Racial Impact on Infections and Deaths due to COVID-19 in New York City

Forthcoming in *Harvard Technology Review*

Yunseo Choi
with Prof. James Unwin

MIT PRIMES

October 23, 2020
Racial Disparities of COVID-19 in NYC

CDC Defined Risk Factors

Old age, Underlying health conditions

Deaths

Cases

Rate per 100,000

Whites

Blacks

Traditional risk factors alone do not explain the disparity.
Racial Disparities of COVID-19 in NYC

CDC Defined Risk Factors
- Old age, Underlying health conditions
Racial Disparities of COVID-19 in NYC

CDC Defined Risk Factors
- Old age, Underlying health conditions

Traditional risk factors alone do not explain the disparity.
Environmental Factors

Previous Studies
- Limited access to local healthcare (G. Gee, 2002)
- Poor water and air quality (K. Beyer, 2016)
- Stress (L. L. Black et al., 2015)

Natural Questions
- How does environment/neighborhood play a role?
- How can we quantify and compare neighborhoods?
Environmental Factors

Previous Studies

- Limited access to local healthcare (G. Gee, 2002)
- Poor water and air quality (K. Beyer, 2016)
- Stress (L. L. Black et al., 2015)
Environmental Factors

Previous Studies

- Limited access to local healthcare (G. Gee, 2002)
- Poor water and air quality (K. Beyer, 2016)
- Stress (L. L. Black et al., 2015)
Environmental Factors

Previous Studies

- Limited access to local healthcare (G. Gee, 2002)
- Poor water and air quality (K. Beyer, 2016)
Environmental Factors

Previous Studies

- Limited access to local healthcare (G. Gee, 2002)
- Poor water and air quality (K. Beyer, 2016)
- Stress (L. L. Black et al., 2015)
Environmental Factors

Previous Studies

- Limited access to local healthcare (G. Gee, 2002)
- Poor water and air quality (K. Beyer, 2016)
- Stress (L. L. Black et al., 2015)

Natural Questions
Environmental Factors

Previous Studies
- Limited access to local healthcare (G. Gee, 2002)
- Poor water and air quality (K. Beyer, 2016)
- Stress (L. L. Black et al., 2015)

Natural Questions
- How does environment/neighborhood play a role?
Environmental Factors

Previous Studies
- Limited access to local healthcare (G. Gee, 2002)
- Poor water and air quality (K. Beyer, 2016)
- Stress (L. L. Black et al., 2015)

Natural Questions
- How does environment/neighborhood play a role?
- How can we quantify and compare neighborhoods?
Residential Redlining

Source. Washington Post
Residential Redlining

Barred Black individuals from entering White communities

Source. Washington Post
Residential Redlining

● Barred Black individuals from entering White communities
● Transparency through Home Mortgage Disclosure Act (HMDA) in 1975

Source. Washington Post
Current Efforts & Importance

Disease Modeling

Only a handful make use of the HMDA database.

COVID-19 Response

Identify individuals according to traditional risk factors.

Importance of Our Work

Provide a new measure to quantify the vulnerability of a community.

Ensure that racial differences are not what guarantee good healthcare through policies.
Disease Modeling

- Only a handful make use of the HMDA database
Current Efforts & Importance

Disease Modeling

- Only a handful make use of the HMDA database

COVID-19 Response

- Identify individuals according to traditional risk factors

Importance of Our Work
- Provide a new measure to quantify the vulnerability of a community
- Ensure that racial differences are not what guarantee good healthcare through policies
Current Efforts & Importance

Disease Modeling
- Only a handful make use of the HMDA database

COVID-19 Response
- Identify individuals according to traditional risk factors

Importance of Our Work
- Provide a new measure to quantify the vulnerability of a community
- Ensure that racial differences are not what guarantee good healthcare through policies
Multi-level Logistical Regression

Census Tracts were numbered from 1 to 2095. Individuals were sorted into Census Tracts and were numbered from 1 to 208,960.
Multi-level Logistical Regression

Data

- Census Tracts were numbered from 1 to 2095.
- Individuals were sorted into Census Tracts and were numbered from 1 to 208,960.
Data

- Census Tracts were numbered from 1 to 2095.
- Individuals were sorted into Census Tracts and were numbered from 1 to 208,960.
Multi-level Logistical Regression

Data
- Census Tracts were numbered from 1 to 2095.
- Individuals were sorted into Census Tracts and were numbered from 1 to 208,960.

