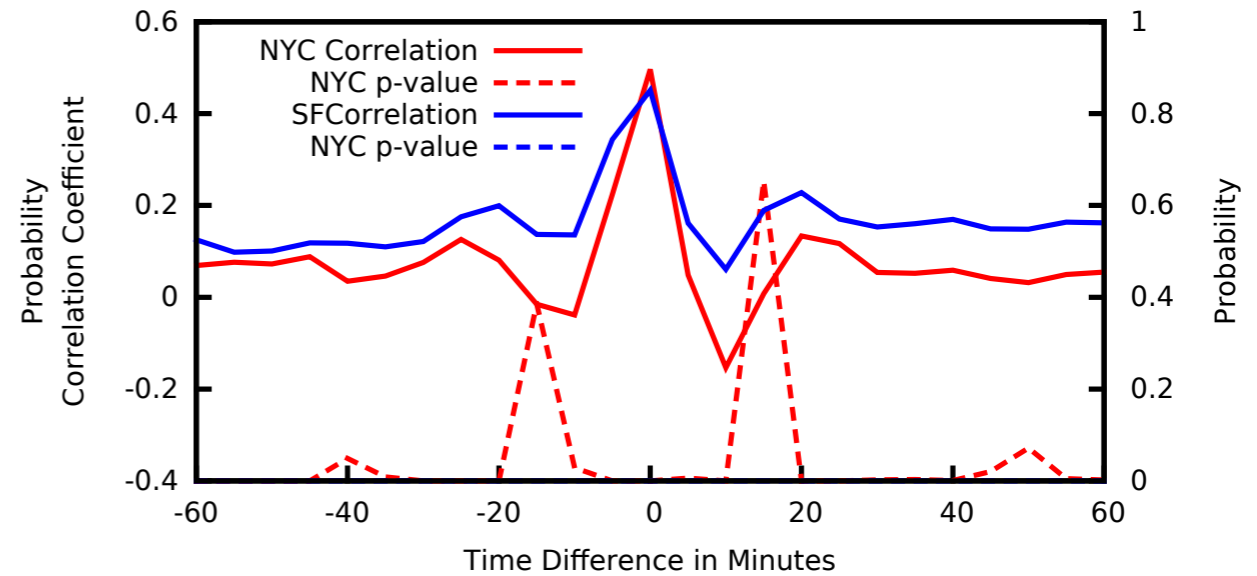


Features correlated with surge multipliers

Features correlated with surge multipliers

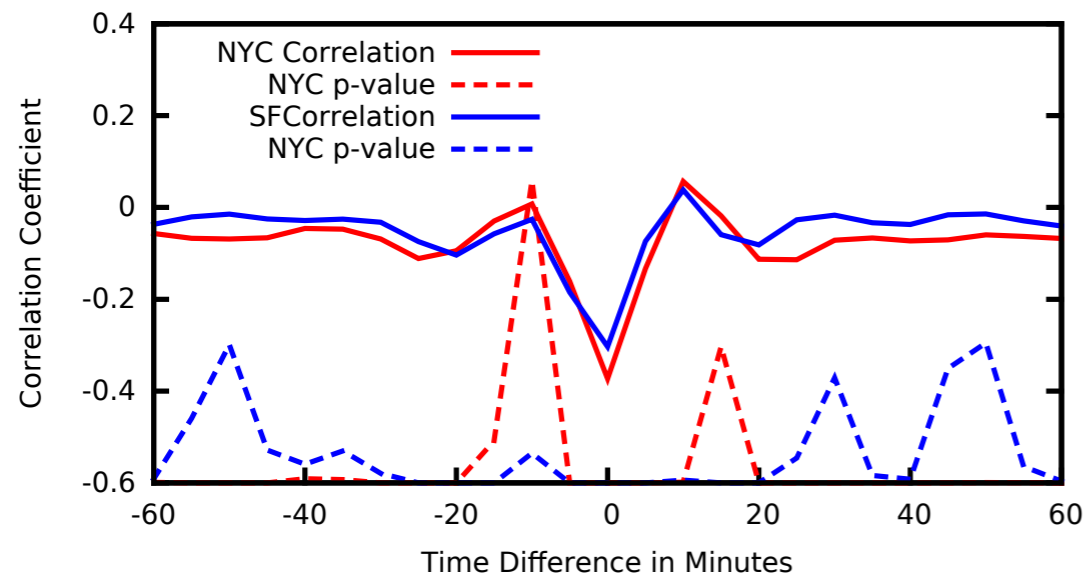


(Supply - demand) vs surge

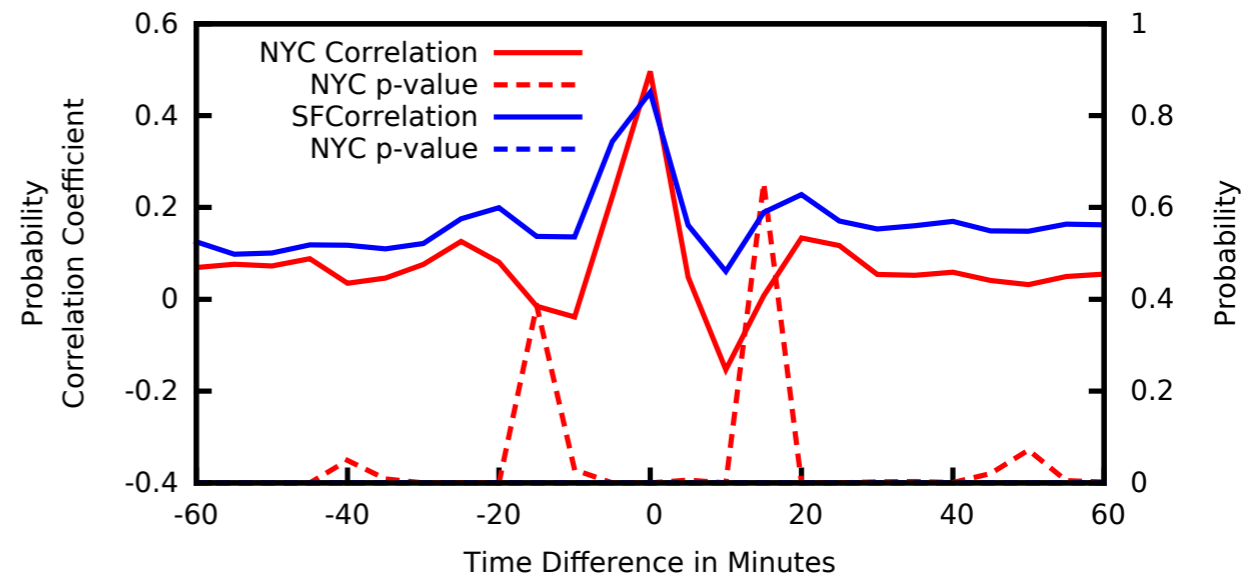


EWT vs surge

Features correlated with surge multipliers



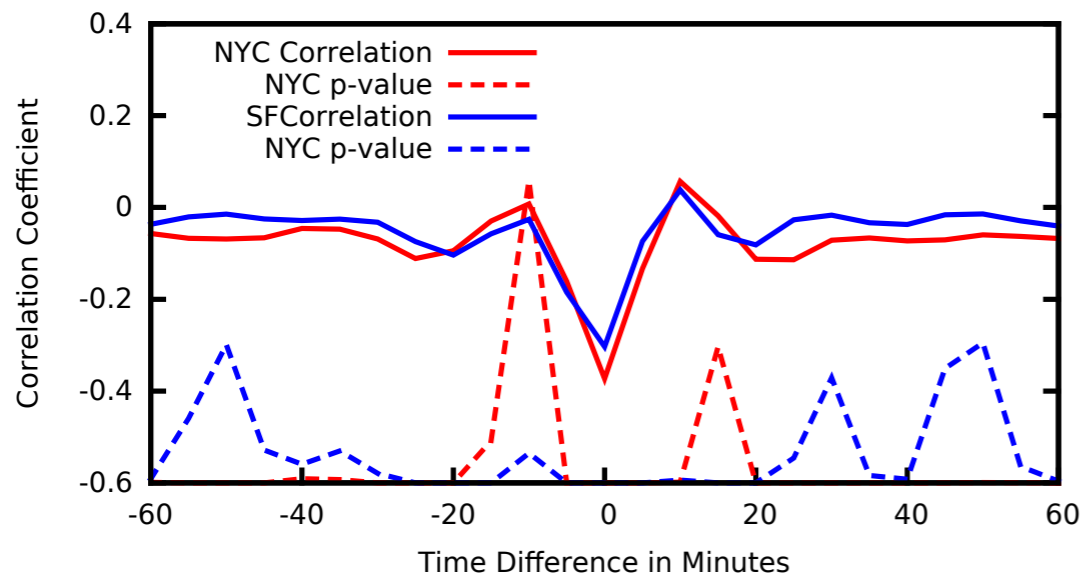
(Supply - demand) vs surge



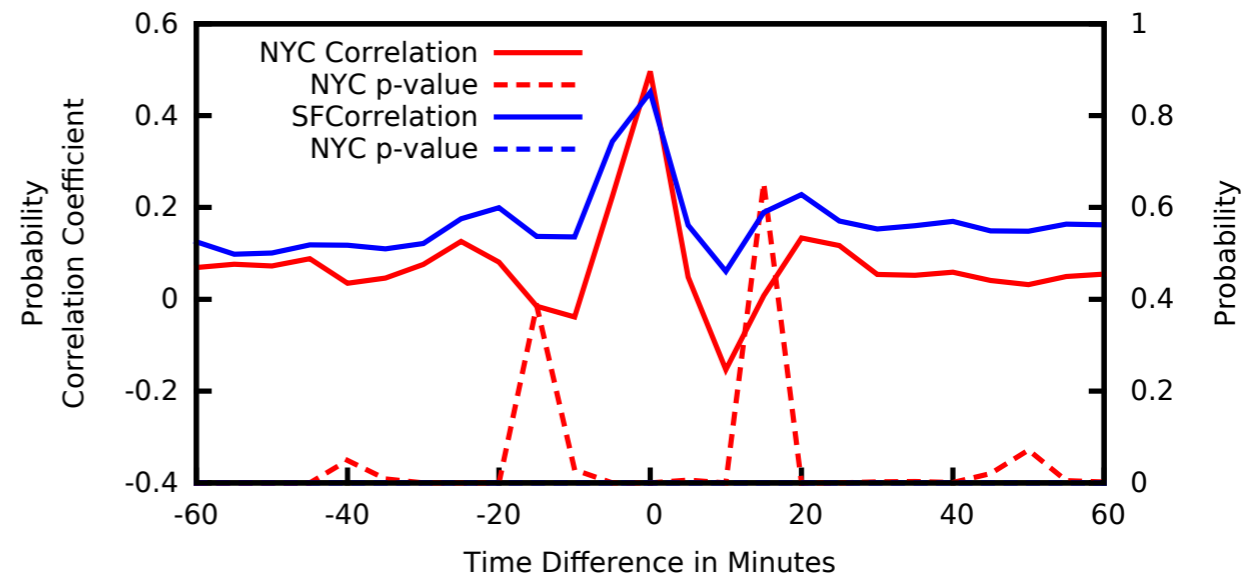
EWT vs surge

- Moderate correlation when time difference is 0

Features correlated with surge multipliers



(Supply - demand) vs surge



EWT vs surge

- Moderate correlation when time difference is 0
- Zero correlation in other windows: responsive but noisy

Can we predict surge
multipliers?

Can we predict surge multipliers?

- Useful variables

Can we predict surge multipliers?

- Useful variables
 - Supply/demand difference, EWT, surge history

Can we predict surge multipliers?

- Useful variables
 - Supply/demand difference, EWT, surge history
- Model

Can we predict surge multipliers?

- Useful variables
 - Supply/demand difference, EWT, surge history
- Model
 - Linear regression

Can we predict surge multipliers?

- Useful variables
 - Supply/demand difference, EWT, surge history
- Model
 - Linear regression
- Performance

Can we predict surge multipliers?

