

---

# **Covid-19: Testing Inequality in New York City**

S. Schmitt-Grohé

Ken Teoh

Martín Uribe

Columbia University

April 24, 2020

Our question: Did higher income give privileged access to tests for Covid-19?

Media seems to suggest as much: for example,

NYTimes headline on March 18, 2020:

Need a Coronavirus Test? Being Rich and Famous May Help.

Analysis focuses on Covid-19 testing inequality in NYC because:

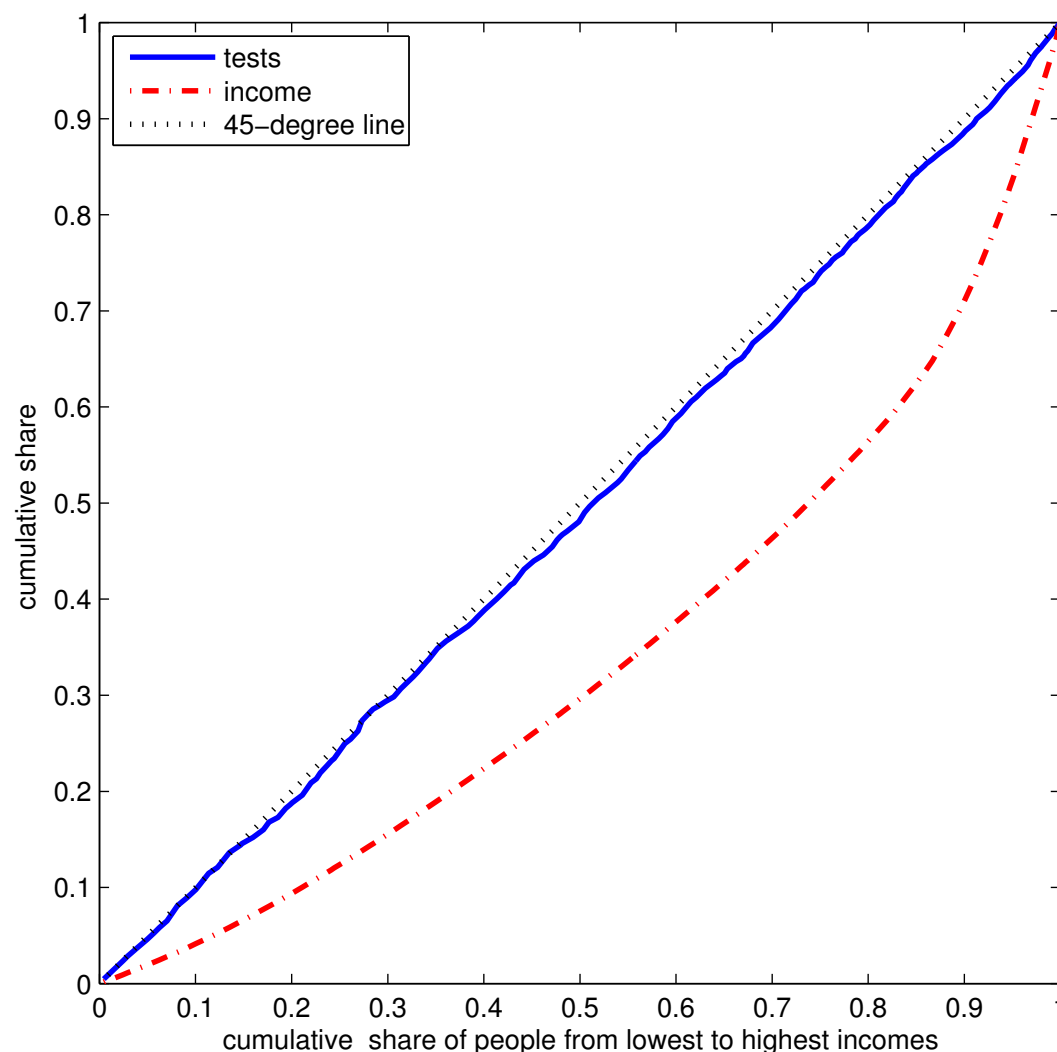
- NYC is the center of the coronavirus outbreak in the United States
- same health policies and regulations across its 177 zip codes
- sizeable variation in per capita income across zip codes
- same agencies produce data for all zip codes, making it comparable

## Data:

- cumulative number of New York City residents who were ever tested for Covid-19 and the number of residents who tested positive as of April 2, 2020 and April 13, 2020 for each of the city's 177 zip codes. Source: New York City Department of Health and Mental Hygiene (DOHMH).
- per capita income in dollars of 2018 past 12 months, population, and racial composition as of 2018 for each NYC zip code. Source: American Community Survey.
- ★ A positive test result means that the patient has the coronavirus and a **negative test** result means that the patient does **not** have the coronavirus. (ie., negative test means good news!)