Multi-level
Level 1: NYC. Level 2: Census Tracts.
Multi-level Logistical Regression

(Schematic)

Curve of best fit

\[ P(y=1) = e^{\alpha x + \beta} \]

\[ \log\left( \frac{P(y=1)}{1 - P(y=1)} \right) = \alpha x + \beta \]
Multi-level Logistical Regression

(Schematic)

\[ P(y=1) = e^{\alpha x + \beta} \]

\[ \log \left( \frac{P(y=1)}{1 - P(y=1)} \right) = \alpha x + \beta \]
Multi-level Logistical Regression

(Schematic)

Outcome

Acceptance

Denial

Loan income

P(y=1) = e^{\alpha x + \beta} e^{\alpha x + \beta} + 1

P(y=1) = \frac{1}{1 + e^{\alpha x + \beta}}

Curve of best fit
Multi-level Logistical Regression

(Schematic)

Curve of best fit

\[ P(y=1) = \frac{e^{\alpha x + \beta}}{e^{\alpha x + \beta} + 1} \]
Multi-level Logistical Regression

(Schematic)

Curve of best fit

\[ P(y=1) = \frac{e^{\alpha x + \beta}}{e^{\alpha x + \beta} + 1} \]

\[ \log \left( \frac{P(y=1)}{1-P(y=1)} \right) = \alpha x + \beta \]
Our Model

Level 1 Equation

\[
\log\left( \frac{p_{ij}}{1 - p_{ij}} \right) = \beta_{0j}^r + \beta_{1j}^r \cdot r_{ij} + \beta_{2j}^s \cdot s_{ij} + \beta_{3j}^l \cdot l_{ij}
\]

- \( r_{ij} \): race of the applicant \( i \) in census tract \( j \) (1 = white, 0 = black).
- \( s_{ij} \): sex of the applicant (1 = male, 0 = female).
- \( l_{ij} \): ratio of requested loan to income.

Level 2 Equation

\[
\beta_{kj} = \gamma_{k0} + u_{kj} \quad \text{for } k > 0.
\]

- \( \gamma_{k0} \) is fixed over NYC.
- \( u_{kj} \) shows the variation across Census Tracts.

Racial variations of COVID in NYC
Our Model

Level 1 Equation

\[
\log[p_{ij}/(1 - p)_{ij}] = \beta_{0j} + \beta_{1j}r_{ij} + \beta_{2j}s_{ij} + \beta_{3j}l_{ij}
\]

- \(r_{ij}\): race of the applicant \(i\) in census tract \(j\) (1 = white, 0 = black).
- \(s_{ij}\): sex of the applicant (1 = male, 0 = female).
- \(l_{ij}\): ratio of requested loan to income.
Our Model

Level 1 Equation

\[
\log[p_{ij}/(1 - p)_{ij}] = \beta_0j + \beta_1jr_{ij} + \beta_2js_{ij} + \beta_3jl_{ij}
\]

- \(r_{ij}\): race of the applicant \(i\) in census tract \(j\) (1 = white, 0=black).
- \(s_{ij}\): sex of the applicant (1 = male, 0=female).
- \(l_{ij}\): ratio of requested loan to income.

Level 2 Equation

\[
\beta_{kj} = \gamma_{k0} + u_{kj} \text{ for } k > 0.
\]

- \(\gamma_{k0}\) is fixed over NYC.
- \(u_{kj}\) shows the variation across Census Tracts.
Our Model

Level 1 Equation

\[
\log[p_{ij}/(1 - p)_{ij}] = \beta_0 + \beta_1 r_{ij} + \beta_2 s_{ij} + \beta_3 l_{ij}
\]

- \(r_{ij}\): race of the applicant \(i\) in census tract \(j\) (1 = white, 0 = black).
- \(s_{ij}\): sex of the applicant (1 = male, 0 = female).
- \(l_{ij}\): ratio of requested loan to income.