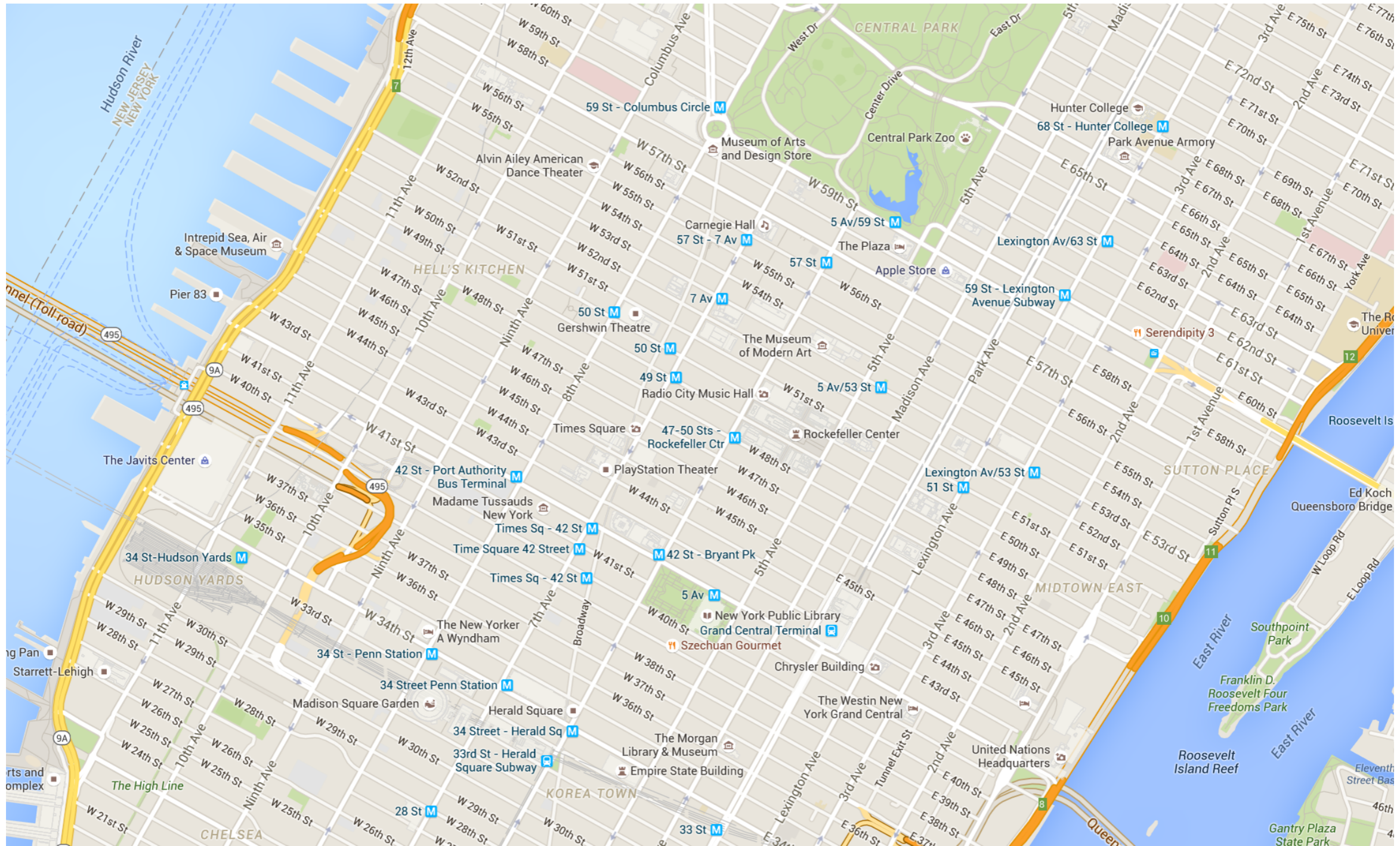
- Useful variables
 - Supply/demand difference, EWT, surge history
- Model
 - Linear regression
- Performance
 - R^2 scores range from 0.37 — 0.57

Can we predict surge multipliers?

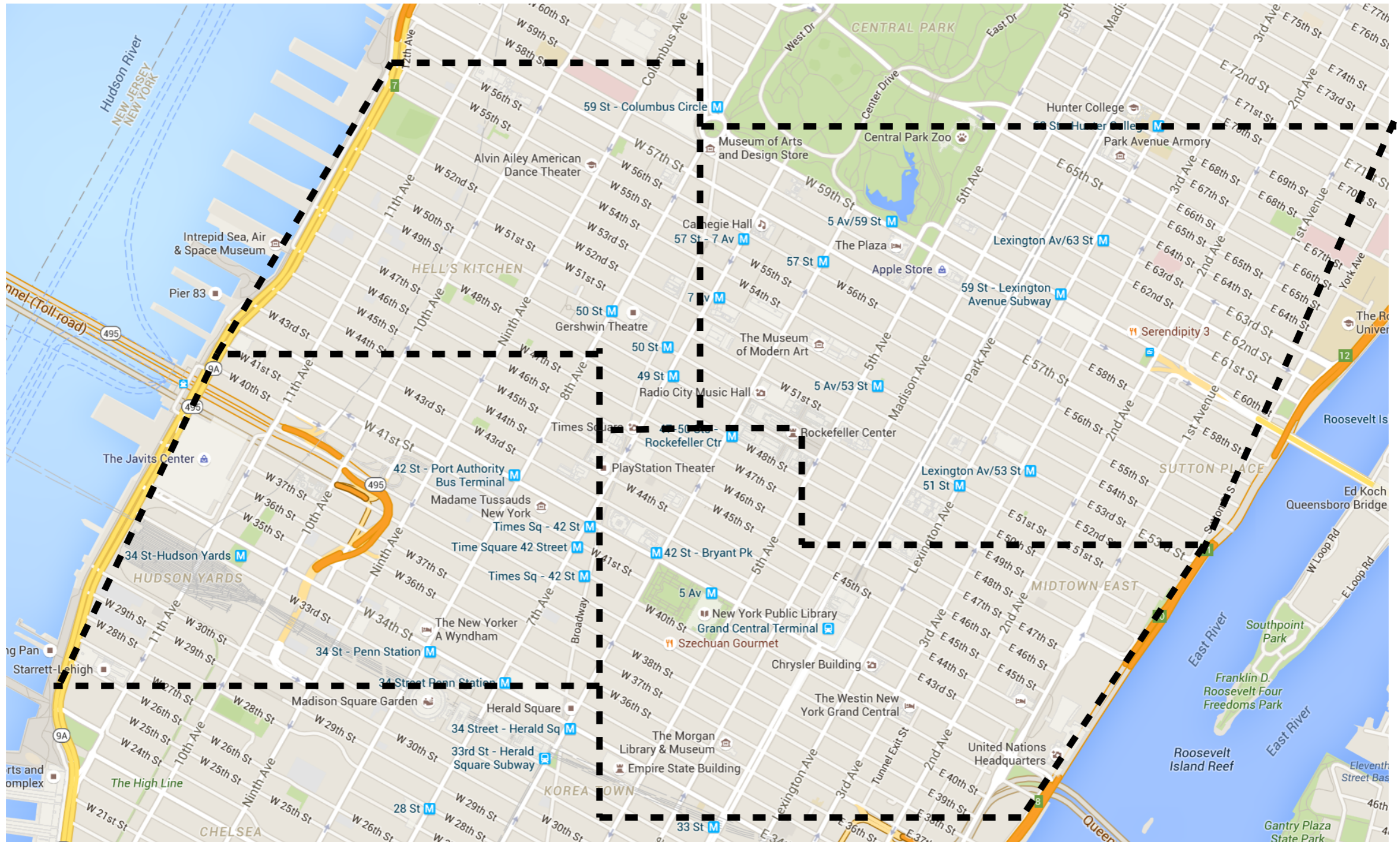
- Useful variables
 - Supply/demand difference, EWT, surge history
- Model
 - Linear regression
- Performance
 - R^2 scores range from 0.37 — 0.57
 - Poor predicative power: missing variable (unfulfilled demand)

Can we avoid surge pricing?

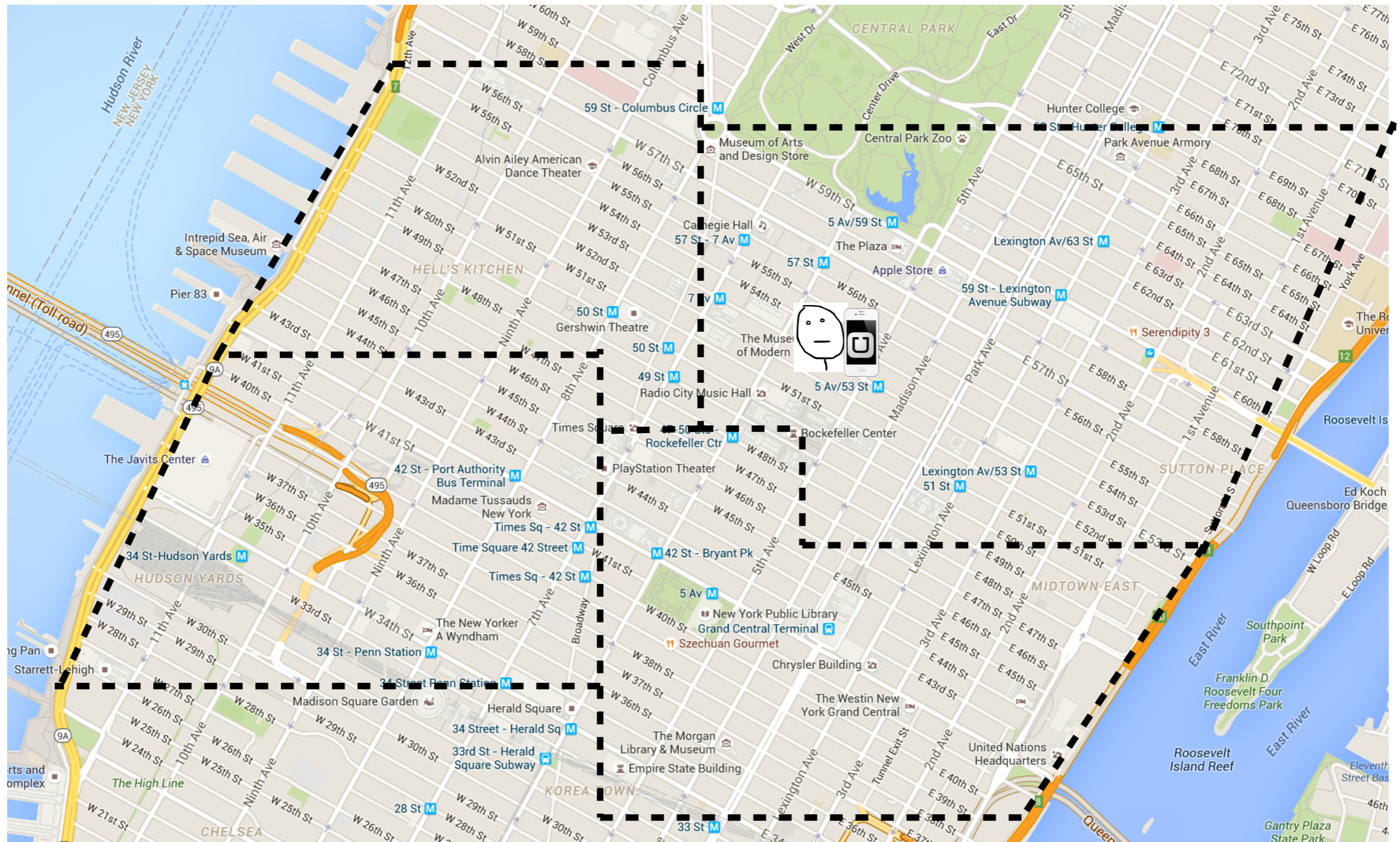
Can we avoid surge pricing?



Can we avoid surge pricing?



Can we avoid surge pricing?



Can we avoid surge pricing?