Inequality in the number of coronavirus tests administered across income?

# Lorenz Curves of Covid-19 Tests and Mean Income Across New York City Zip Codes



	bottom 10%	top 10%
Income	0.04	0.29
Tests	0.10	0.11

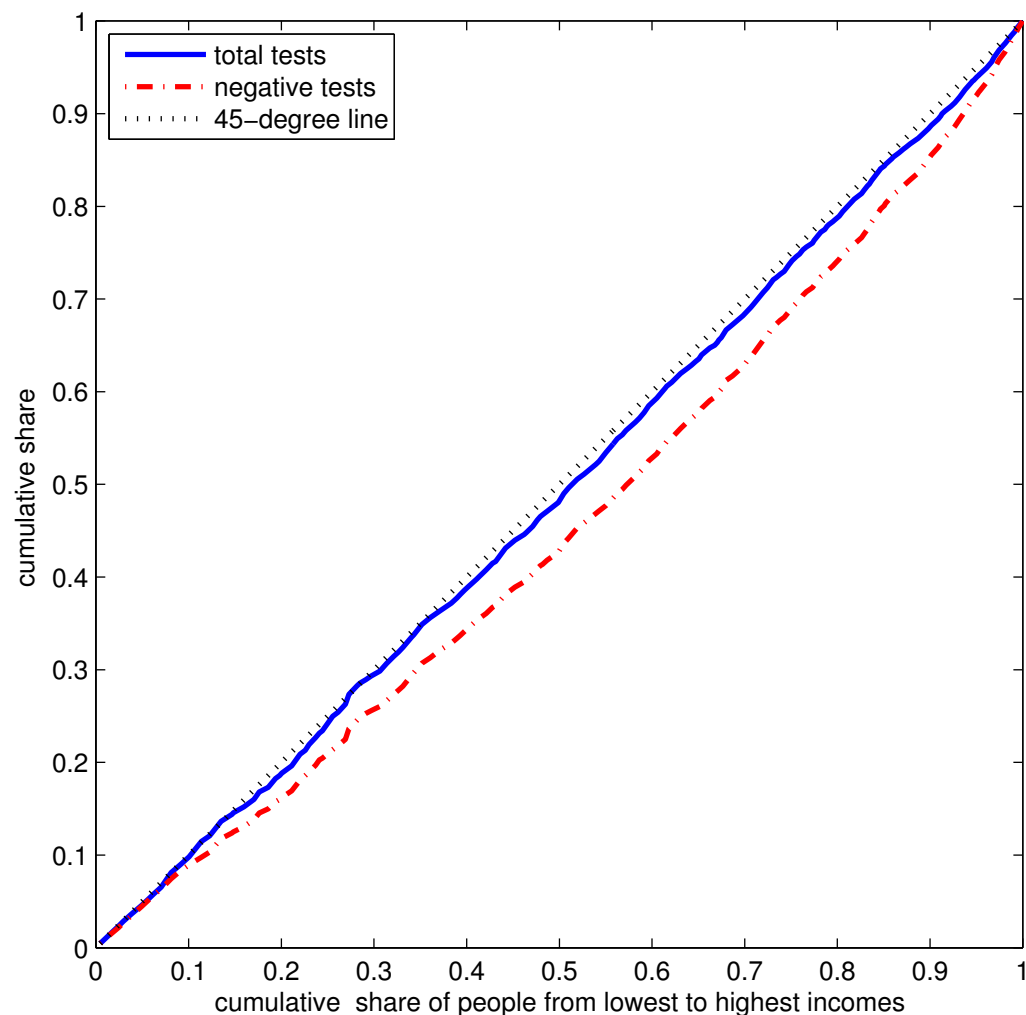
Gini income = 0.32

Gini tests = 0.02

Notes. Own calculations based on data from the New York City Department of Health and Mental Hygiene, as of April 2, 2020, and American Community Survey. Replication code `gini_testing.m` in `stu_covid19.zip`.