Level 2 Equation

\[
\beta_{kj} = \gamma_{k0} + u_{kj} \quad \text{for } k > 0.
\]

- \(\gamma_{k0}\) is fixed over NYC.
- \(u_{kj}\) shows the variation across Census Tracts.
Our Model

Level 1 Equation

\[ \log \left[ \frac{p_{ij}}{(1 - p)_{ij}} \right] = \beta_{0j} + \beta_{1j} r_{ij} + \beta_{2j} s_{ij} + \beta_{3j} l_{ij} \]

- \( r_{ij} \): race of the applicant \( i \) in census tract \( j \) (1 = white, 0 = black).
- \( s_{ij} \): sex of the applicant (1 = male, 0 = female).
- \( l_{ij} \): ratio of requested loan to income.

Level 2 Equation

\[ \beta_{kj} = \gamma_{k0} + u_{kj} \text{ for } k > 0. \]

- \( \gamma_{k0} \) is fixed over NYC.
- \( u_{kj} \) shows the variation across Census Tracts.
# Our Model

## Level 1 Equation

\[
\log\left(\frac{p_{ij}}{1 - p_{ij}}\right) = \beta_0 + \beta_1 r_{ij} + \beta_2 s_{ij} + \beta_3 l_{ij}
\]

- \(r_{ij}\): race of the applicant \(i\) in census tract \(j\) (1 = white, 0 = black).
- \(s_{ij}\): sex of the applicant (1 = male, 0 = female).
- \(l_{ij}\): ratio of requested loan to income.

## Level 2 Equation

\[
\beta_{kj} = \gamma_{k0} + u_{kj} \text{ for } k > 0.
\]

- \(\gamma_{k0}\) is fixed over NYC.
- \(u_{kj}\) shows the variation across Census Tracts.
Our Model

Level 1 Equation

\[
\log[p_{ij} / (1 - p)_{ij}] = \beta_0 j + \beta_1 j r_{ij} + \beta_2 j s_{ij} + \beta_3 j l_{ij}
\]

- \( r_{ij} \): race of the applicant \( i \) in census tract \( j \) (1 = white, 0 = black).
- \( s_{ij} \): sex of the applicant (1 = male, 0 = female).
- \( l_{ij} \): ratio of requested loan to income.

Level 2 Equation

\[
\beta_{kj} = \gamma_{k0} + u_{kj} \text{ for } k > 0.
\]

- \( \gamma_{k0} \) is fixed over NYC.
- \( u_{kj} \) shows the variation across Census Tracts.
Our Model

**Level 1 Equation**

\[
\log\left[ \frac{p_{ij}}{1 - p_{ij}} \right] = \beta_{0j} + \beta_{1j} r_{ij} + \beta_{2j} s_{ij} + \beta_{3j} l_{ij}
\]

- \( r_{ij} \): race of the applicant \( i \) in census tract \( j \) (1 = white, 0 = black).
- \( s_{ij} \): sex of the applicant (1 = male, 0 = female).
- \( l_{ij} \): ratio of requested loan to income.

**Level 2 Equation**

\[
\beta_{kj} = \gamma_{k0} + u_{kj} \text{ for } k > 0.
\]

- \( \gamma_{k0} \) is fixed over NYC.
- \( u_{kj} \) shows the variation across Census Tracts.
Our Model

### Level 1 Equation

\[ \log[p_{ij}/(1 - p)_{ij}] = \beta_{0j} + \beta_{1j} r_{ij} + \beta_{2j} s_{ij} + \beta_{3j} l_{ij} \]

- \( r_{ij} \): race of the applicant \( i \) in census tract \( j \) (1 = white, 0 = black).
- \( s_{ij} \): sex of the applicant (1 = male, 0 = female).
- \( l_{ij} \): ratio of requested loan to income.

### Level 2 Equation

\[ \beta_{kj} = \gamma_{k0} + u_{kj} \text{ for } k > 0. \]

- \( \gamma_{k0} \) is fixed over NYC.
- \( u_{kj} \) shows the variation across Census Tracts.
Our Model

Level 1 Equation

\[ \log[p_{ij}/(1 - p)_{ij}] = \beta_0^j + \beta_1^j r_{ij} + \beta_2^j s_{ij} + \beta_3^j l_{ij} \]

- \( r_{ij} \): race of the applicant \( i \) in census tract \( j \) (1 = white, 0 = black).
- \( s_{ij} \): sex of the applicant (1 = male, 0 = female).
- \( l_{ij} \): ratio of requested loan to income.

