

Counting things (02-04)

Matthew J. Salganik
Department of Sociology
Princeton University

Soc 596: Computational Social Science
Fall 2016



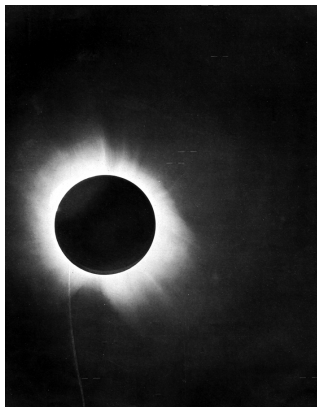
Beneath the jargon and technical details, lots of research is just counting things.

Beneath the jargon and technical details, lots of research is just counting things.

What to count?

- ▶ Interesting
- ▶ Important

- ▶ Theory-driven data analysis
- ▶ Data-driven theorizing



What makes this interesting and important is not in the data itself

https://commons.wikimedia.org/wiki/File:1919_eclipse_positive.jpg

WHY YOU CAN'T FIND A TAXI IN THE RAIN AND OTHER LABOR SUPPLY LESSONS FROM CAB DRIVERS*

HENRY S. FARBER

Farber 2015. "Why you can't find a taxi in the rain and other labor supply lessons from cab drivers." *Quarterly Journal of Economics*,
<http://dx.doi.org/10.1093/qje/qjv026>

- ▶ Strategic research site (Merton)

- ▶ Strategic research site (Merton)
- ▶ Tests competing theory

- ▶ Strategic research site (Merton)
- ▶ Tests competing theory
- ▶ Replication and extension

- ▶ Strategic research site (Merton)
- ▶ Tests competing theory
- ▶ Replication and extension
- ▶ Best-case for big data

Careful counting of hours worked and hourly wages
Better data leads to better estimate

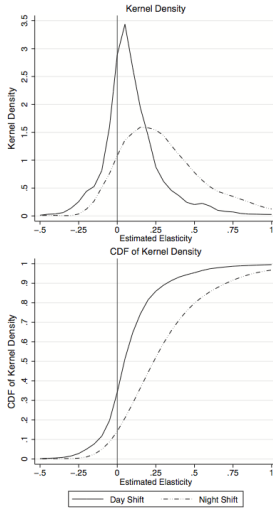
How does he use the size of his data?

How does he use the size of his data?

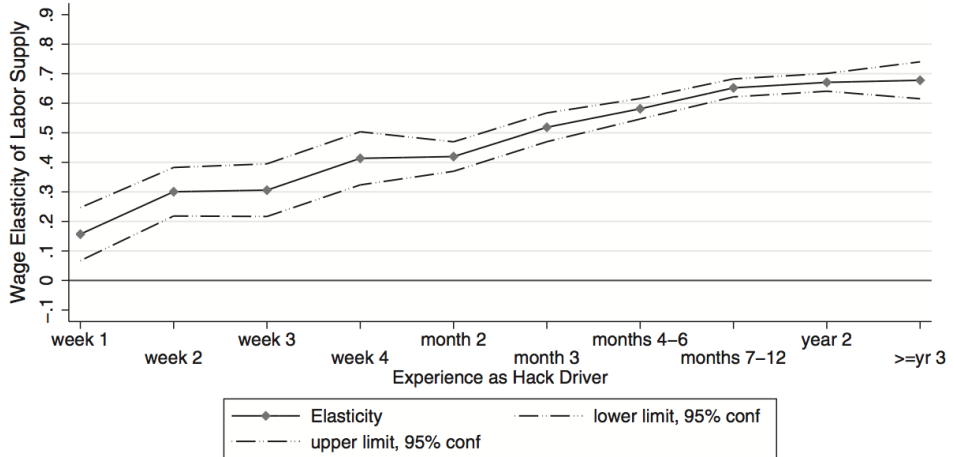
heterogeneity, mechanisms

Key ideas from social science:
heterogeneity, mechanisms

Kernel Density and CDF, IV Estimates of Individual Elasticities



New Drivers — All Shifts



Compare this to everything else people have done with NYC taxi data