American Journal of Epidemiology

Vol. 190, No. 7
https://doi.org/10.1093/aje/kwaa277
Advance Access publication:
December 29, 2020

Original Contribution

Socioeconomic Disparities in Subway Use and COVID-19 Outcomes in New York City

Karla Therese L. Sy*, **Micaela E. Martinez***, **Benjamin Rader**, **and Laura F. White**

* Correspondence to Karla Therese L. Sy, Department of Epidemiology, Boston University, 715 Albany Street, The Talbot Building T3E & T4E, Boston MA, 02118 (e-mail: rsy@bu.edu); or Dr. Micaela E. Martinez, Department of Environmental Health Sciences, Columbia University Mailman School of Public Health, New York NY 10032 (e-mail: mem2352@cumc.columbia.edu).

Initially submitted June 29, 2020; accepted for publication December 11, 2020.

Using data from New York City from January 2020 to April 2020, we found an estimated 28-day lag between the onset of reduced subway use and the end of the exponential growth period of severe acute respiratory syndrome coronavirus 2 within New York City boroughs. We also conducted a cross-sectional analysis of the associations between human mobility (i.e., subway ridership) on the week of April 11, 2020, sociodemographic factors, and coronavirus disease 2019 (COVID-19) incidence as of April 26, 2020. Areas with lower median income, a greater percentage of individuals who identify as non-White and/or Hispanic/Latino, a greater percentage of essential workers, and a greater percentage of health-care essential workers had more mobility during the pandemic. When adjusted for the percentage of essential workers, these associations did not remain, suggesting essential work drives human movement in these areas. Increased mobility and all sociodemographic variables (except percentage of people older than 75 years old and percentage of health-care essential workers) were associated with a higher rate of COVID-19 cases per 100,000 people, when adjusted for testing effort. Our study demonstrates that the most socially disadvantaged not only are at an increased risk for COVID-19 infection, they lack the privilege to fully engage in social distancing interventions.

COVID-19; health disparities; infectious disease; New York City; SARS-CoV-2; social determinants of health; social epidemiology

Abbreviations: ACS, American Community Survey; aRR, adjusted risk ratio; CI, confidence interval; COVID-19, coronavirus disease 2019; IQR, interquartile range; NYC, New York City; RR, risk ratio; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2; SES, socioeconomic status; ZCTA, zip code tabulation area.

As of August 28, 2020, there were more than 16 million confirmed cases of coronavirus disease 2019 (COVID-19), the disease caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection, worldwide (1). Large metropolitan areas in the United States, including New York City (NYC), have been hit particularly hard. A number of factors promote virus transmission in cities, including high population density and a dense network of individuals, which support long chains of sustained disease transmission (2, 3). The spread of SARS-CoV-2 in big cities like NYC could also be exacerbated by the reliance on public transportation, where riders are sometimes tightly packed in confined spaces, physically unable to space appropriately apart for social distancing. Outside of the pandemic period, 45%–51% of NYC residents reported the subway as their primary transportation to work (4, 5), and more than 1 billion rides are taken annually (6).

Emerging evidence suggests inequities based on race/ethnicity and socioeconomics have put poorer communities and communities of color at higher risk of SARS-CoV-2 infection (7). This may be due to a community's overall ability to stay at home or shelter in place, which may not be feasible or safe for everyone, because of social vulnerabilities (8–10). Moreover, only 25% of workers in the United States are estimated to be able to transition to remote work, meaning there is a continued need for essential workers to leave their home, putting them at increased risk of exposure to SARS-CoV-2 and increasing the likelihood of local transmission within the community. There is potential for increased SARS-CoV-2 exposure in communities of low socioeconomic status (SES), due to a more limited ability to shelter in place (11), which we term "social distancing inequity." Importantly, social distancing inequity may further exacerbate existing health disparities, compounding structural inequalities in the United States. There are, however, limited published data on the intersection of SES and the ability to shelter in place.

The *New York Times* reported that at least 40% of residents fled the wealthiest neighborhoods of NYC after the pandemic hit, whereas few residents from lower income neighborhoods left the city (12). This indicates that declines in subway ridership likely reflect the ability to either stay home within NYC or leave the city to less-dense residences, both afforded by wealth. Public transportation use during the pandemic may facilitate transmission within the city because it reflects not only close contacts on the subway but also the amount to which individuals are required to leave their homes for work and other essential activities. Previous studies of infectious disease transmission have also used public transportation data as a proxy for human mobility and travel patterns (13). Here, we consider subway ridership as a measure of human mobility and the ability of individuals to follow social distancing measures within NYC. We used a publicly available database of subway ridership to explore the associations between human mobility, sociodemographic factors, and COVID-19 incidence.