Inequality in test outcomes?

## Lorenz Curves of Negative Tests and Total Tests for Covid-19 Across New York City Zip Codes



	bottom 10%	top 10%
Tests	0.10	0.11
Negatives	0.09	0.15

Gini tests = 0.02  
Gini negatives = 0.09

Notes. Own calculations based on data from the New York City Department of Health and Mental Hygiene, as of April 2, 2020, and American Community Survey. Replication code `gini_negatives_testing.m` in `stu_covid19.zip`.



## Gini Coefficients

Income	0.32
Covid-19 Tests	0.02
Covid-19 Negative Test Results	0.09

Note. Own calculations based on data from the New York City Department of Health and Mental Hygiene as of April 2, 2020, and American Community Survey. Replication code `gini_testing.m` and `gini_negatives_testing.m` in `stu_covid19.zip`.

By Bayes Law:

$$P(neg|test)P(test) = P(test|neg)P(neg)$$

Bottom decile versus top decile

$$\begin{aligned} \frac{P^{poor}(neg|test)}{Prich(neg|test)} \times \frac{P^{poor}(test)}{Prich(test)} &= \frac{P^{poor}(test|neg)}{Prich(test|neg)} \times \frac{P^{poor}(neg)}{Prich(neg)} \\ \frac{0.09}{0.15} \times \frac{0.10}{0.11} &= \frac{P^{poor}(test|neg)}{Prich(test|neg)} \times \frac{P^{poor}(neg)}{Prich(neg)} \\ 0.5 &= \frac{P^{poor}(test|neg)}{Prich(test|neg)} \times \frac{P^{poor}(neg)}{Prich(neg)} \end{aligned}$$

As shown on the previous slide, our estimates imply that

$$P^{rich}(neg) = 2 \times \frac{P^{poor}(test|neg)}{P^{rich}(test|neg)} \times P^{poor}(neg)$$

- If  $P^{poor}(test|neg) = P^{rich}(test|neg)$ , then

$$P^{rich}(neg) = 2 \times P^{poor}(neg)$$

⇒ Top decile is twice as likely to NOT have the coronavirus than the bottom decile. Inequity in the sense that the bottom decile gets tested at a lower rate than corresponds to incidence rate.

Thus egalitarian testing by share in population need not imply egalitarian testing by incidence

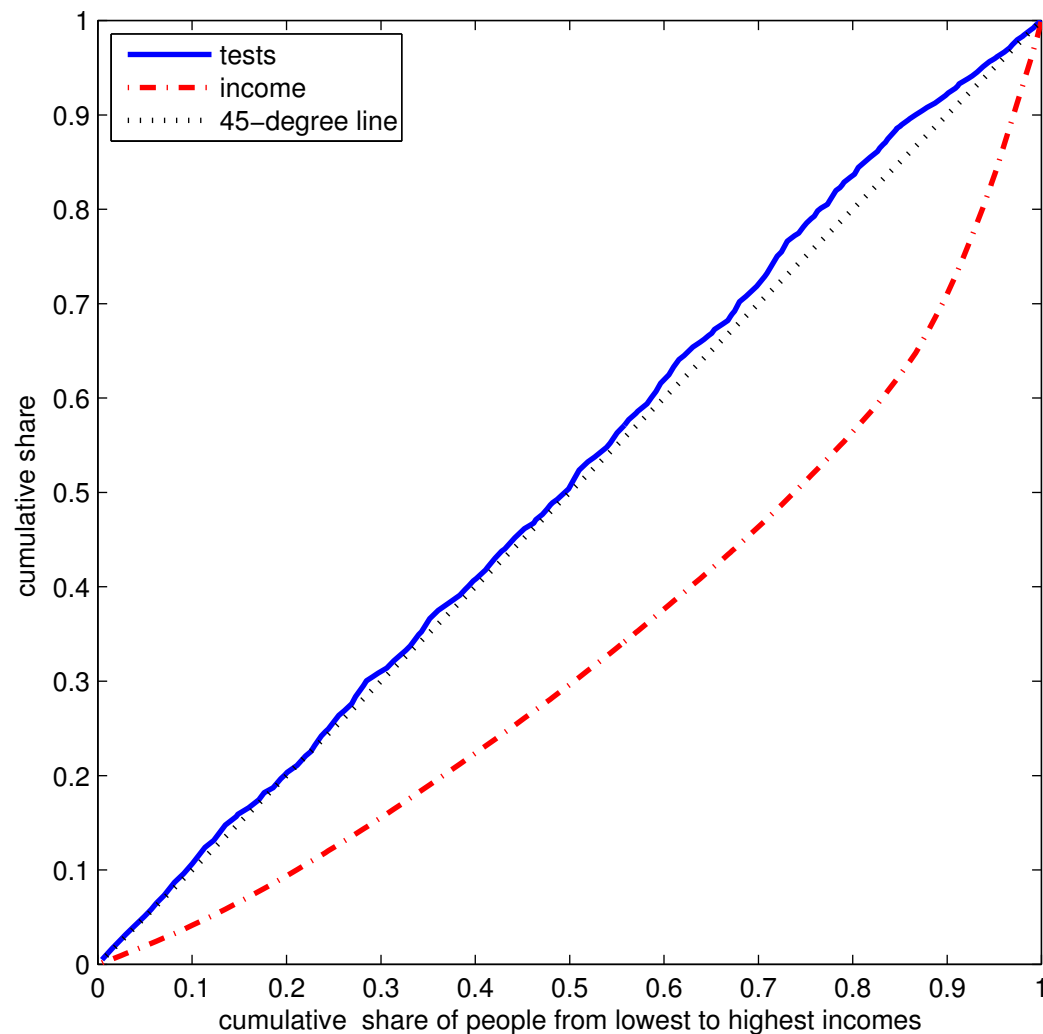
- If  $P^{poor}(test|neg) < P^{rich}(test|neg)$ , then 2 is an upper bound for relative incidence.

