Taxi Trips in DC

To tip or not to tip, that is the question

Emefa Agodo & Keita Miyaki

Questions

- What determines the taxi fare?
  - Mileage would explain most of meterfare, but anything else?
- What determines the tips?
  - Who are stingy and who are generous?
  - Is tipping time-varying?

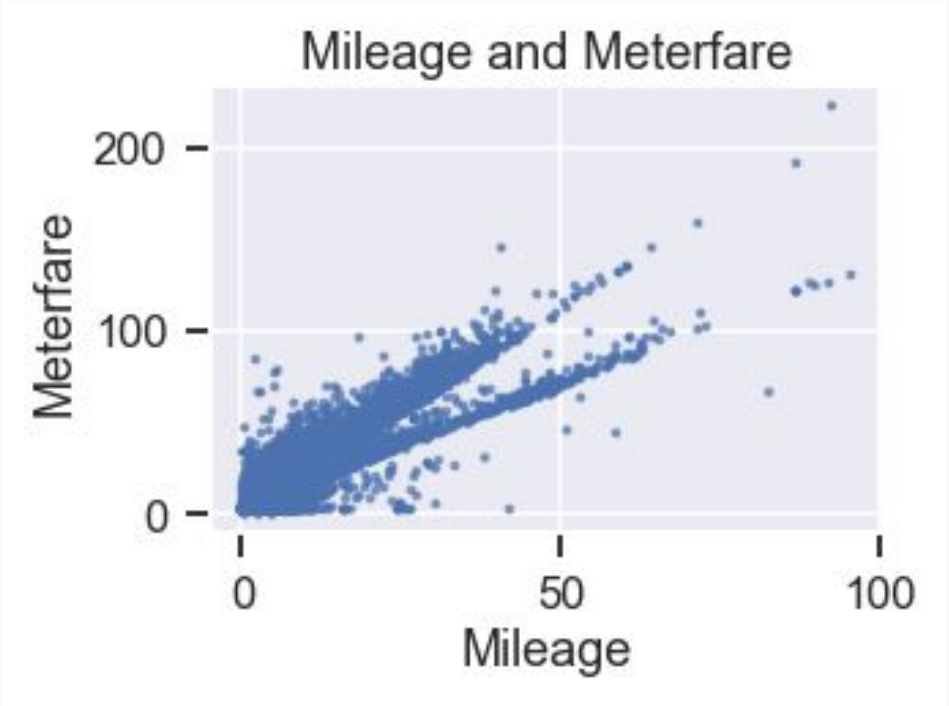
Dataset

- Taxicab Trips Sampling in July 2017
  - Provided by the DC Government
  - 228,611 observations over July 23 to July 30, 2017.
  - 196,493 observations after cleaning
- Data cleaning
  - Removed negative meterfare, total, mileage, and trip duration
  - Removed total over $1,000, mileage over 100 miles, trip duration over 2 hours, and average mileage/minute over 2 (=120 mph)

What determines the taxi fare?