METHODS

In this study, we used 2 geographic scales: borough (county) and zip code tabulation areas (ZCTAs). First, we conducted a longitudinal analysis using weekly subway ridership and daily COVID-19 cases in 4 NYC boroughs: Bronx, Brooklyn, Manhattan, and Queens. We excluded Staten Island in all analyses because there are no subway connections from Staten Island to the other boroughs.

Next, a cross-sectional analysis was conducted at the ZCTA level using COVID-19 case data reported as of April 26, 2020, and mobility data on the week of April 11, 2020. Zip code tabulation areas follow census-block boundaries, but they are independent of all other statistical and government entities. New York City comprises 5 boroughs and 214 ZCTAs. We removed 29 ZCTAs with no population, based on the American Community Survey (ACS) data, because they are associated with individual buildings (14). We removed from the analysis 61 ZCTAs without a subway station, under the assumption that in areas with no subway stations, people are less likely to travel to a different ZCTA with a subway. According to ACS data for 2018, the median percent subway usage in the 61 ZCTAs was 16.0% (interquartile range (IQR), 6.1–27.6) compared with 50.0% in the ZCTAs with subway stations (IQR, 39.9–59.4) (4). The final analytic data set for the ZCTA-level analysis included 124 ZCTAs.

Data

Mobility data by borough. Weekly transportation data are collated by the Metropolitan Transportation Authority New York City transit and are publicly available (15). The data comprise the number of MetroCard swipes made each week for the 472 subway stations in the NYC subway system, and we aggregated the swipes for each ZCTA. Subway stations were geocoded by the *ggmap* package in R, and the coding was checked manually for accuracy. Subway use for each borough was calculated as the mean of standardized subway ridership of the individual ZCTAs.

We estimated a standardized change in subway use for each ZCTA during the outbreak (week of March 7, 2020 to April 11, 2020) by subtracting the mean weekly subway use pre-shutdown (week of January 4, 2020 to February 29, 2020) from the number of subway swipes each week and dividing by the pre-shutdown period standard deviation; that is: (weekly ridership – weekly pre-shutdown ridership)/(weekly pre-shutdown standard deviation). Standardizing the ridership by the pre-shutdown standard deviation allowed us to view variation in mobility relative to baseline variation. Seasonal differences in subway use did not need to be accounted for, because our study was done within a season, and subway ridership data in previous years have shown little variation during the months of our study (Web Figure 1) (available at https://doi.org/10.1093/aje/kwaa277).

Mobility data by neighborhood. For the cross-sectional ZCTA analysis, our measure of mobility was the standardized change in subway use during the week of April 11, 2020, calculated as: (weekly ridership on April 11, 2020 – weekly pre-shutdown ridership)/(weekly pre-shutdown standard deviation).

COVID-19 data by borough. Longitudinal, daily incident case counts and daily tests at the borough level were available for March 2, 2020, until April 26, 2020, from publicly available data from the NYC Department of Health and Mental Hygiene (16). Our main outcome for the borough analysis was the log of the cumulative case counts, which enabled us to assess the timing of the exponential growth period. We also describe 3 daily time series of 3 additional COVID-19 outcomes: 1) the rate of positive COVID-19 cases per 100,000 population, 2) the percentage of positive tests out of the total number of tests, and 3) the rate of total tests per 100,000 population. The outcome percentage of positive tests took testing capabilities into account, and the rate of testing was used to assess COVID-19 testing volume in each borough.

COVID-19 data by neighborhood. Cumulative COVID-19 case and test data were collated for each ZCTA as of April 26, 2020, from the NYC Department of Health and Mental Hygiene (16). We defined the main COVID-19 outcome as the rate of positive COVID-19 cases per 100,000 population, which we adjusted for number of tests. We also considered 2 additional outcomes relating to COVID-19: 1) the percentage of positive tests among all tests, and 2) the rate of total tests per 100,000 population.

Demographic and SES by neighborhood. We obtained demographic and SES information from the most recent 2014–2018 5-year ACS at the ZCTA level, published by the US Census Bureau (4). We used the R package *tidycensus* to obtain the data. We extracted estimates of population size; median individual income; percentage of individuals older than 75 years; percentage of the population that identifies as Black or African American, Asian, American Indian and Alaska Native, Native Hawaiian and Other Pacific Islander, and/or Hispanic or Latino (i.e., non-White and/or Hispanic/Latino); percentage uninsured; percentage with a high school education or less; percent non–health-care essential workers (employed in the following areas: agriculture, forestry, fishing and hunting, mining, construction, manufacturing, wholesale trade, retail trade, transportation, warehousing, and utilities), and percentage of health-care essential workers (health-care practitioners and technical occupations). Essential worker classifications were based on ACS categories that most closely aligned with the New York State guidance on essential businesses or entities (17).