Still, egalitarian testing by share in population does not imply egalitarian testing by incidence unless the poor have less than half the chance of getting tested conditional on being negative than the rich.

Dynamics: How do results change if we consider data up until April 13, 2020?

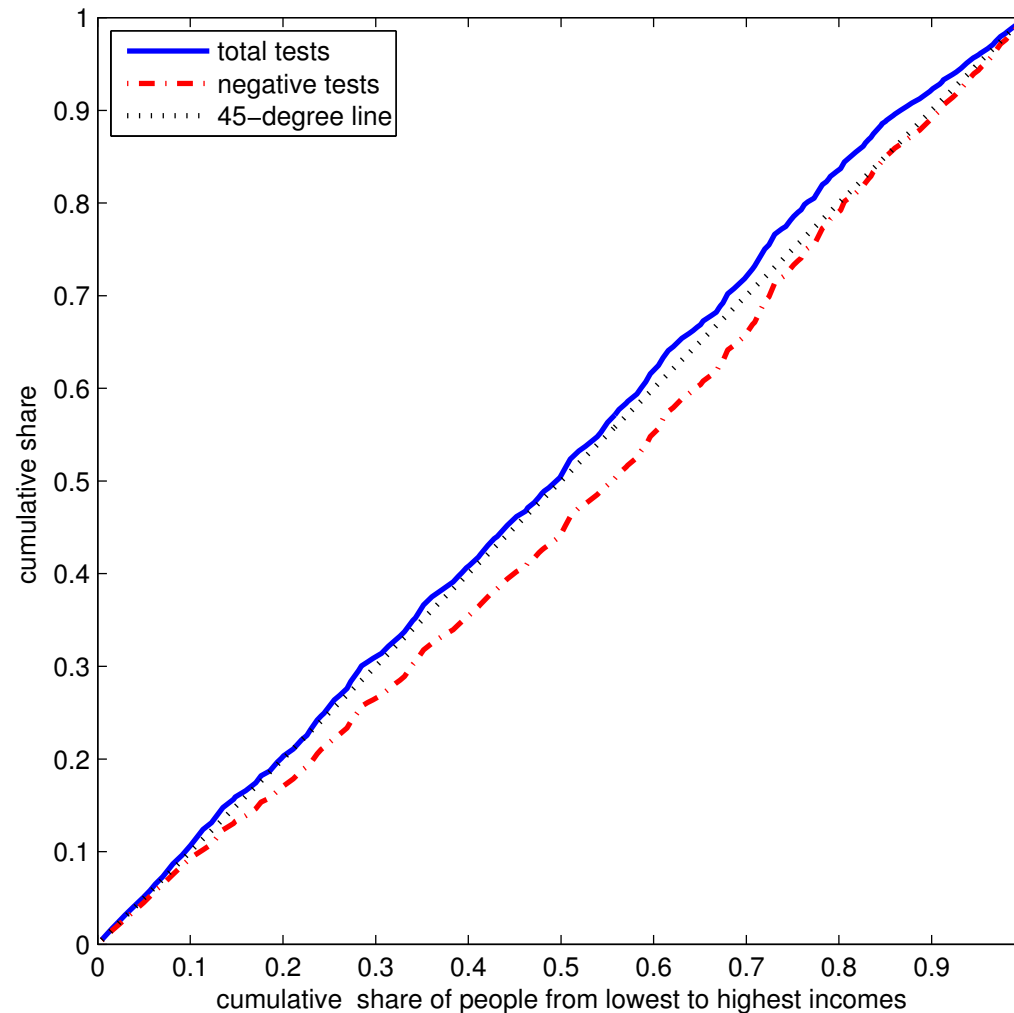
Total tests: 73,215 by April 2; 182,099 by April 13