Level 2 Equation

\[ \beta_{kj} = \gamma_{k0} + u_{kj} \text{ for } k > 0. \]

- \( \gamma_{k0} \) is fixed over NYC.
- \( u_{kj} \) shows the variation across Census Tracts.
Redlining Index Map
Redlining Index Map

Fig. Redlining index

Fig. Per capita income
Redlining Index

- Ranged from 1.70 to 2.48
- A correlation of 0.68 with per capita income
Mortgage Discrimination in NYC 2013-17

Summary

Although mortgage discrimination is not institutionalized by the government, it is institutionalized in practice.
Summary

Although mortgage discrimination is not institutionalized by the government, it is institutionalized in practice.

<table>
<thead>
<tr>
<th>Year</th>
<th>N</th>
<th>Percent denied</th>
<th>N</th>
<th>Percent denied</th>
<th>Redlining Index (95 CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2013</td>
<td>9930</td>
<td>40.2</td>
<td>46475</td>
<td>23.8</td>
<td>1.88 (1.77, 1.99)</td>
</tr>
<tr>
<td>2014</td>
<td>7203</td>
<td>37.8</td>
<td>29848</td>
<td>23.4</td>
<td>1.93 (1.81, 2.01)</td>
</tr>
<tr>
<td>2015</td>
<td>7487</td>
<td>34.8</td>
<td>32249</td>
<td>20.8</td>
<td>1.95 (1.83, 2.07)</td>
</tr>
<tr>
<td>2016</td>
<td>8090</td>
<td>37.1</td>
<td>32930</td>
<td>20.6</td>
<td>2.19 (2.06, 2.33)</td>
</tr>
<tr>
<td>2017</td>
<td>7200</td>
<td>29.9</td>
<td>27548</td>
<td>17.0</td>
<td>2.06 (1.92, 2.22)</td>
</tr>
</tbody>
</table>
Summary

Although mortgage discrimination isn’t institutionalized by the government, it is institutionalized in practice.

<table>
<thead>
<tr>
<th>Year</th>
<th>N</th>
<th>Percent denied</th>
<th>N</th>
<th>Percent denied</th>
<th>Redlining Index (95 CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2013</td>
<td>9930</td>
<td>40.2</td>
<td>46475</td>
<td>23.8</td>
<td>1.88 (1.77, 1.99)</td>
</tr>
<tr>
<td>2014</td>
<td>7203</td>
<td>37.8</td>
<td>29848</td>
<td>23.4</td>
<td>1.93 (1.81, 2.01)</td>
</tr>
<tr>
<td>2015</td>
<td>7487</td>
<td>34.8</td>
<td>32249</td>
<td>20.8</td>
<td>1.95 (1.83, 2.07)</td>
</tr>
<tr>
<td>2016</td>
<td>8090</td>
<td>37.1</td>
<td>32930</td>
<td>20.6</td>
<td>2.19 (2.06, 2.33)</td>
</tr>
<tr>
<td>2017</td>
<td>7200</td>
<td>29.9</td>
<td>27548</td>
<td>17.0</td>
<td>2.06 (1.92, 2.22)</td>
</tr>
</tbody>
</table>
## Summary

Although mortgage discrimination isn’t institutionalized by the government, it is institutionalized in practice.

<table>
<thead>
<tr>
<th>Year</th>
<th></th>
<th>Applicant Race</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Black</td>
<td></td>
<td>White</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>Percent denied</td>
<td>N</td>
<td>Percent denied</td>
<td>Redlining Index (95 CI)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2013</td>
<td></td>
<td>9930</td>
<td>40.2</td>
<td>46475</td>
<td>23.8</td>
<td>1.88 (1.77, 1.99)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2014</td>
<td></td>
<td>7203</td>
<td>37.8</td>
<td>29848</td>
<td>23.4</td>
<td>1.93 (1.81, 2.01)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2015</td>
<td></td>
<td>7487</td>
<td>34.8</td>
<td>32249</td>
<td>20.8</td>
<td>1.95 (1.83, 2.07)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2016</td>
<td></td>
<td>8090</td>
<td>37.1</td>
<td>32930</td>
<td>20.6</td>
<td>2.19 (2.06, 2.33)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2017</td>
<td></td>
<td>7200</td>
<td>29.9</td>
<td>27548</td>
<td>17.0</td>
<td>2.06 (1.92, 2.22)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**Summary**

Although mortgage discrimination isn’t institutionalized by the government, it is institutionalized in practice.