Statistical analyses

Mobility and COVID-19 across boroughs. We conducted descriptive statistical tests on weekly mobility and daily COVID-19 outcomes and conducted Kruskal-Wallis 1-way analysis of variance tests to assess whether COVID-19 outcomes were significantly different between boroughs. A segmented regression was fit to the mobility data for each borough to estimate the timing of the change in subway travel, or the "breakpoint" (18). We also estimated the end of the exponential growth period as the breakpoint in the log of cumulative cases for a segmented regression for each borough. Using daily COVID-19 case data at the borough level allowed us to granularly assess whether the timing of social distancing may have affected the end of the exponential growth period.

Mobility, sociodemographic variables, and COVID-19 across neighborhoods. We conducted descriptive analyses of NYC subway use and the 3 main cross-sectional COVID-19 outcomes as of April 26, 2020, across ZCTAs. We then examined the cross-sectional association between neighborhood-level ACS sociodemographic variables and change in mobility, using a generalized linear regression model. To assess the association between mobility and the rate of positive COVID-19 tests per 100,000 population, we used a negative binomial generalized linear model with a log link, with ZCTA population as the offset. The same model was used for the outcome rate of total tests per 100,000 population. The analysis of the percentage of positive tests out of the total number of tests was conducted with a generalized linear model with a binomial distribution and a logit link, with total number of tests as weights. All tests of statistical significance were 2 sided, and analyses were conducted in R, version 4.0.0 (R Foundation for Statistical Computing, Vienna, Austria) (19).

Because of the high multicollinearity of the neighborhood-level sociodemographic factors (Web Figure 2), we refrained from including all variables in the model that would produce unstable regression estimates with high variance. Thus, for each regression, we a priori selected 1 mediator to include to assess the direct effect of each predictor on the outcome. We hypothesized that the percentage of the population working in essential services would be the main mediator driving the association between sociodemographic factors and mobility. We also hypothesized that income may be the key mediator of the association between neighborhood-level features and COVID-19. Moreover, we adjusted for testing effort in models with the outcome rate of positive COVID-19 tests per 100,000 population, to account for differential testing within ZCTAs.

RESULTS

Characterizing change in human mobility

The mean subway use during the pre-shutdown period was more than 25 million swipes per week. Overall, this decreased 69.7% to fewer than 8 million by the week of April 11, 2020. The timeline and extent of the reduction in subway use after March 4, 2020, varied greatly among the ZCTAs (Web Figure 3). The mean standardized change in mobility from baseline to April 11 among all ZCTAs was –20.33 (IQR, –25.59 to –16.04). The Bushwick/Bedford-Stuyvesant neighborhood in Brooklyn had the greatest standardized reduction in mobility, with a decrease of 42.86, followed by Upper West Side, Manhattan (–40.47). The areas with the least reductions in mobility were Rockaway, Queens, with a standardized decrease of 2.92, and Fort George, Manhattan (–4.02). Figure 1 shows the geographic variability in standardized changes in mobility during the early stages of the pandemic in NYC (20, 21).

The change in mobility also varied among the boroughs (Web Figure 3). The mean standardized decrease in mobility was greatest in Manhattan neighborhoods with a decline of 22.4 (IQR, –25.65 to –18.44), followed by a reduction of 22.2 (IQR, –28.17 to –14.72) in Brooklyn, 19.3 (IQR, –25.08 to –14.72) in Queens, and 18.6 (IQR, –21.36 to –16.76) in the Bronx.

Characterizing COVID-19 outcomes

By April 26, the neighborhood with the highest rate of COVID-19 (4,044.41 cases/100,000 population) was East Elmhurst, Queens, and the neighborhood in which residents had the highest probability of testing positive for COVID-19 (68.67%) was Corona, Queens. The area with the greatest rate of COVID-19 testing was Co-op City in the Bronx, with 8,262 cases per 100,000 population. The Bronx had the highest incidence of COVID-19 among the 4 boroughs, with 2,472 cases per 100,000 population, followed by Queens (2,149 cases per 100,000 population), Brooklyn (1,653 cases per 100,000 population), and Manhattan (1,292 cases per 100,000 population). The 3 COVID-19 outcomes (i.e., rate of positive COVID-19 cases per 100,000 population, the percentage of positive tests out of the total number of tests, and the rate of total tests per 100,000 population) were all significantly different among the 4 boroughs ($P < 0.001$, $P = 0.03$, and $P = 0.01$, respectively) (Web Figure 4).