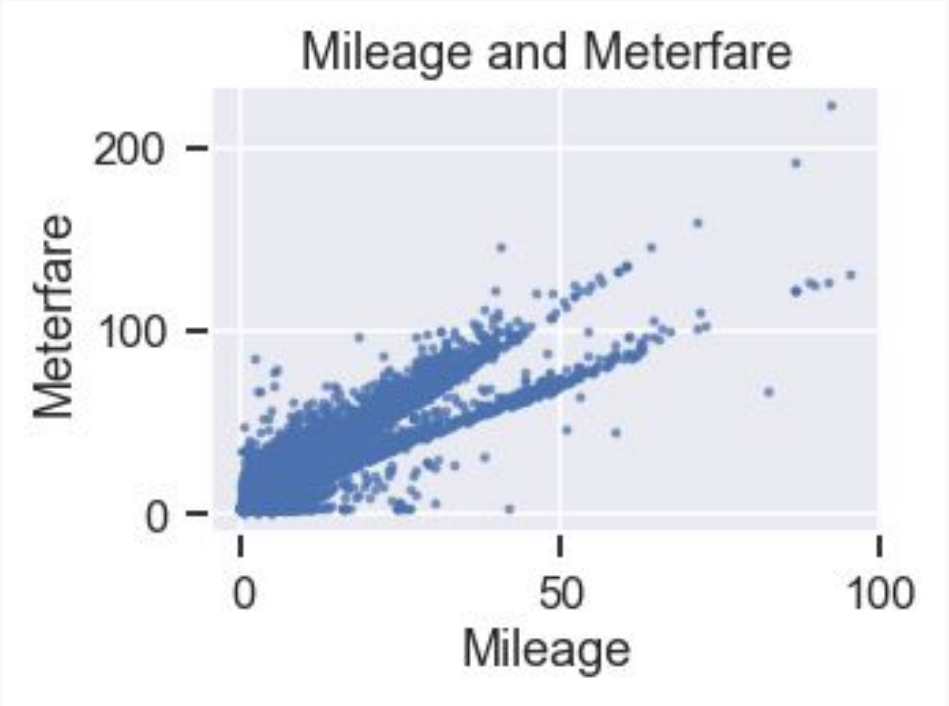
- Meterfare is largely determined by mileage

Meterfare =
$5.90 + $1.98 * mile + u



What determines the taxi fare?

- But there looks to be two different trend lines?

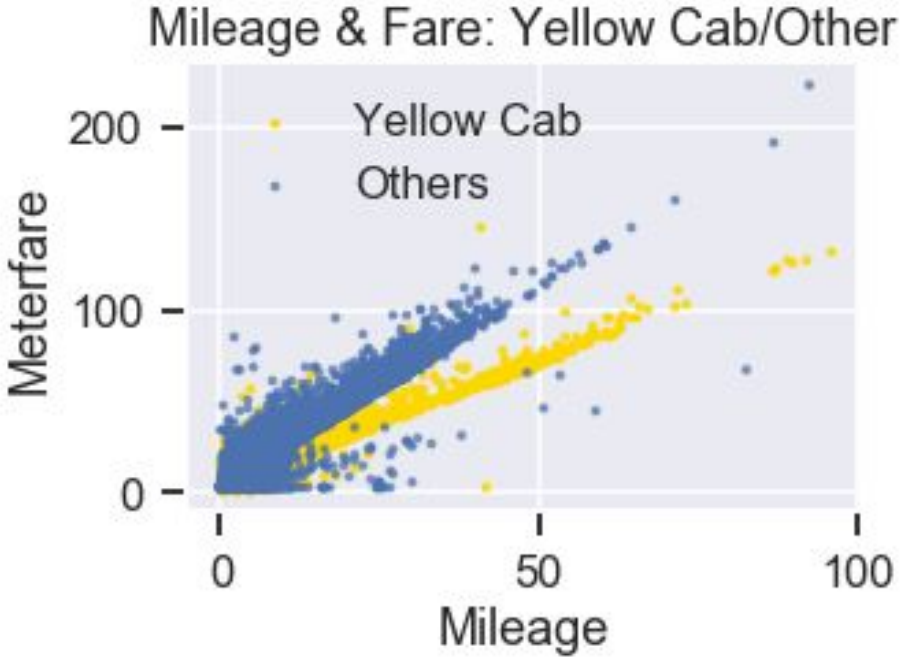


What determines the taxi fare?