## Lorenz Curves of Covid-19 Tests and Mean Income Across New York City Zip Codes as of April 13, 2020



Gini: -0.03; bottom decile 11% of tests, top decile 8% of tests.

## Lorenz Curves of Negative Tests and Total Tests for Covid-19 Across New York City Zip Codes, April 13, 2020



Gini=0.05; bottom decile 8% of negatives, top decile 11% of negatives.

## Conclusions:

- NYC population is tested at equal rates across income, Gini coefficient of testing is 0.02.
- Lower income residents have lower share of negative test results, Gini coefficient 0.09. (top decline get 15 percent of all negative tests)
- If lower income groups have higher incidence of coronavirus, then results suggest they get tested less than proportionally to incidence.
- Controlling for income, race does not affect testing or test outcomes much. (Not shown.)

Extras



## Summary Statistics of Zip Code Level Data

Statistic	Covid-19 Total Tests per 100,000	Covid-19 Share of Negative Tests (%)	Per Capita Income dollars of 2018	Population
Mean	908	49	44287	47645
Median	860	49	31779	42653
Std.Dev.	268	9	31919	26698
Max	2390	75	147547	112425
Min	450	23	13394	3028

Notes. Summary statistics are computed across the 177 New York City zip codes. Total tests and negative tests are cumulative as of April 2, 2020. Replication code `summary_statistics.m` in `stu_covid19.zip`.

$y_i^c$  = per capita income in zip code  $i = 1, \dots, 177$ , sorted in ascending order, so that  $y_i^c < y_{i+1}^c$  for any  $1 \leq i < 177$  and let  $p_i$  be the population of zip code  $i$ .

$y_i$  = total income in zip code  $i$

$$y_i = y_i^c p_i$$

The share of income of zip code  $i$  in total New York City income is defined as

$$s_i^y = \frac{y_i}{\sum_{i=1}^{177} y_i}$$

The cumulative income share up to the  $i$ th poorest zip code, denoted  $S_i^y$ , is given by

$$S_i^y = \sum_{k=1}^i s_k^y$$

Similarly, the population share of the  $i$ th poorest zip code, denoted  $s_i^p$  is given by

$$s_i^p = \frac{p_i}{\sum_{i=1}^{177} p_i}.$$

And the cumulative population share up to the  $i$ th poorest zip code, denoted  $S_i^p$ , is given by

$$S_i^p = \sum_{k=1}^i s_k^p.$$

Let  $\tau_i$  denote the number of Covid-19 tests in the  $i$ th poorest zip code. Then, the share of tests in zip code  $i$ , denoted  $s_i^\tau$  is given by

$$s_i^\tau = \frac{\tau_i}{\sum_{i=1}^{177} \tau_i}$$

And the corresponding cumulative share up to the  $i$ th poorest zip code, denoted  $S_i^\tau$ , is

$$S_i^\tau = \sum_{k=1}^i s_k^\tau.$$

The Lorenz curves plots the variables  $S_i^y$  and  $S_i^T$  (vertical axis) against the variable  $S_i^p$  (horizontal axis).