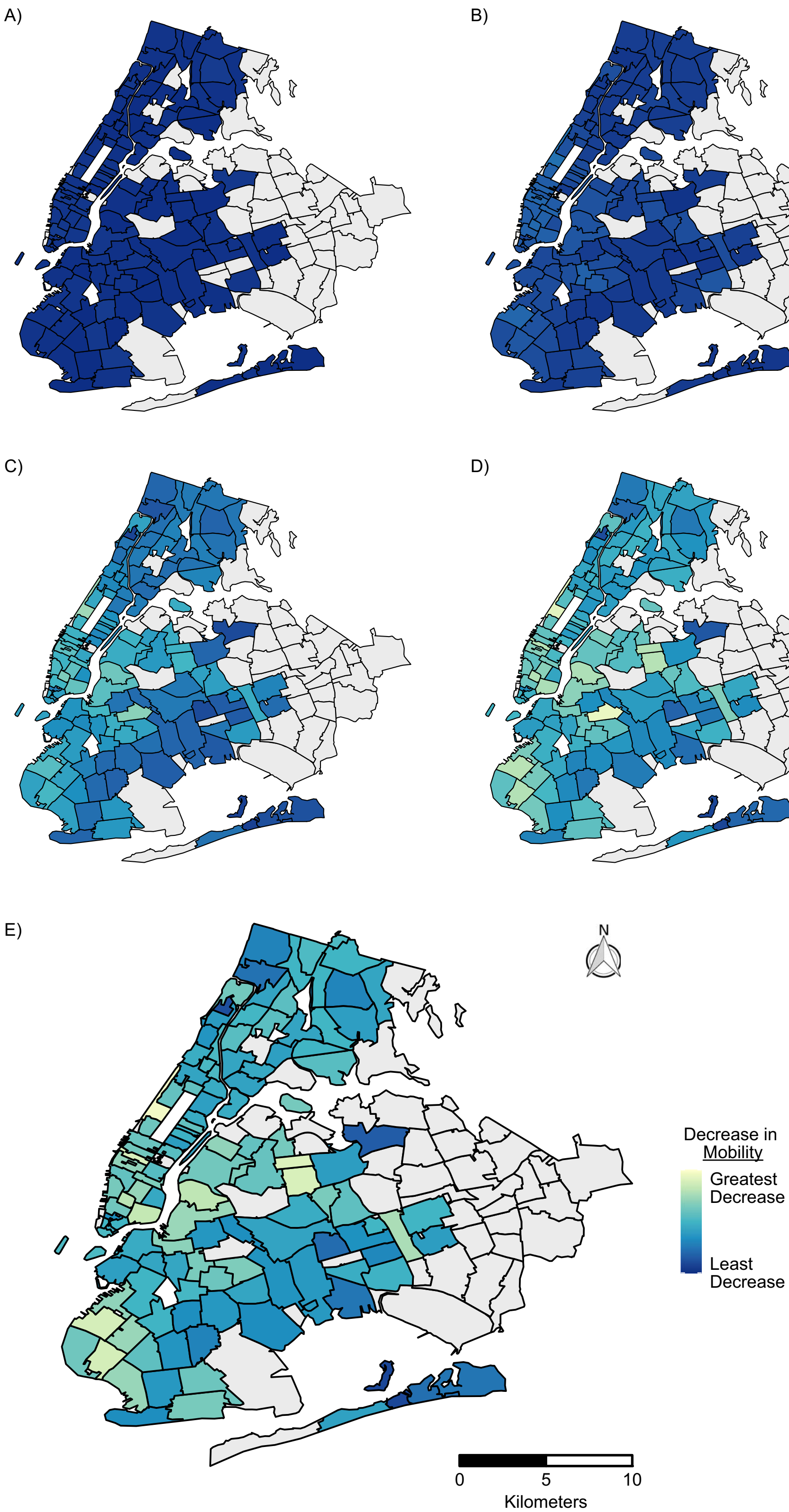


Figure 1. New York City reduction in subway use in zip code tabulation areas during the coronavirus disease 2019 outbreak on the week of A) February 29, 2020; B) March 7, 2020; C) March 14, 2020; D) March 21, 2020; and E) April 11, 2020. Reductions were calculated as the change in subway use relative to the pre-shutdown period and standardized by the pre-shutdown standard deviation. B–D) Maps correspond to key New York City executive orders, as follows: B) local state of emergency, restricted gatherings exceeding 500 persons; C) city school closures; and D) stay-at-home order, nonessential businesses closure (20, 21).