- Yellow cabs have different fare system

Meterfare =
$5.97 + $1.41 * mile + u

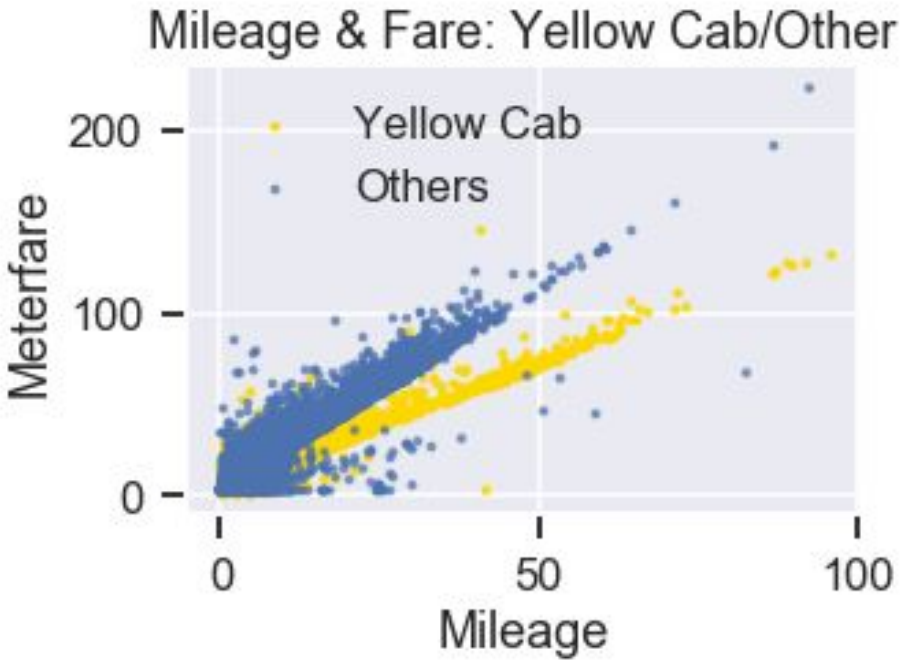
Meterfare=
$5.42 + $2.25 * mile + u



What determines the taxi fare?

Dependent Variable: Meterfare

Variables	OLS 1	OLS 2
Mileage	1.98***	2.25***
	(0.009)	(0.006)
Yellow Cab (dummy)		0.55***
		(0.04)
Mileage * Yellow Cab (dummy)		-0.83***
		(0.009)
Constant	5.90**	5.42***
	(0.03)	(0.02)
N	196,493	196,493
Adj. R Squared	0.868	0.909



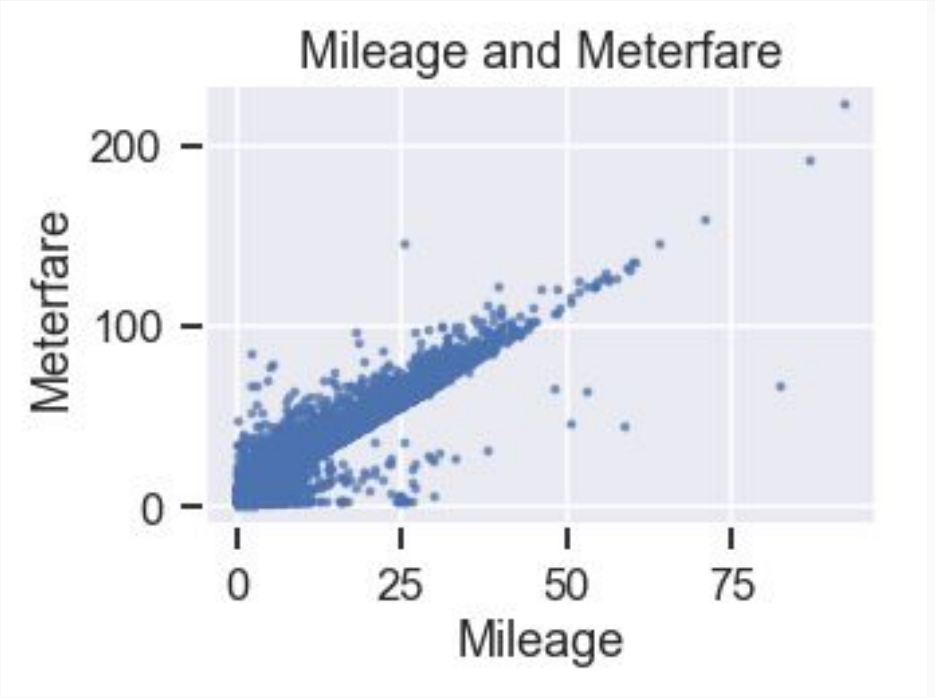
Wait a minute!