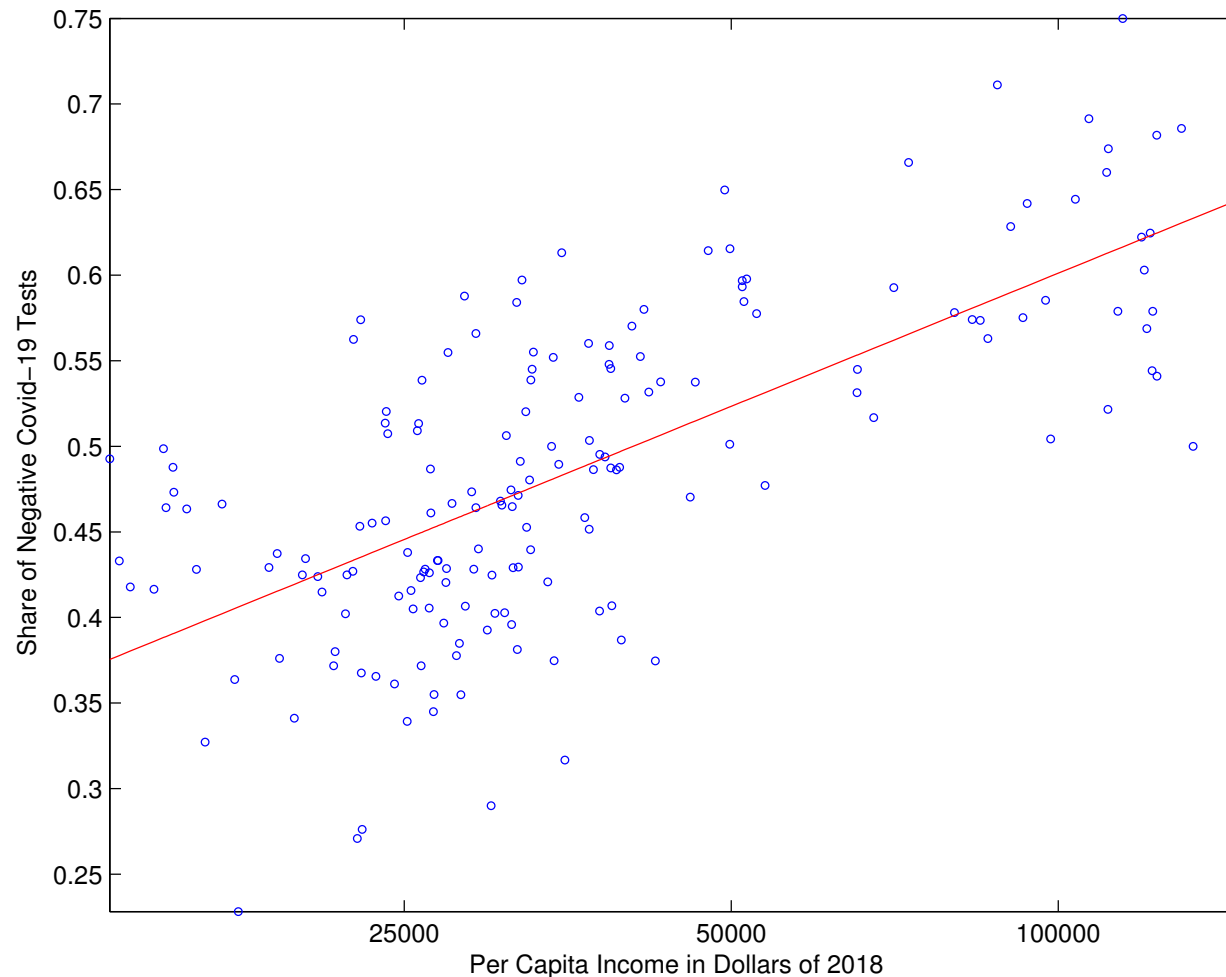
The Gini coefficient of the income distribution across zip codes is measured as

$$\text{Gini coefficient of income distribution} = 1 - \frac{\sum_{i=1}^{177} s_i^p S_i^y}{\sum_{i=1}^{177} s_i^p S_i^p},$$

and the Gini coefficient of the Covid-19 testing distribution across income levels by