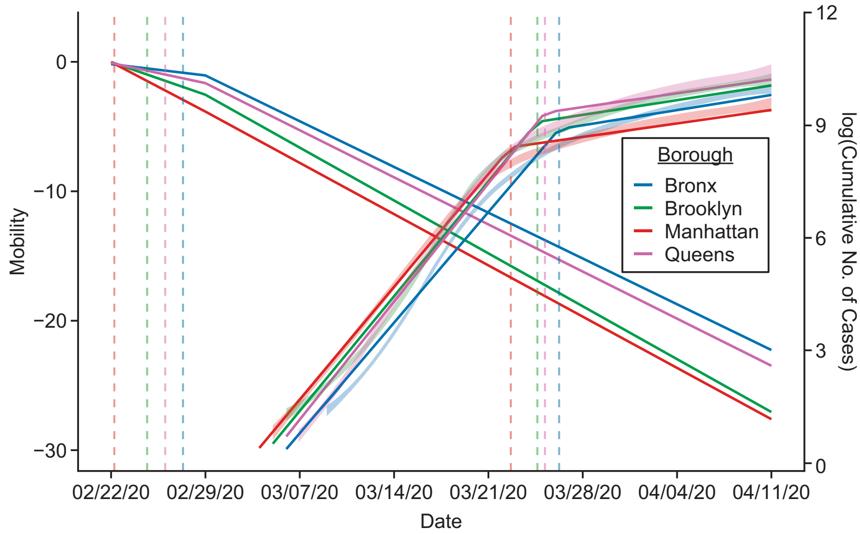


Figure 2. Segmented regression for subway use and log of cumulative cases, by borough, between February 22, 2020, and April 11, 2020. Opaque lines represent the fitted regression estimates, and transparent loess smoothed lines denote empirical case data. Vertical dashed lines indicate the breakpoints of subway use (i.e., date of onset of decline in subway use) and of log of cumulative reported cases (end date of exponential growth period) for each borough.

Borough analysis

Segmented regression of change in mobility and COVID-19 outcomes. Subway ridership reduction in all 4 boroughs occurred within 6 days of each other and was estimated to occur first in Manhattan starting on February 22 (95% confidence interval (CI): February 19, February 24). This was followed by Brooklyn on February 24 (95% CI: February 21, February 27), Queens on February 25 (95% CI: February 22, February 29), and the Bronx on February 27 (95% CI: February 25, February 29) (Figure 2).

There was a delay of approximately 1 month (mean = 28.62 days) between the start of mobility reduction and end of the exponential growth period of reported cases (Figure 2). Manhattan stabilized first on March 22 (95% CI: March 22, March 23), Brooklyn on March 24 (95% CI: March 24, March 25), Queens on March 25 (95% CI: March 24, March 25), and the Bronx on March 26 (95% CI: March 25, March 26). The date of decrease in human movement in the boroughs (Web Figure 5A) corresponded qualitatively to mobility on the week of April 11 (Web Figure 5B), and boroughs in which travel decreased earlier experienced an earlier end of the exponential growth period (Web Figure 5C).

ZCTA analysis

Sociodemographic variables and mobility. The ZCTAs with the highest median income tended to have the greatest decrease in mobility (Figure 3) (20, 21). In unadjusted analyses, lower median income and greater percentages of people working in essential services, health-care essential workers, and non-White/Hispanic individuals were associated with a smaller decrease in mobility. In analyses adjusted for percentage of essential service workers, there were no associations with mobility (Figure 4, Web Table 1).

Mobility, sociodemographic variables, and COVID-19 outcomes. Rate of positive cases per 100,000 population. Smaller decreases in mobility during the pandemic period were significantly associated with the rate of positive cases per 100,000 population in each ZCTA (risk ratio (RR) = 1.13, 95% CI: 1.04, 1.23). This association was slightly attenuated in analysis adjusted for testing (adjusted risk ratio (aRR) = 1.12, 95% CI: 1.05, 1.20), and decreased further when adjusted for testing and median income (aRR = 1.06, 95% CI: 1.01, 1.13) (Figure 4, Web Table 1).

In analysis adjusting for the number of tests performed, all sociodemographic variables except percentage of population uninsured and percentage older than 75 years were independently associated with the rate of positive COVID-19 cases per 100,000 population (Figure 4**,** Web Table 1). An increase in median individual income of $10,000 was associated with a 9% decrease in the rate of COVID-19 (aRR = 0.91, 95% CI: 0.89, 0.94), and an increase in 10% of the population in each ZCTA working in all essential services was associated with a nearly 2-fold increase in the rate of positive cases per 100,000 population (aRR = 1.78, 95% CI: 1.54, 2.07). A greater percentage of the population working in non–health-care essential services and a greater percentage who were non-White/Hispanic, uninsured, and educated to the level of high school or less also increased the rate of positive COVID-19 cases per 100,000 population when adjusted for testing (Figure 4, Web Table 1). When we adjusted for both testing and median individual income, the associations remained except for the percentages of the