2.25 / 1.41 = **1.60**

What is this number?

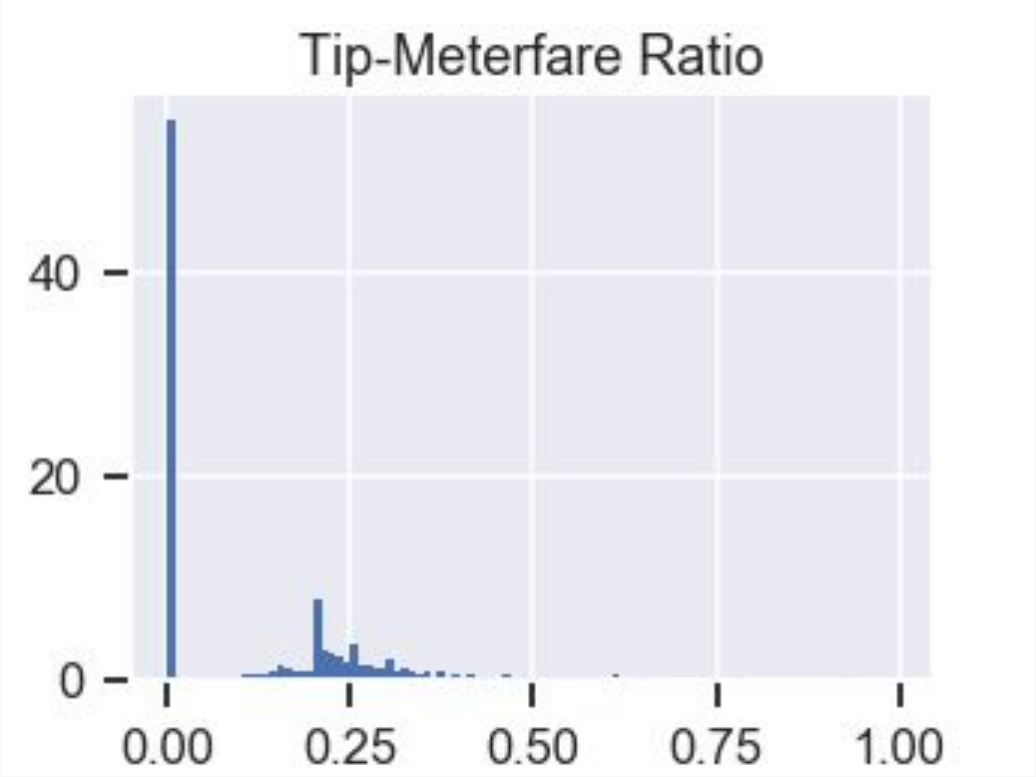
Yellow Cab Reported Trip Mileage in Kilometers, while Others Reported in Miles

- It was a clear measurement error
- After correcting the data, the scatter plot looks much cleaner



Do you tip?

- 54% of our samples indicated no tip
- Median tip amount is around 23% of meterfare for those who tip



Who tips?