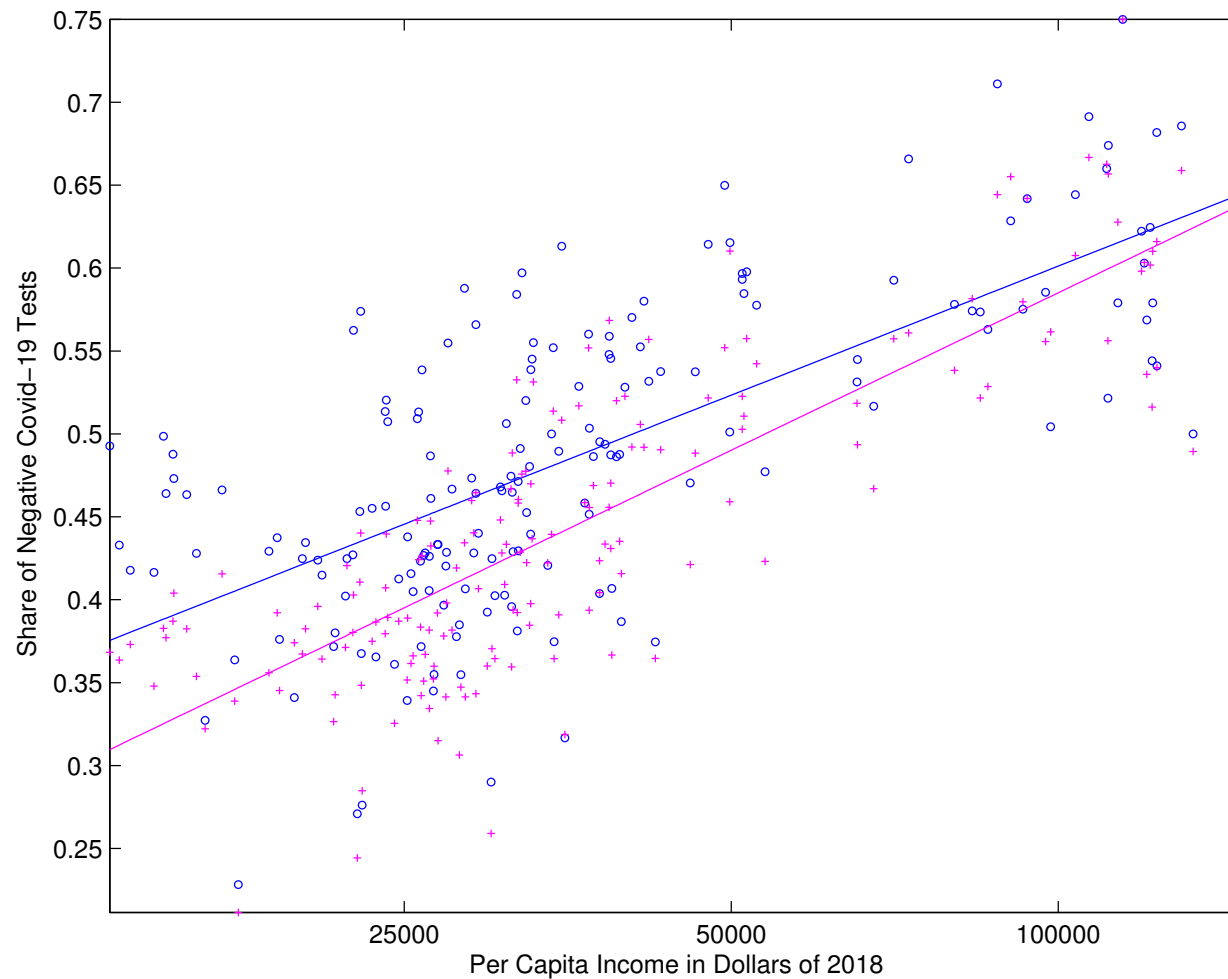
$$\text{Gini coefficient of Covid-19 testing across income levels} = 1 - \frac{\sum_{i=1}^{177} s_i^p S_i^T}{\sum_{i=1}^{177} s_i^p S_i^p}$$

## Share of Negative Tests and Mean Income Per Capita Across New York City Zip Codes



Notes. Own calculations based on data from the New York City Department of Health and Mental Hygiene, as of April 2, 2020, and American Community Survey. The negative share is defined as the number of negative test results divided by the total number of tests. The solid line is the OLS regression. Replication code `negatives_vs_income.m` in `stu_covid19.zip`.

# Share of Negative Tests and Mean Income Per Capita Across New York City Zip Codes, April 2 versus April 13



blue: April 1; magenta: April 13