<table>
<thead>
<tr>
<th>Year</th>
<th>Applicant Race</th>
<th>Redlining Index (95 CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Black</td>
<td>White</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>Percent denied</td>
</tr>
<tr>
<td>2013</td>
<td>9930</td>
<td>40.2</td>
</tr>
<tr>
<td>2014</td>
<td>7203</td>
<td>37.8</td>
</tr>
<tr>
<td>2015</td>
<td>7487</td>
<td>34.8</td>
</tr>
<tr>
<td>2016</td>
<td>8090</td>
<td>37.1</td>
</tr>
<tr>
<td>2017</td>
<td>7200</td>
<td>29.9</td>
</tr>
</tbody>
</table>
COVID-19 Maps

Source of data: Updated daily by the NYC Government for each ZCTA.

Fig. Rate of infection

Fig. % of positive tests

Fig. Rate of deaths

Racial variations of COVID in NYC

Results 10 / 14
Source of data:
Updated daily by the NYC Government for each ZCTA.
COVID-19 Maps

Source of data:
Updated daily by the NYC Government for each ZCTA.

Fig. Rate of infection
Source of data:
Updated daily by the NYC Government for each ZCTA.

**Fig.** Rate of infection

**Fig.** % of positive tests
COVID-19 Maps

Source of data:
Updated daily by the NYC Government for each ZCTA.

Fig. Rate of infection
Fig. % of positive tests
Fig. Rate of deaths
Scatterplots (May 20th)

- Rate of infection: -0.54
- % of positive tests: -0.64
- Rate of deaths: -0.43

Correlation with Index
Scatterplots (May 20\textsuperscript{th})

**Fig.** Rate of infection
Scatterplots (May 20\textsuperscript{th})

Fig. Rate of infection

Fig. % of positive tests
Scatterplots (May 20th)

Fig. Rate of infection
Fig. % of positive tests
Fig. Rate of deaths

Correlation with Index
Rate of infection: -0.54
% of positive tests: -0.64
Rate of deaths: -0.43
Scatterplots (May 20th)

**Fig.** Rate of infection

**Fig.** % of positive tests

**Fig.** Rate of deaths

**Correlation with Index**

- Rate of infection: -0.54
- % of positive tests: -0.64
- Rate of deaths: -0.43
## Correlation over time

<table>
<thead>
<tr>
<th>Date</th>
<th>Pearson Correlation Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>4/1</td>
<td></td>
</tr>
<tr>
<td>4/16</td>
<td></td>
</tr>
<tr>
<td>5/1</td>
<td></td>
</tr>
<tr>
<td>5/16</td>
<td></td>
</tr>
<tr>
<td>5/31</td>
<td></td>
</tr>
<tr>
<td>6/15</td>
<td></td>
</tr>
<tr>
<td>6/30</td>
<td></td>
</tr>
</tbody>
</table>

- Redlining vs. Pos. Tests
- Redlining vs. Rate of Infection
- Redlining vs. Rate of Deaths

### Racial variations of COVID in NYC

Results 12 / 14
Correlation over time

The graph illustrates the Pearson Correlation Coefficient over time between different variables and redlining.

- **Redlining vs. Pos. Tests**
- **Redlining vs. Rate of Infection**
- **Redlining vs. Rate of Deaths**

The correlation coefficients range from -0.2 to -1.0, showing a negative trend over time.
Conclusion

Possible explanations:
- Medical resources are not distributed equally.
- Residents of redlined neighborhoods are less likely to seek medical assistance.
Possible explanations:

Source. The New Yorker
Conclusion

Possible explanations:

- Medical resources are not distributed equally.

Source. The New Yorker
Possible explanations:

- Medical resources are not distributed equally.
- Residents of redlined neighborhoods are less likely to seek medical assistance.

Source: The New Yorker
Acknowledgements

- Prof. James Unwin
- Dr. Slava Gerovitch
- Prof. Pavel Etingof
- Dr. Tanya Khovanova
- MIT PRIMES
- My sister, Yuji