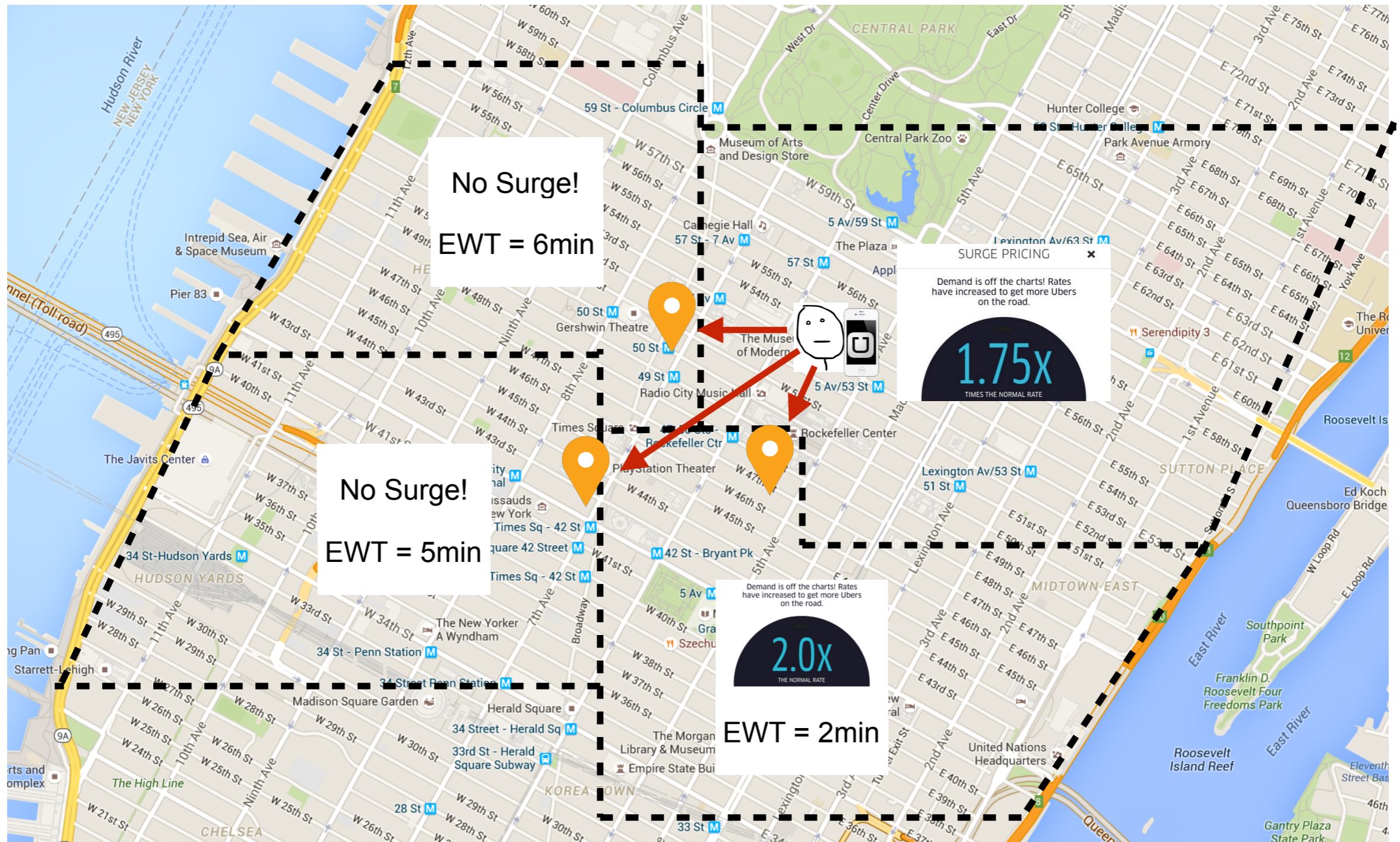
Can we avoid surge pricing?



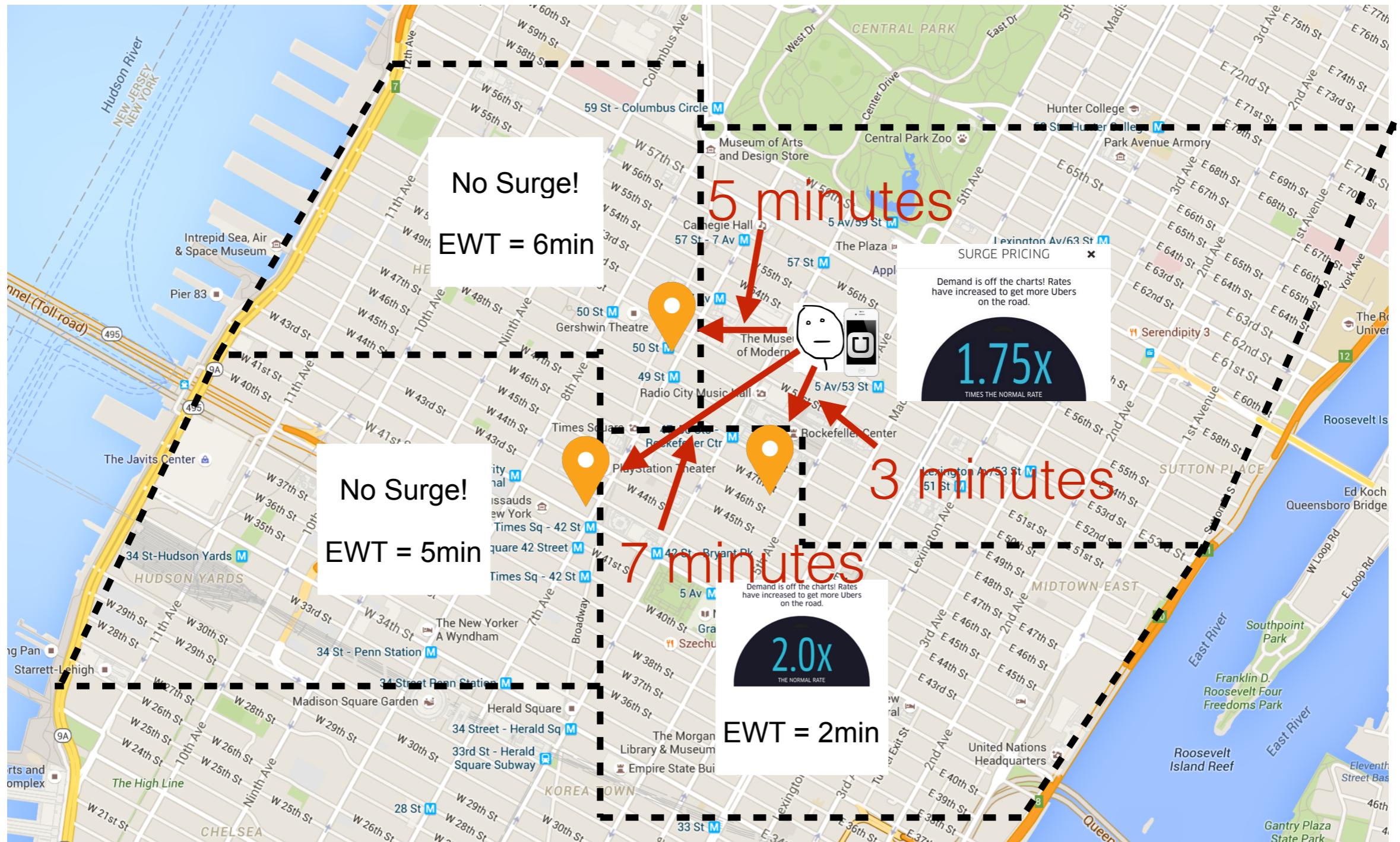
Can we avoid surge pricing?



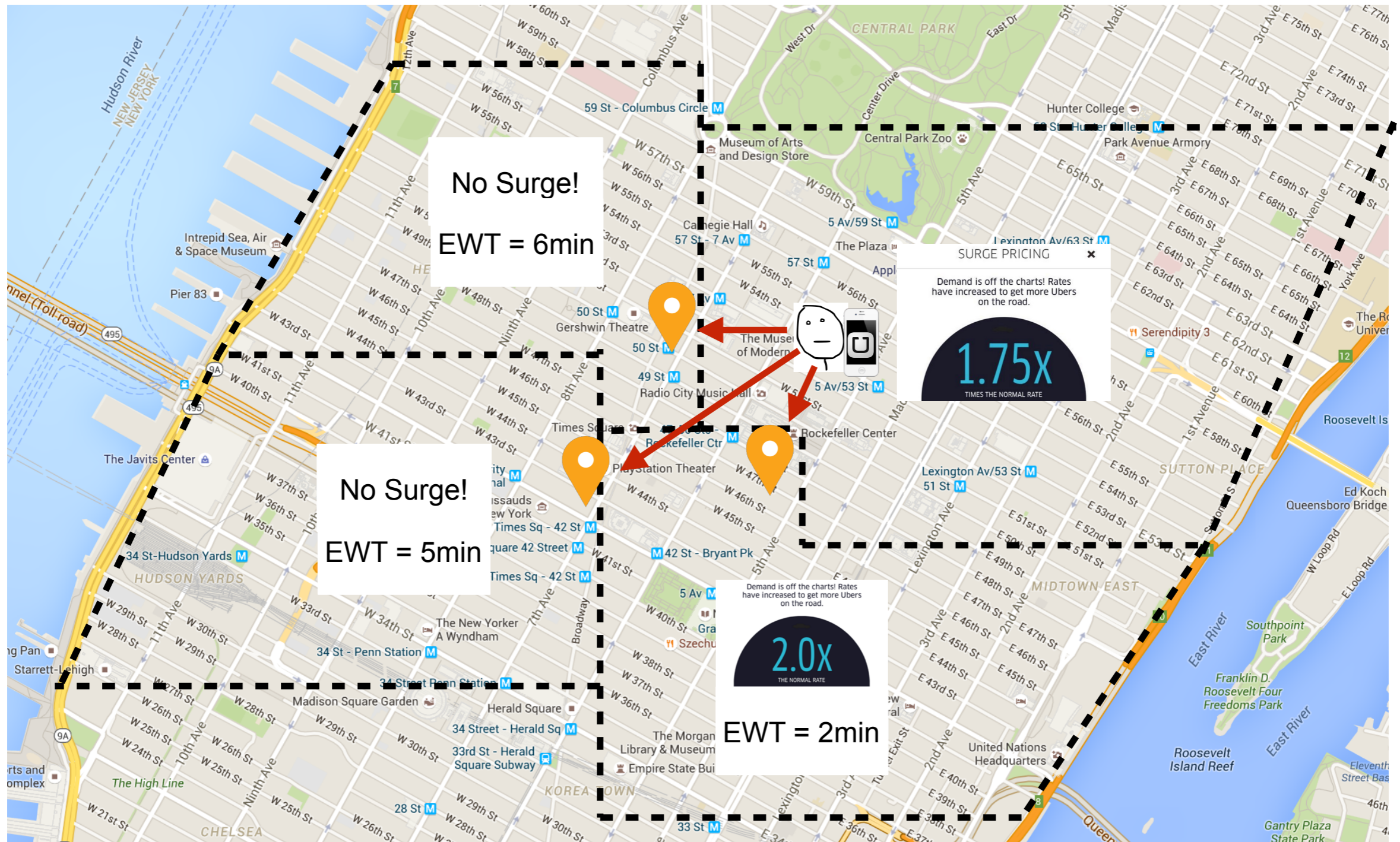
Can we avoid surge pricing?