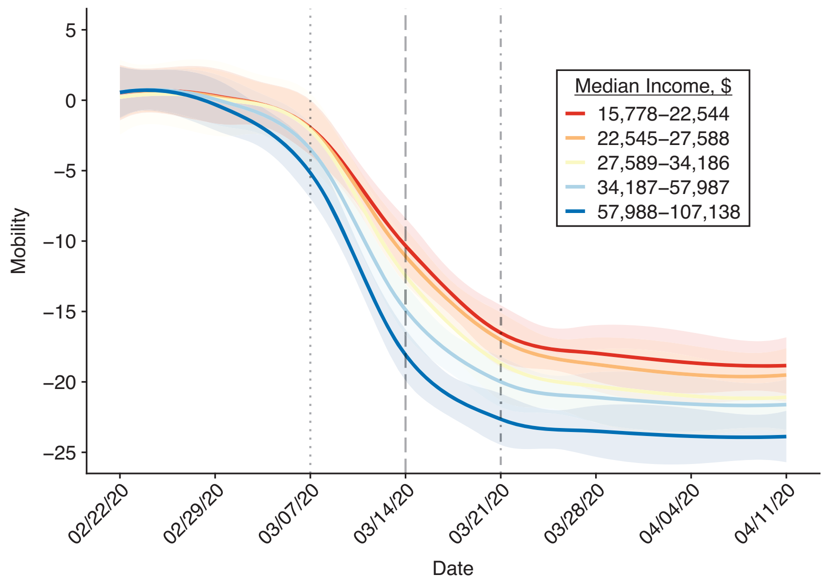


Figure 3. Change in subway use by median income quantiles between February 22, 2020, and April 11, 2020. Loess smoothed lines and associated 95% confidence intervals were fitted over each income group. Vertical lines indicate timing of policies implemented in New York City; the dotted line indicates local state of emergency, the dashed line represents city schools closure, and the dotted-dashed line indicates stay-at-home order (20, 21).

population who were non-White/Hispanic and who were uninsured; however, the percentage of health-care essential workers was associated with increased reported COVID-19 (Web Table 1).

Secondary COVID-19 outcomes (proportion positive among tested and rate of tests per 100,000 population). Smaller decreases in human movement correlated with a greater proportion of COVID-19–positive tests in the unadjusted analysis (RR = 1.04, 95% CI: 1.03, 1.05); when adjusting for income, the proportion of COVID-19 cases among people tested became negatively associated with subway use (aRR = 0.98, 95% CI: 0.97, 0.99) (Web Table 2). All sociodemographic variables were associated with the proportion of positive tests in unadjusted analyses (Web Figure 6, Web Table 2). When adjusted for median income, all associations remained significant.

A smaller decrease in mobility was also associated with an increased rate of COVID-19 testing in both unadjusted (RR = 1.11, 95% CI: 1.03, 1.18) and adjusted (aRR = 1.07, 95% CI: 1.01, 1.14) analyses (Web Table 2). In unadjusted analysis, all sociodemographic variables were associated with the rate of testing (Web Figure 6, Web Table 2). The associations with percentage of the population >75 years old and percentage working in essential services (all, health-care, and non–health-care) remained when adjusted for median income (Web Table 2).

DISCUSSION

On March 22, 2020, after the New York State on PAUSE executive order, a stay-at-home guidance was issued by NYC Mayor Bill de Blasio (21). However, in this study, we show that human movement started declining almost a full month before the guidance was issued and that the order's timing approximately coincided with the end of the exponential growth period of COVID-19. Although these findings suggest many people socially distanced on their own accord in response to the global reports of COVID-19, an individual's ability to fully participate in nonpharmaceutical interventions varied and was associated with SARS-CoV-2 transmission.

Here we report evidence of substantial social distancing inequities throughout NYC neighborhoods. Communities with smaller decreases in mobility were more likely to be socioeconomically disadvantaged, have a greater percentage of persons of color, and have a greater percentage of essential workers, indicating that marginalized communities had reduced ability to shelter in place. However, these associations did not remain when we adjusted analyses for the percentage of the population employed in essential services, suggesting that disparities in mobility reductions are driven by essential work and reduced privilege to socially distance. Furthermore, these same communities were associated with greater COVID-19 burden, even when analyses were adjusted for income, demonstrating inequities in both opportunity to reduce exposure and eventual COVID-19 infection.