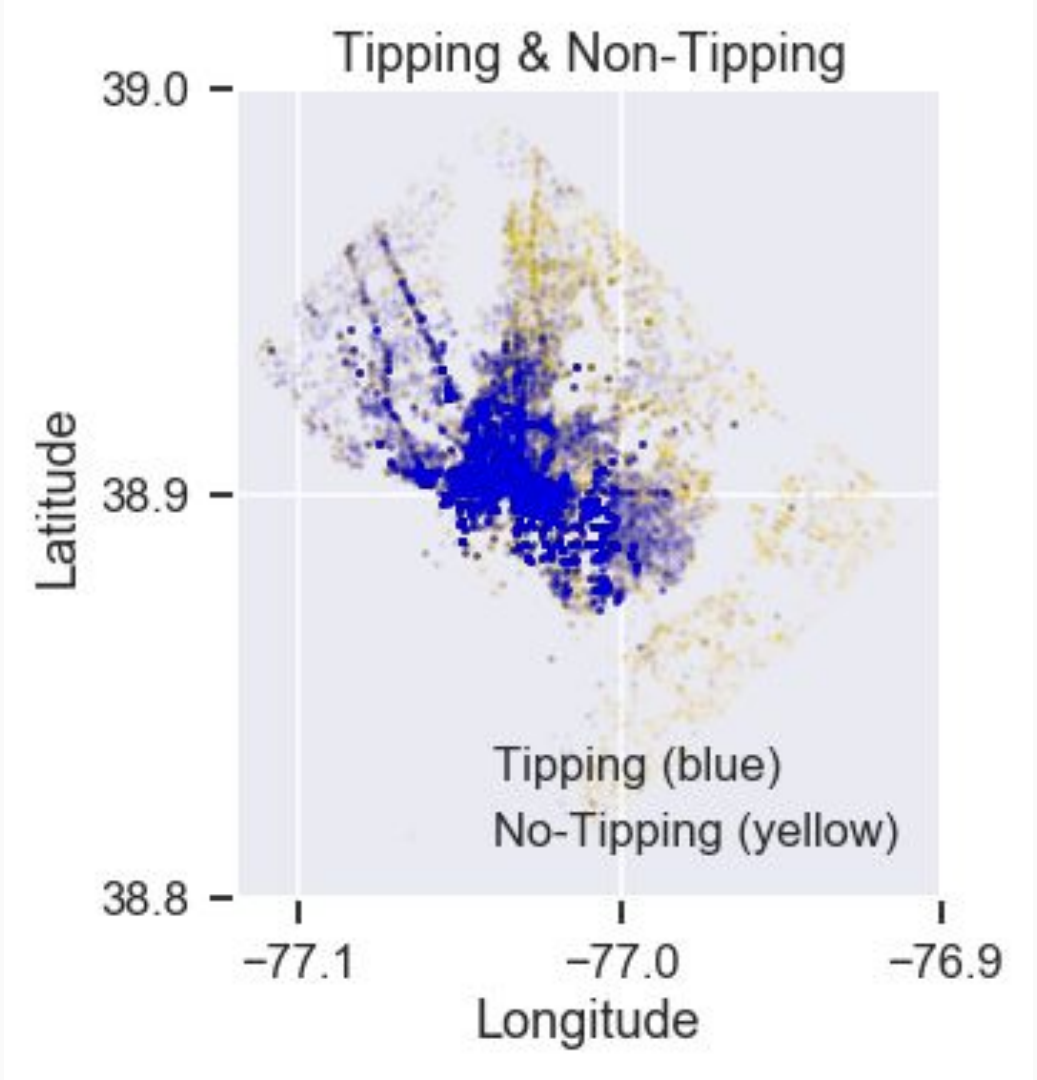
- Let's run Logistic regressions
- Dependent var:
  - 1 for positive tip
  - 0 for zero tip

	Logit 1	Logit 2
Airport	0.5006***	0.4415***
	(0.0315)	(0.0511)
Credit Card	7.0683***	7.2235***
	(0.0492)	(0.0574)
Voucher	6.3998***	6.5849***
	(0.2452)	(0.2800)
Trip Mileage	-0.0208***	-0.0653***
	(0.0021)	(0.0060)
Drop-off Quadrant NW		0.2220***
		(0.0322)
Drop-off Quadrant SE		0.1422***
		(0.0499)
Drop-off Quadrant SW]		0.1126**
		(0.0480)

	Logit 1	Logit 2
Control Date/Hour	Y	Y
Control Provider/Type	Y	Y
Pseudo R-squared	0.6761	0.6895
No. observations	196493	155050

Downtown vs Suburb

- It seems drop-off location matters
- In NE and SE, no-tipping prevails in suburbs?



Who tips?