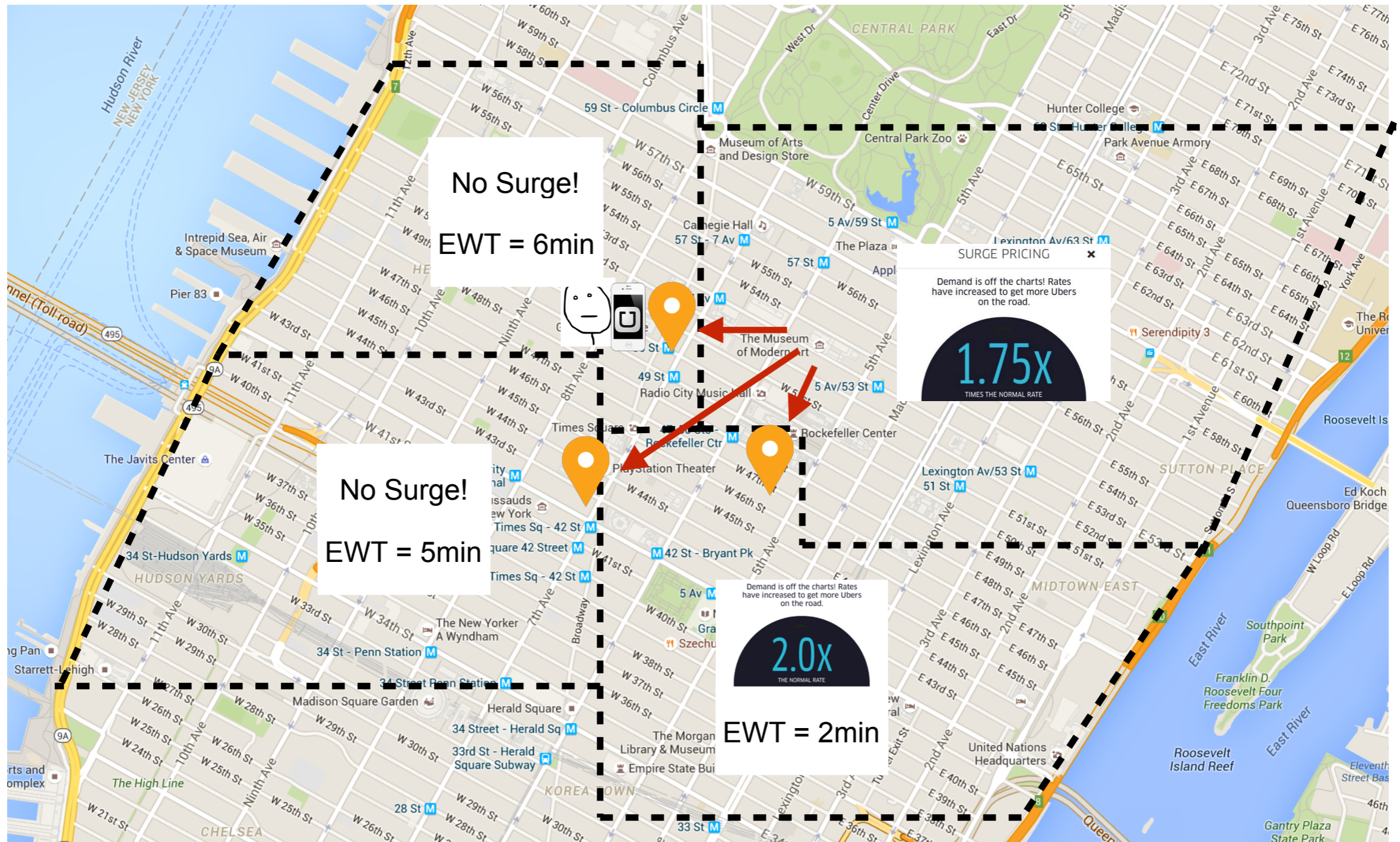
Can we avoid surge pricing?



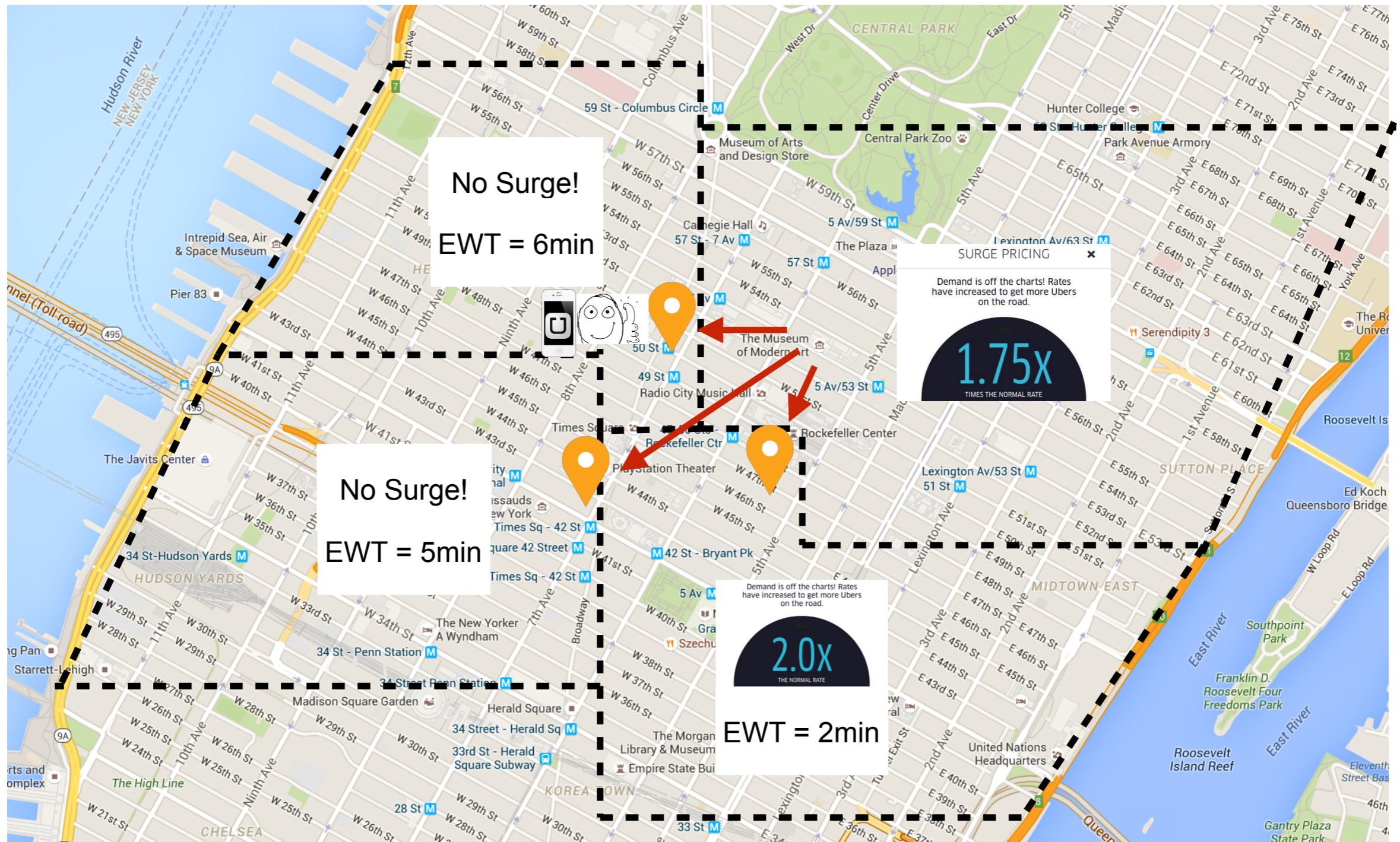
Can we avoid surge pricing?



Can we avoid surge pricing?

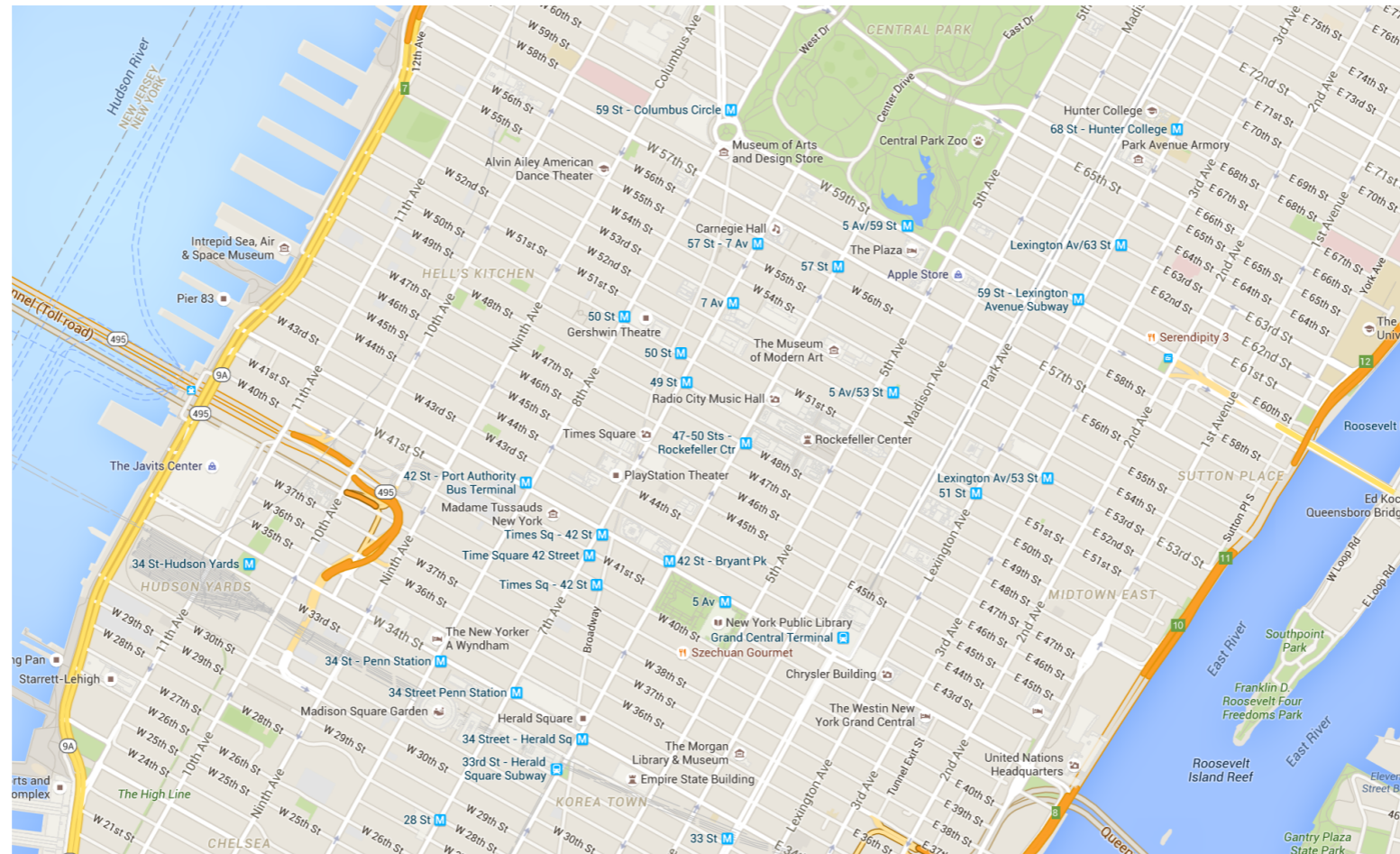


Can we avoid surge pricing?