Overall, the association between sociodemographic factors and the 3 COVID-19 outcomes were consistent and in the directions expected, with the exception of the finding that areas with a higher percentage of essential workers had a lower percentage of individuals testing positive for COVID-19. These areas may have greater access to testing, increasing the denominator in the calculation of percentage

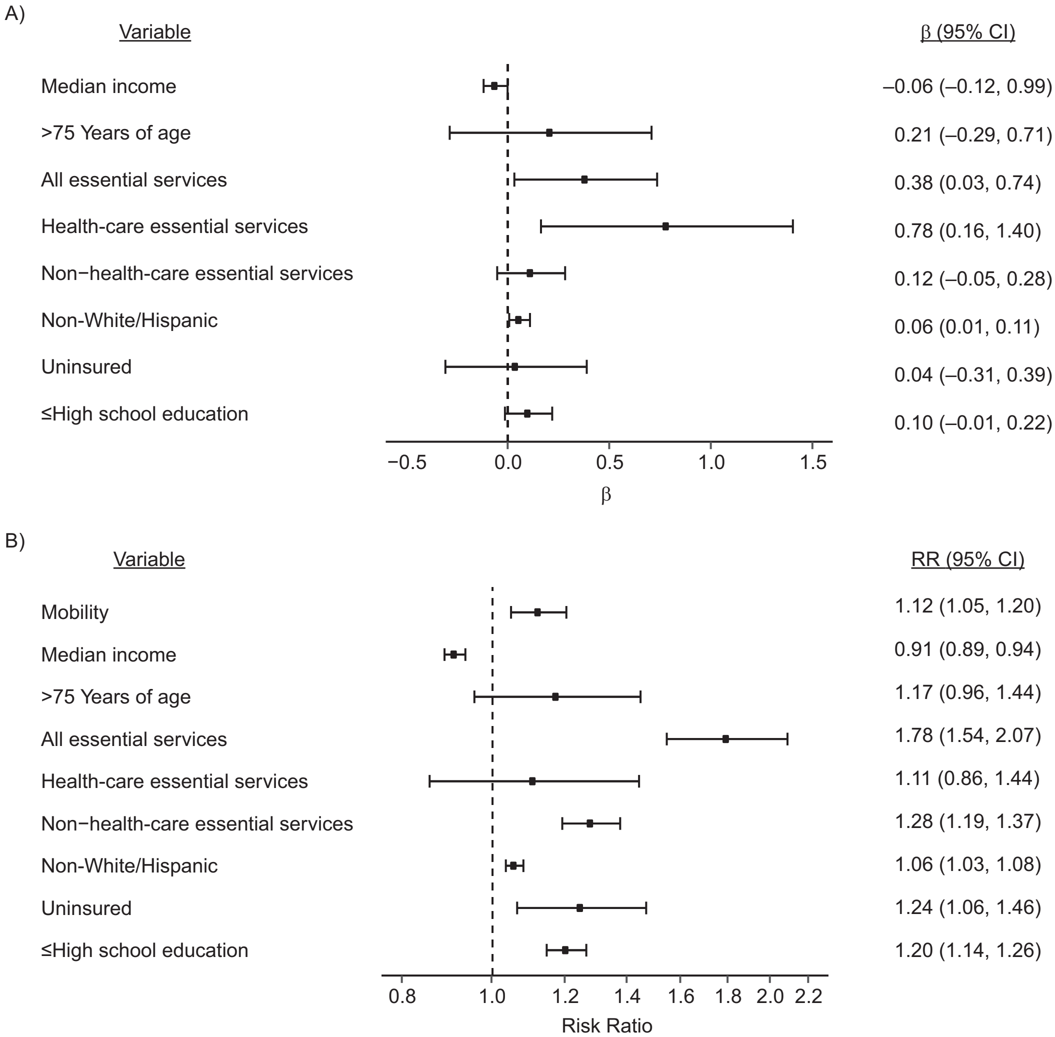


Figure 4. Associations among sociodemographic variables, mobility, and coronavirus disease 2019 (COVID-19) rate per 100,000 population. All COVID-19 models were single-predictor models adjusted for testing to account for differential testing within zip code tabulation areas. The COVID-19 case data were reported as of April 26, 2020, and mobility data were reported the week of April 11, 2020. The subway outcomes were also from single-predictor models (with no adjustments). The estimate for the rate of COVID-19 is a risk ratio (RR) with a null of 1, and the estimate for subway use is a slope (β) with a null of 0. See associated Web table 1 for more details. CI, confidence interval.

testing positive, or greater prioritization of high-risk groups, decreasing the percentage positive among those tested, consistent with our finding of increased testing rates in these areas. Moreover, the rate of SARS-CoV-2 testing is influenced by area-level sociodemographic variables, showing differential testing in NYC even when adjusted for median income.