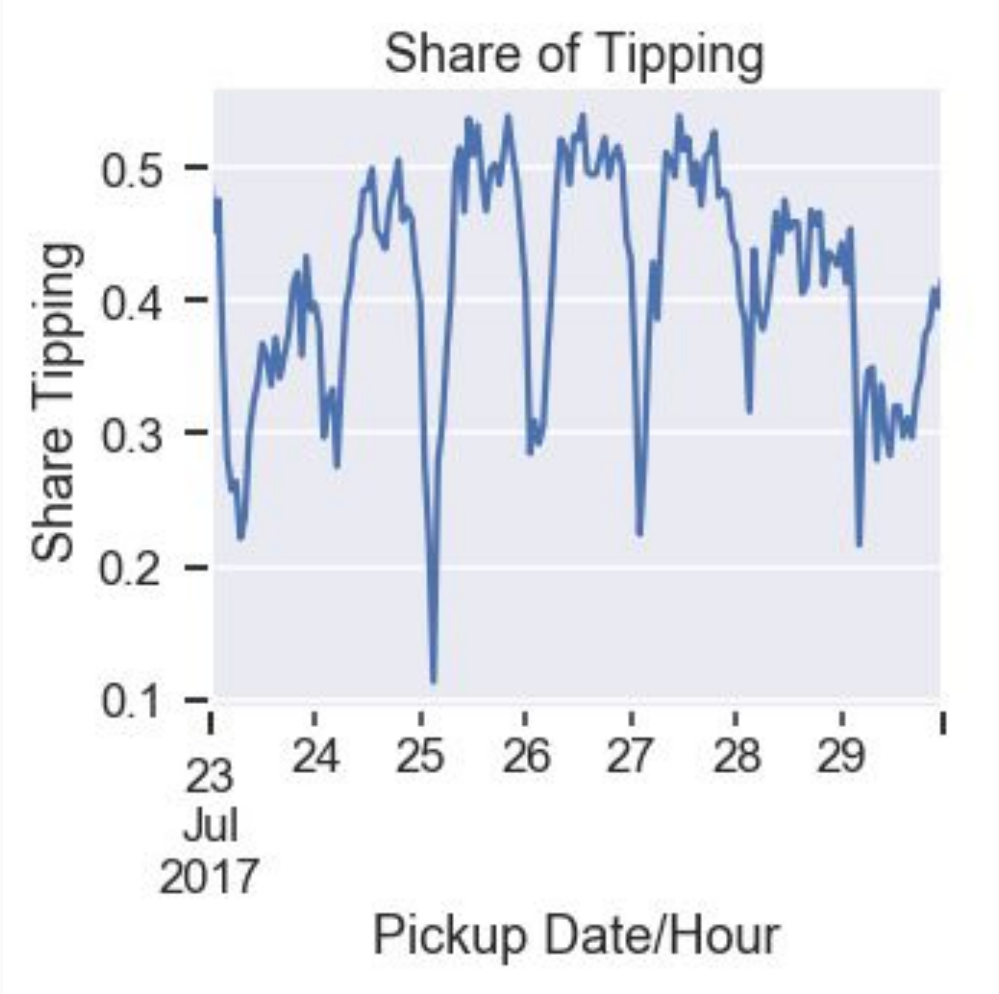
- Let's run Logistic regressions
- Dependent var:
  - 1 for positive tip
  - 0 for zero tip

	Logit 3	Logit 4
Airport	0.3620***	0.3471***
	(0.0506)	(0.0505)
Credit Card	7.2229***	7.2182***
	(0.0575)	(0.0575)
Voucher	6.5895***	6.6009***
	(0.2789)	(0.2786)
Trip Mileage	-0.0381***	-0.0333***
	(0.0058)	(0.0057)

	Logit 3	Logit 4
Distance from Capitol	-8.9893***	-29.9535***
	(0.5791)	(1.8156)
Drop-off Quadrant NW		25.0571***
* Distance from Capitol		(1.8940)
Drop-off Quadrant SE		-8.2814**
* Distance from Capitol		(3.4600)
Drop-off Quadrant SW		12.3487***
* Distance from Capitol		(3.8098)
Control Quadrant	Y	Y
Control Date/Hour	Y	Y
Control Provider/Type	Y	Y
Pseudo R-squared	0.6906	0.6921
No. observations	155050	155050