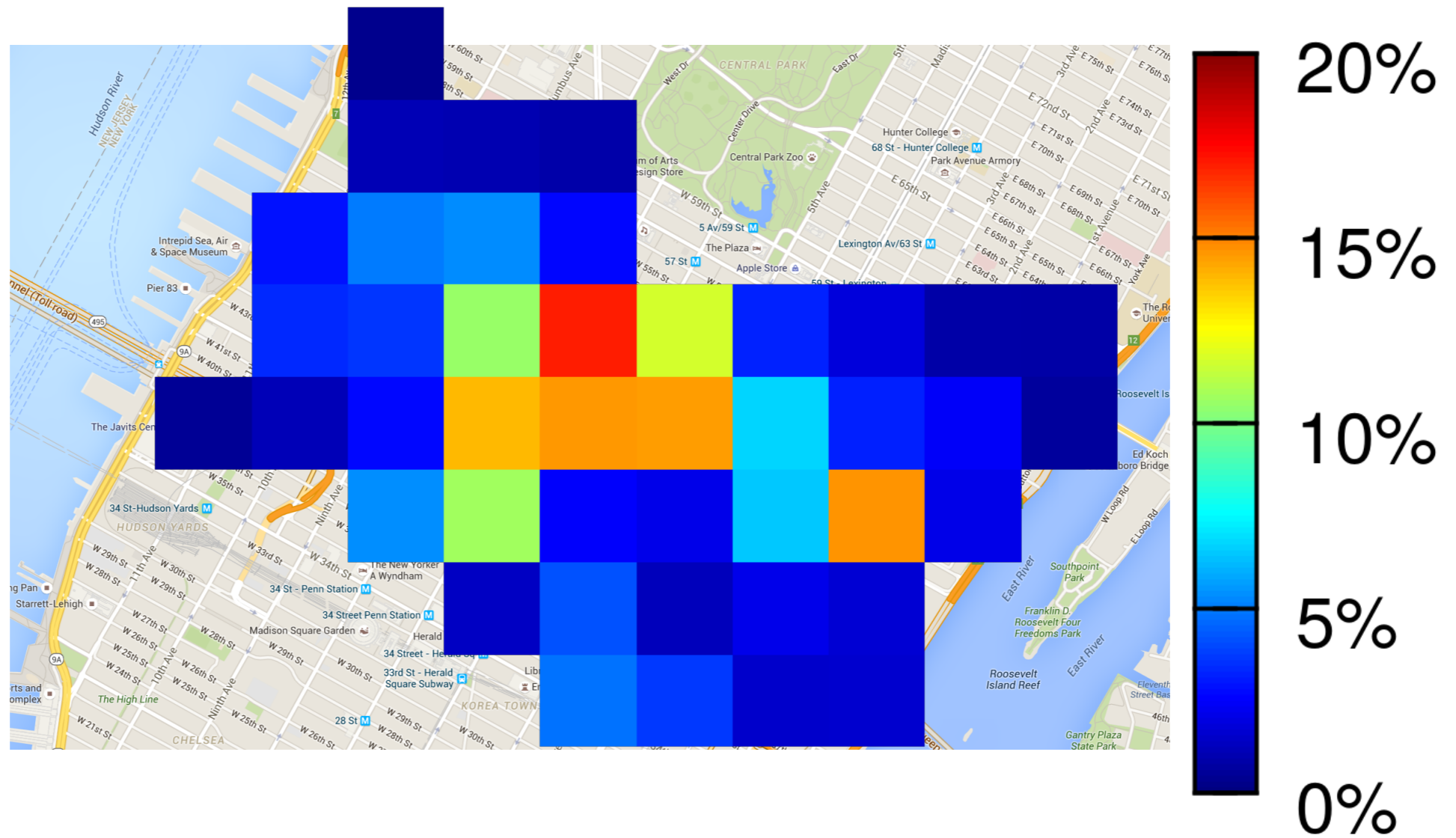


Avoiding surge

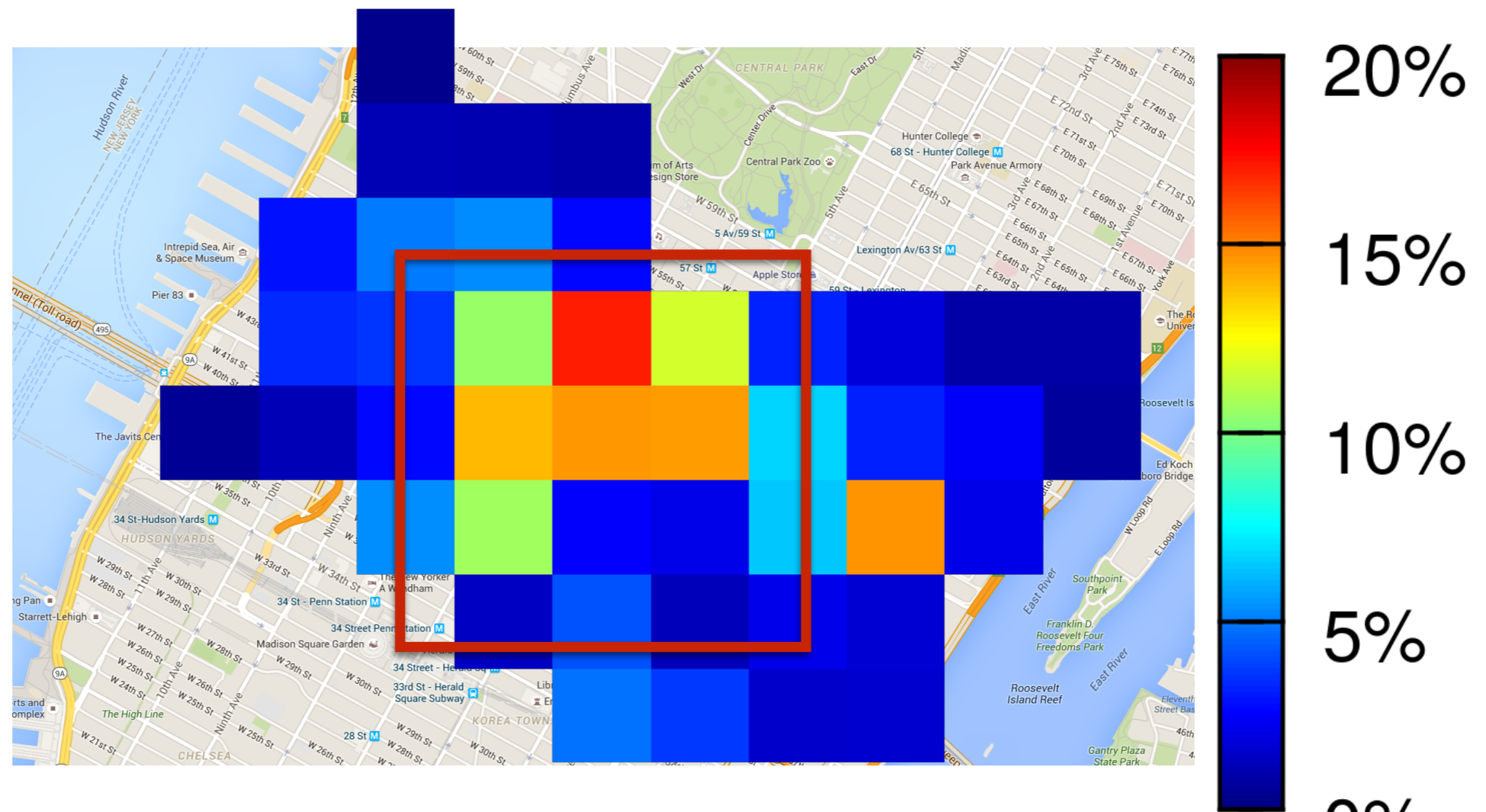
Avoiding surge



Avoiding surge

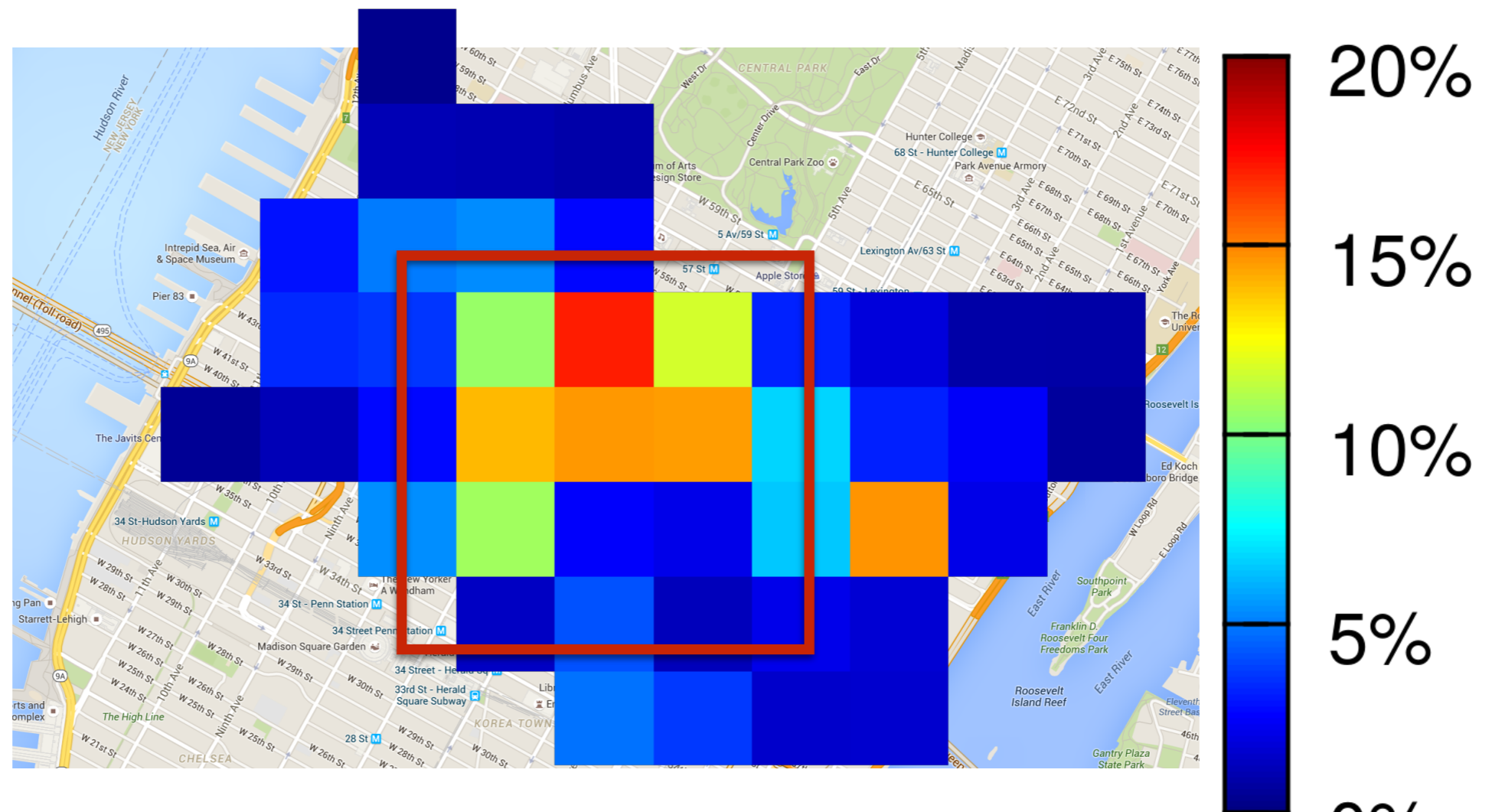


Avoiding surge



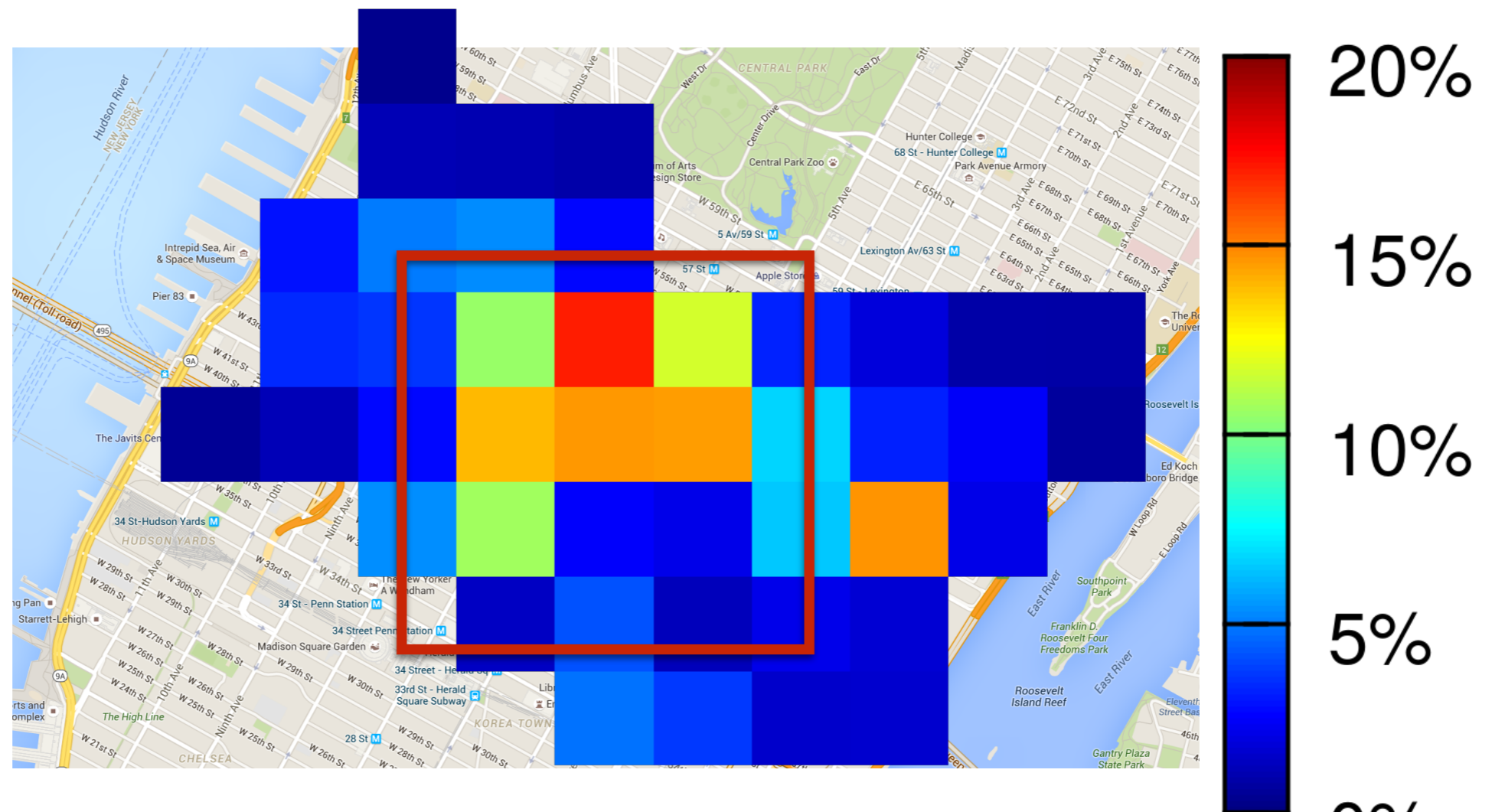
- 10% — 15% of chances saving money

Avoiding surge



- 10% — 15% of chances saving money
- On average, we can save 50% of prices

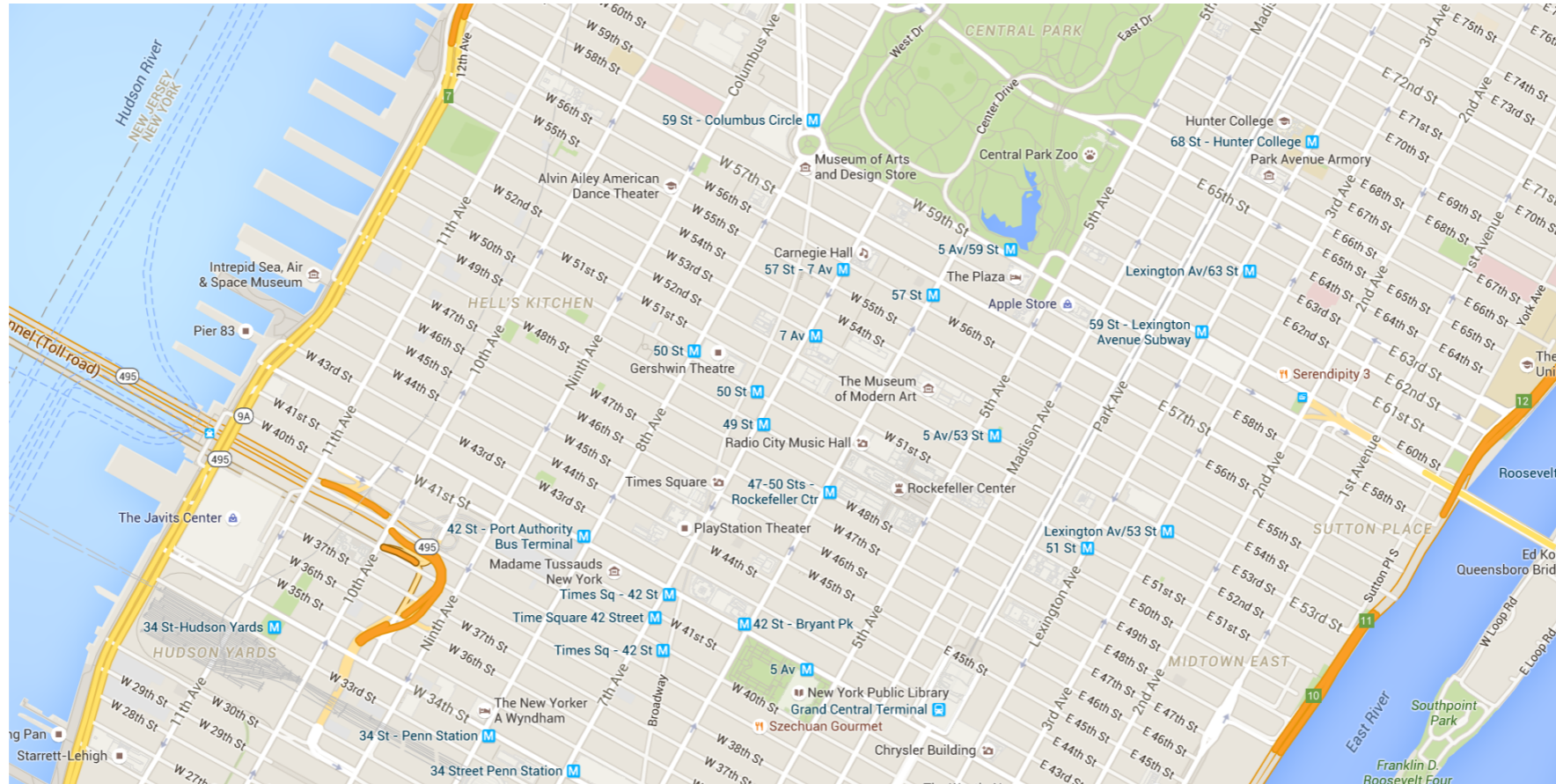
Avoiding surge



- 10% — 15% of chances saving money
- On average, we can save 50% of prices
- We built an app, but cannot release it

What is the impact of surge pricing?
Does it conform with Uber's claim?

What is the impact of surge pricing? Does it conform with Uber's claim?



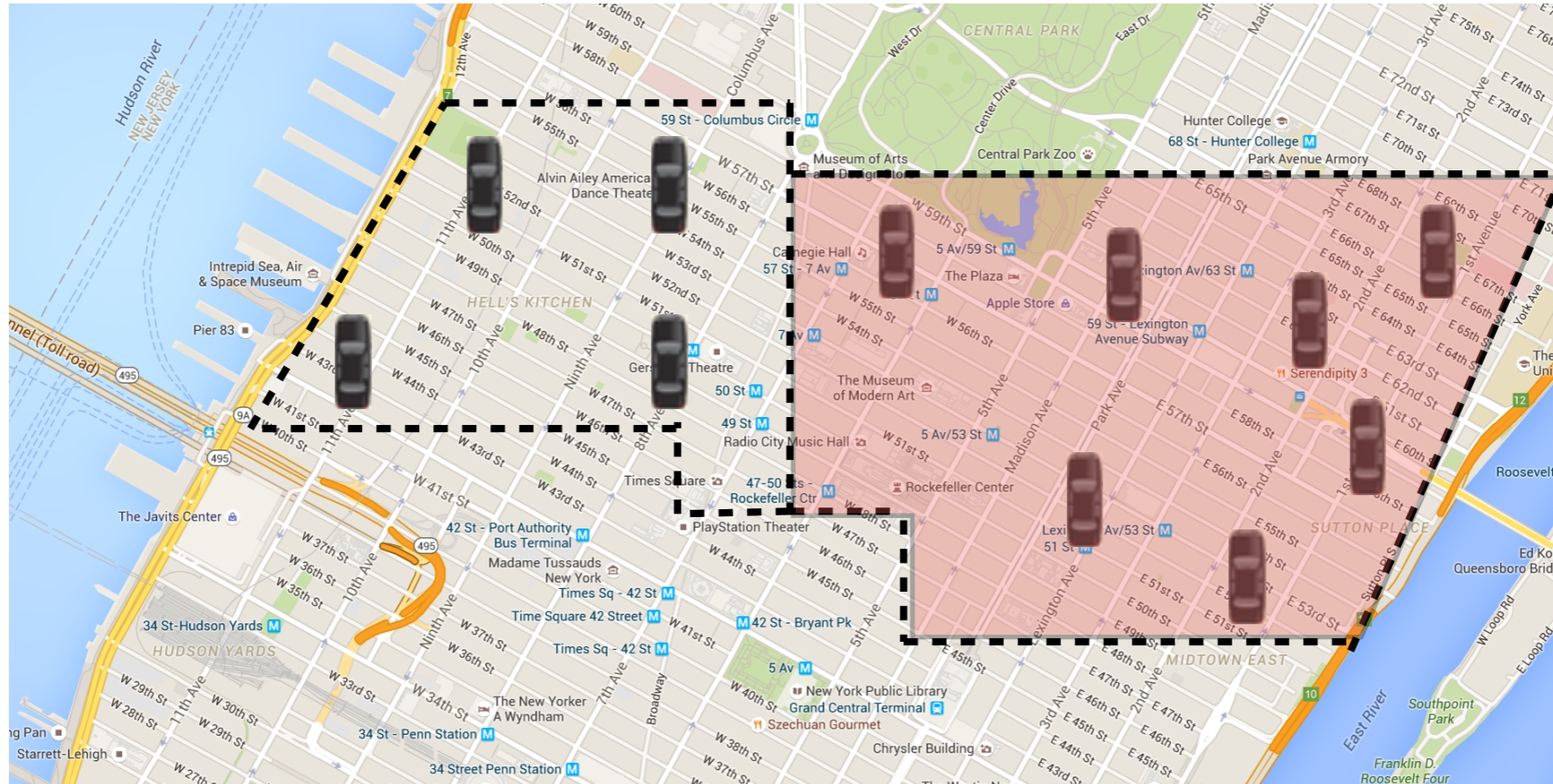
What is the impact of surge pricing? Does it conform with Uber's claim?



What is the impact of surge pricing? Does it conform with Uber's claim?