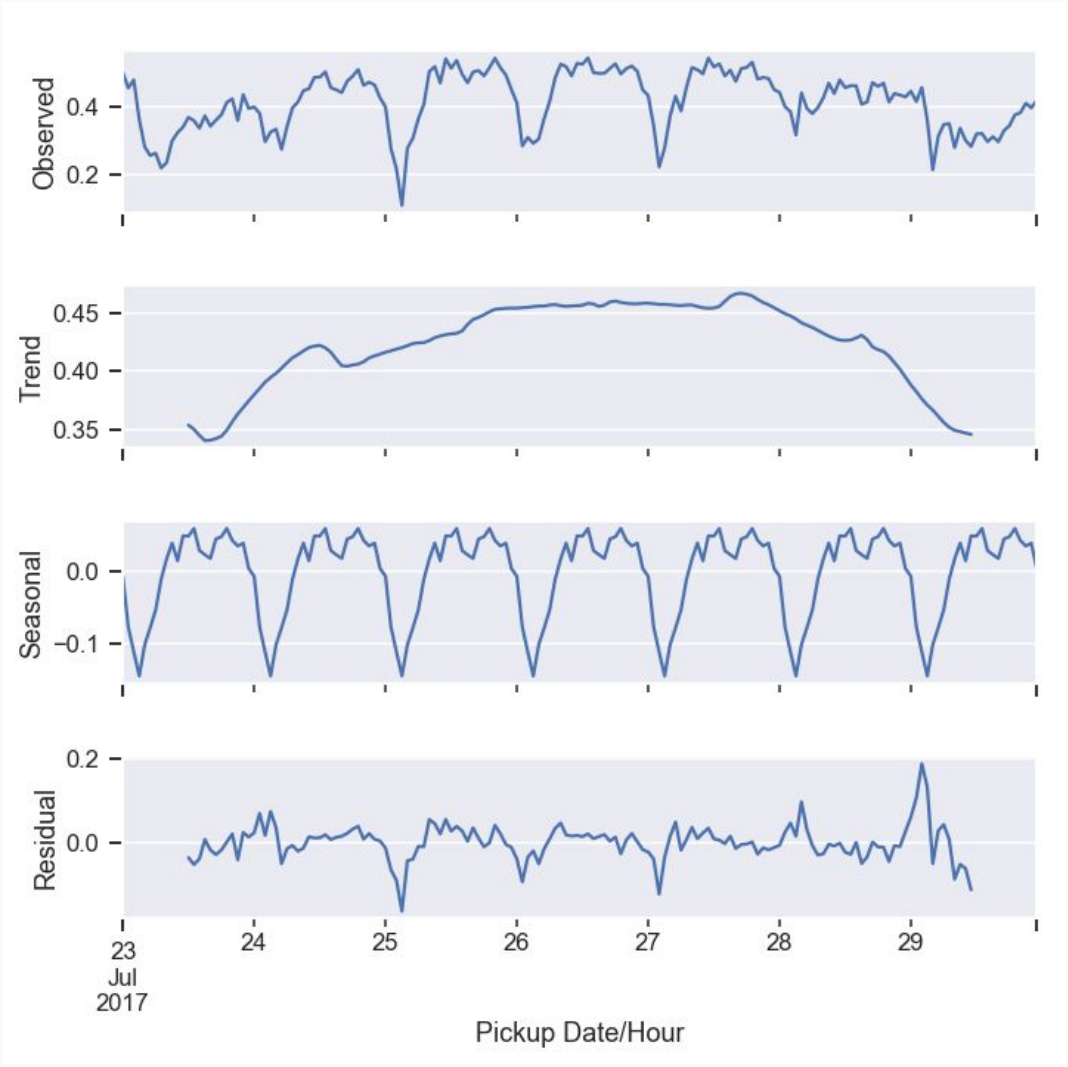
What is the time-varying trend in tipping?

- Share of tipping declines in early mornings



What is the time-varying trend in tipping?

- Share of tipping declines in early mornings



Source

DC Government

https://opendata.dc.gov/datasets/taxicab-trips-sampling-in-july-2017