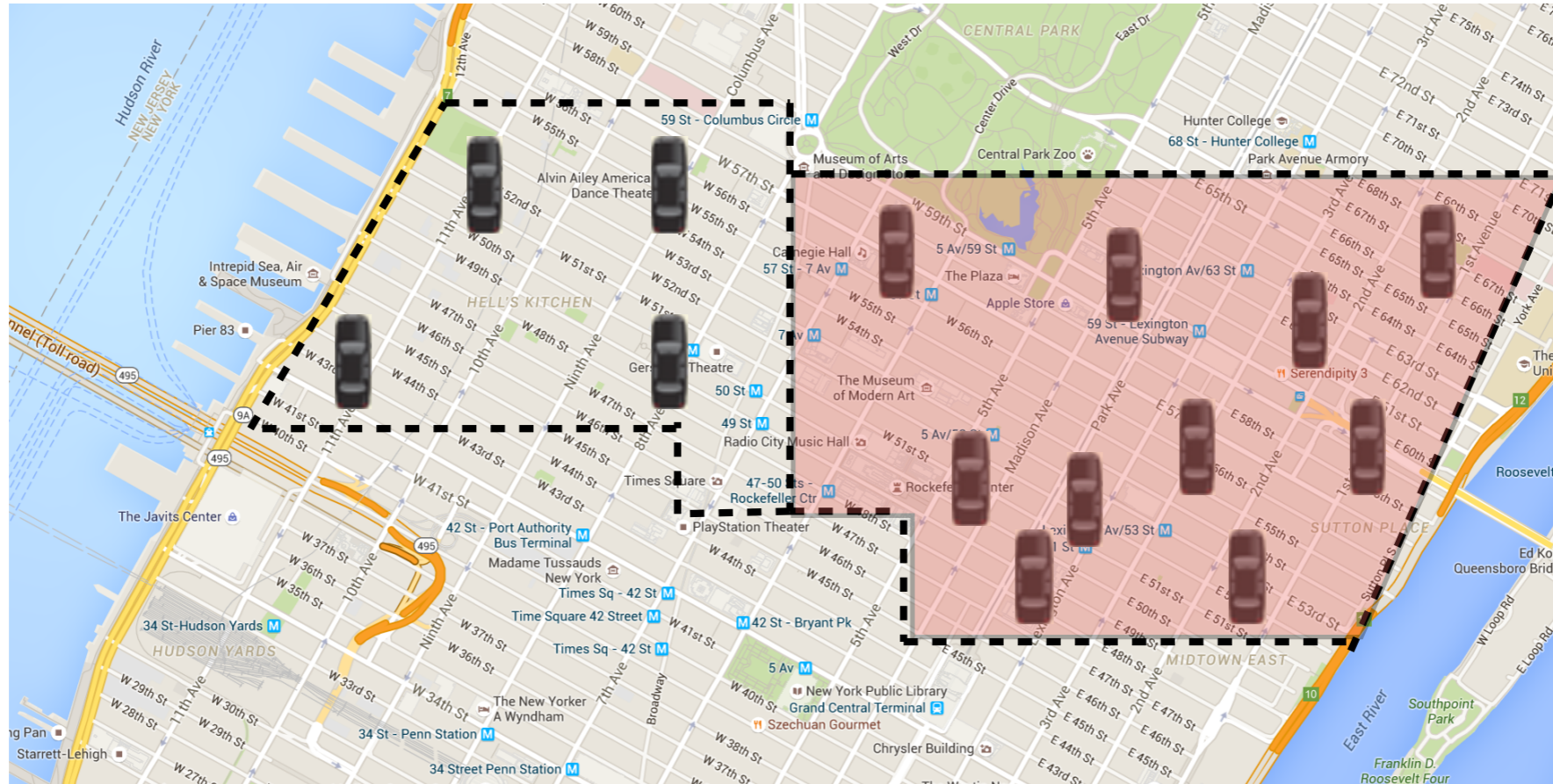


What is the impact of surge pricing? Does it conform with Uber's claim?



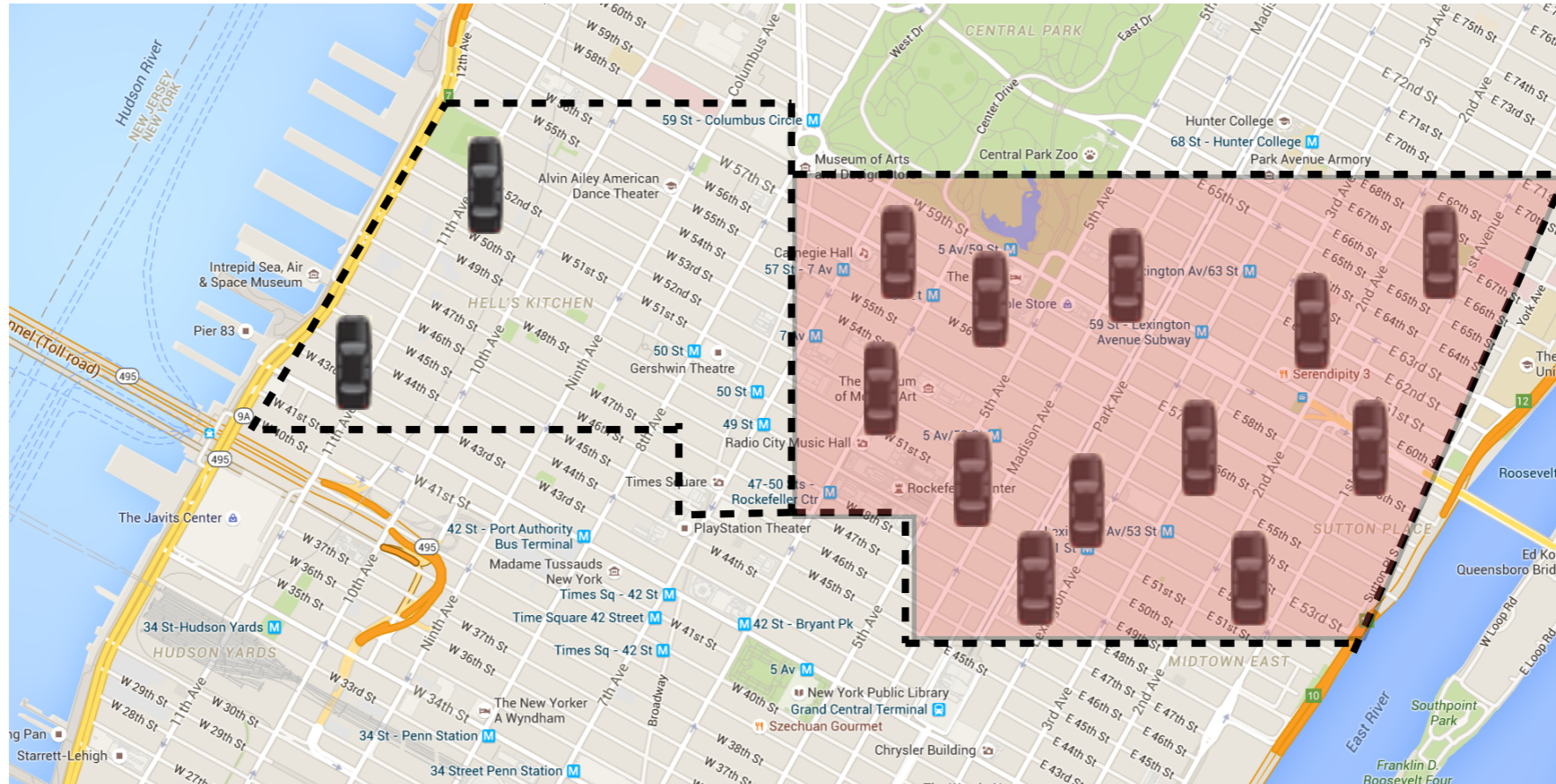
- Supply

What is the impact of surge pricing? Does it conform with Uber's claim?



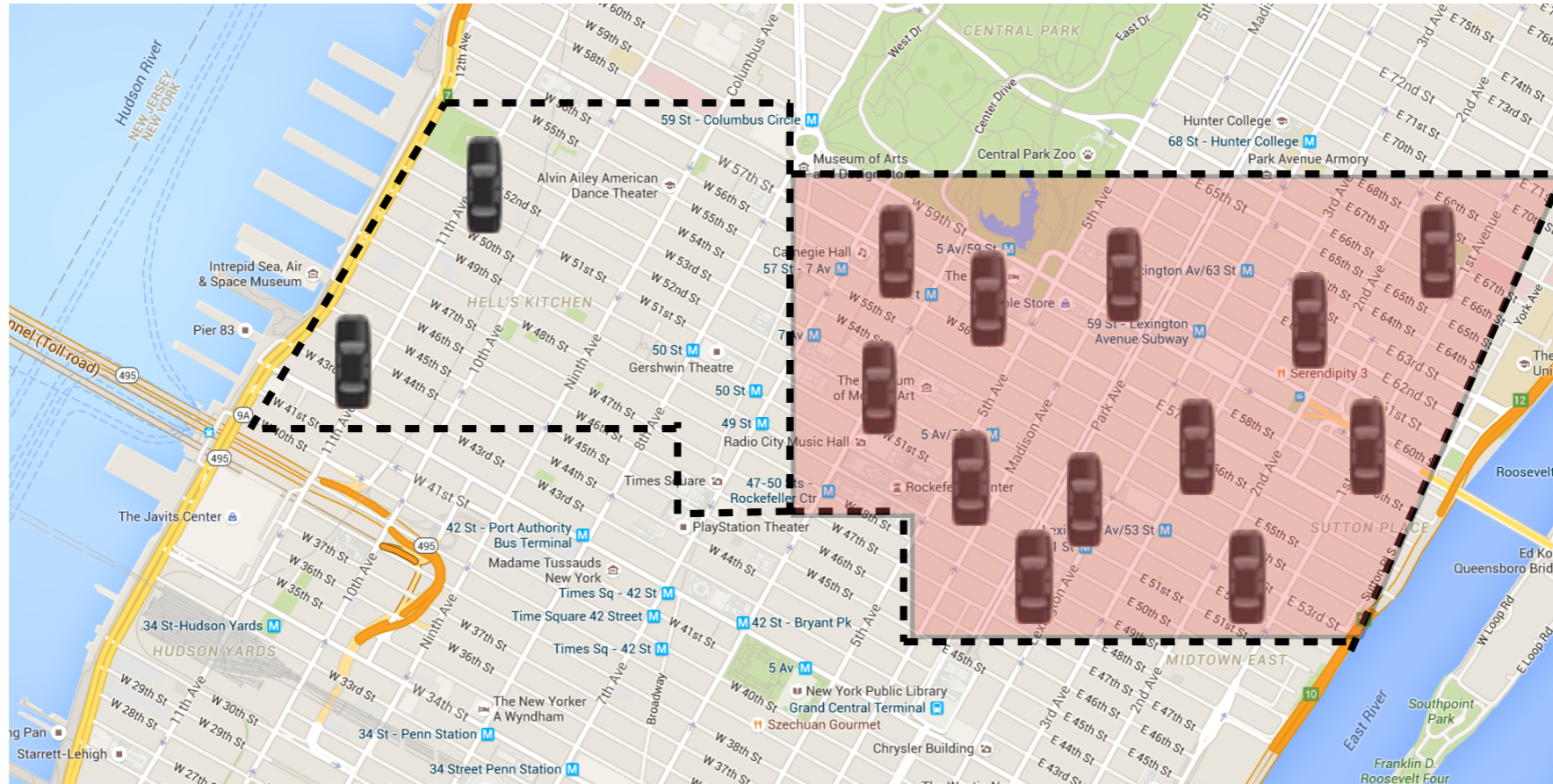
- Supply
 - New: more new log-in cars in surging area

What is the impact of surge pricing? Does it conform with Uber's claim?



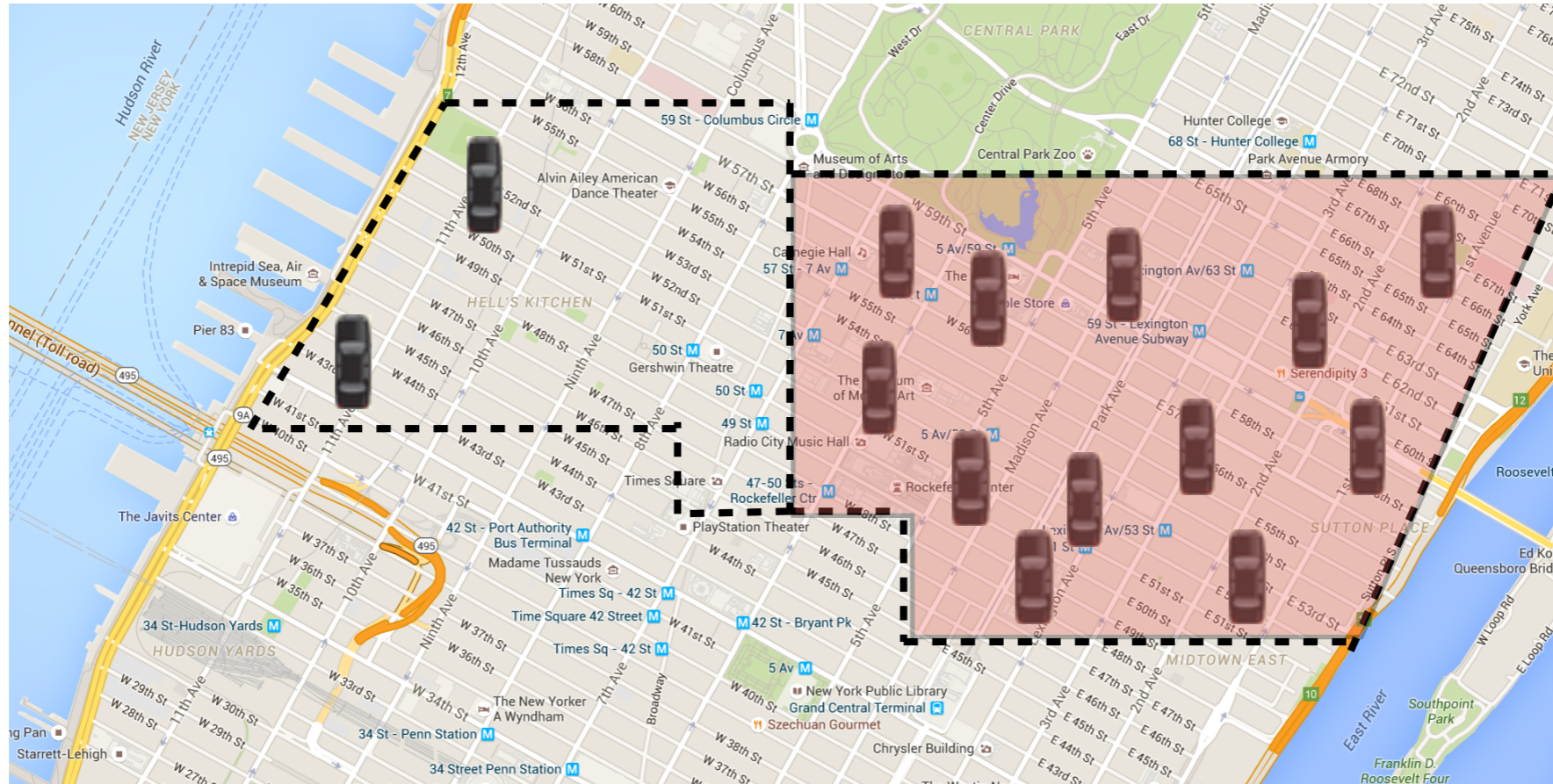
- Supply
 - New: more new log-in cars in surging area
 - Move-in: more move-in cars from nearby areas

What is the impact of surge pricing? Does it conform with Uber's claim?



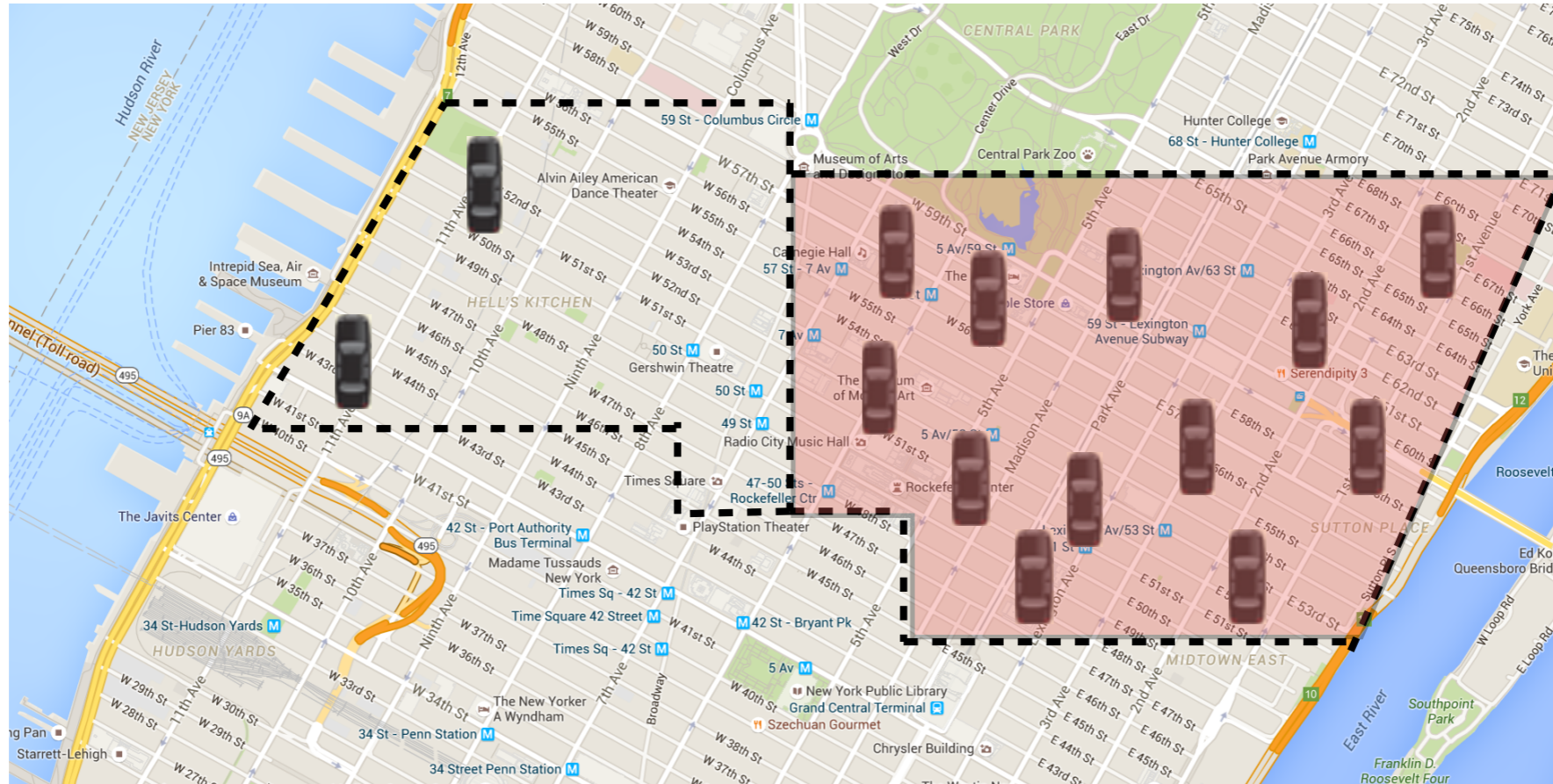
- Supply
 - New: more new log-in cars in surging area
 - Move-in: more move-in cars from nearby areas
 - Move-out: less cars leaving the surging area

What is the impact of surge pricing? Does it conform with Uber's claim?



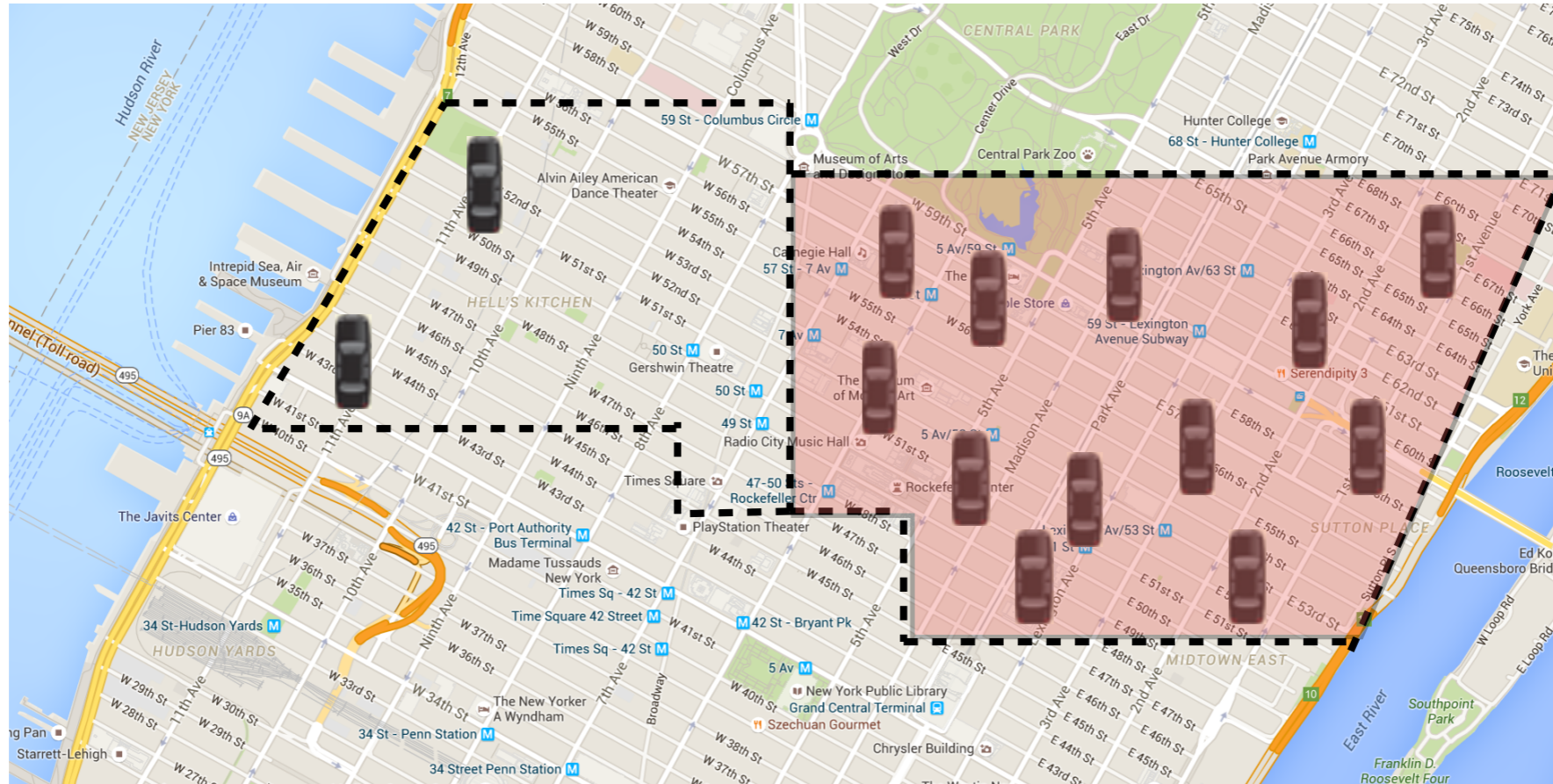
- Supply
 - New: more new log-in cars in surging area
 - Move-in: more move-in cars from nearby areas
 - Move-out: less cars leaving the surging area
- Demand

What is the impact of surge pricing? Does it conform with Uber's claim?



- Supply
 - New: more new log-in cars in surging area
 - Move-in: more move-in cars from nearby areas
 - Move-out: less cars leaving the surging area
- Demand
 - Booked: less cars getting booked in the surging area

What is the impact of surge pricing? Does it conform with Uber's claim?



- Supply
 - New: more new log-in cars in surging area
 - Move-in: more move-in cars from nearby areas
 - Move-out: less cars leaving the surging area
- Demand
 - Booked: less cars getting booked in the surging area
 - Old: more old cars staying the surging area

Impact on supply and demand, in reality

State

**During
surge**

Expected?

Impact on supply and demand, in reality

State	During surge	Expected?
New	+2%	Yes

Impact on supply and demand, in reality

State	During surge	Expected?
New	+2%	Yes
Booked	-7%	Yes

Impact on supply and demand, in reality

State	During surge	Expected?
New	+2%	Yes
Booked	-7%	Yes
Old	+14%	Yes

Impact on supply and demand, in reality

State	During surge	Expected?
New	+2%	Yes
Booked	-7%	Yes
Old	+14%	Yes
Move-in	-13%	No

Impact on supply and demand, in reality

State	During surge	Expected?
New	+2%	Yes
Booked	-7%	Yes
Old	+14%	Yes
Move-in	-13%	No
Move-out	+14%	No

Summary of Uber study

Summary of Uber study

- Presented the first in-depth analysis of Uber
 - Covered dynamics of supply, demand, surge, etc.
 - Detailed implementation of surge pricing algorithm

Summary of Uber study

- Presented the first in-depth analysis of Uber
 - Covered dynamics of supply, demand, surge, etc.
 - Detailed implementation of surge pricing algorithm
- Showed impact of surge on supply and demand

Thanks!

- Questions?