Our findings are consistent with recent literature describing health disparities in COVID-19 (22, 23). Socioeconomic status is also associated with other underlying comorbidities that may heighten vulnerability to COVID-19 infection and death, such as hypertension, obesity, renal disease, heart disease, and diabetes (24–26). Moreover, the ability to stay at home and physically distance is considerably more difficult not only for those engaged in essential work but also for those with other adverse social determinants, such as food insecurity (8, 27), unstable housing (9, 28), or experiencing domestic violence (10, 29). These risk factors are further compounded by variability in testing access and volume across sociodemographic factors, which have been previously demonstrated across the United States (30, 31) and within NYC (32). The interrelationship of socioeconomic disparities, social distancing inequities, chronic diseases, and COVID-19 is complex, but our findings demonstrate that the COVID-19 pandemic disproportionally has affected the poorest and most vulnerable communities in NYC.

This study has several limitations. One limitation of the subway data is that a swipe represents an individual entering the subway station to take a trip, but the NYC subway only requires individuals to swipe when entering the subway;

therefore, we do not know where trips terminated. Moreover, testing and reporting bias may distort the case counts for ZCTAs and boroughs. We attempted to limit the extent of this distortion by adjusting for the number of tests given, but variability in volume is a function of both resource allocation and response to disease incidence, and thus it is impossible to disentangle these biases. In addition, mild and asymptomatic COVID-19 cases are likely underestimated and it is possible that our findings are related to differential ascertainment rather than true prevalence. However, we found that low-income neighborhoods and communities of color in NYC were hit hardest by COVID-19, even when adjusting for testing effort. This conclusion is supported by a recent seroprevalence report that stated lower-income communities and communities of color had a greater percentage of positive COVID-19 antibody tests (33). In addition, we used cross-sectional population-level data in this study; thus, aggregated population-level risk factors must be interpreted carefully with the knowledge that correlations that arise are not necessarily informative regarding the true mechanisms of SARS-CoV-2 transmission at an individual level. More research at the individual level is needed to elucidate these associations, because COVID-19 seems to entrench existing inequalities and health disparities.

Our findings that social distancing inequities and health disparities are associated with SARS-CoV-2 infection are consistent with previous research in NYC (32, 34). To our knowledge, this is one of the first studies to systematically assess the interrelationship among sociodemographic factors, mobility, and COVID-19. We show a 28-day lag time between the dramatic reduction in subway ridership and the end of the exponential growth phase for reported cases of COVID-19, and that heterogeneity in these reductions are associated with SES. Our study provides additional evidence that the most socially disadvantaged and poorest communities are not only at an increased risk for COVID-19 infection but lack the privilege to fully engage in social distancing interventions, potentially compounding already existing health inequalities. Coronavirus disease 2019 is still a rapidly worsening crisis; to effectively fight this pandemic, sociodemographic and health disparities must be addressed.

ACKNOWLEDGMENTS

Author affiliations: Author affiliations: Department of Epidemiology, Boston University School of Public Health, Boston, Massachusetts, United States (Karla Therese L. Sy, Benjamin Rader); Department of Global Health, Boston University School of Public Health, Boston, Massachusetts, United States (Karla Therese L. Sy); Department of Environmental Health Sciences, Columbia University Mailman School of Public Health, New York, New York, United States (Micaela E. Martinez); School of Biosciences and Medicine, Faculty of Health and Medical Sciences, University of Surrey, Guildford, United Kingdom (Micaela E. Martinez); Computational Epidemiology Laboratory, Boston Children’s Hospital, Boston, Massachusetts, United States (Benjamin Rader); and Department of Biostatistics, Boston University School of Public Health, Boston, Massachusetts, United States (Laura F. White).

This work was supported by the National Science Foundation Grants for Rapid Response Research (grant 2029421 (M.E.M.)); M.E.M. also was supported by the National Institutes of Health (NIH) Office of the Director (grant DP5OD023100). B.R. was supported by Google.org and the Tides Foundation (grant TF2003-089662). L.F.W. was supported by NIH grant R01 GM122876.

The authors thank Kathryn Cordiano and Hanbin Wang Jr. for their thoughtful contribution to the manuscript.

Data availability: All data files to reproduce this work is archived at https://doi.org/10.5061/dryad.vhhmgqnrh. Weekly Metropolitan Transportation Authority New York City transit subway data are publicly available (https://data.ny.gov/Transportation/Fare-Card-History-for-Metropolitan-Transportation-/v7qc-gwpn). New York City Department of Health and Mental Hygiene COVID-19 data are available openly (Link: https://github.com/nychealth/coronavirus-data).

The funding bodies had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

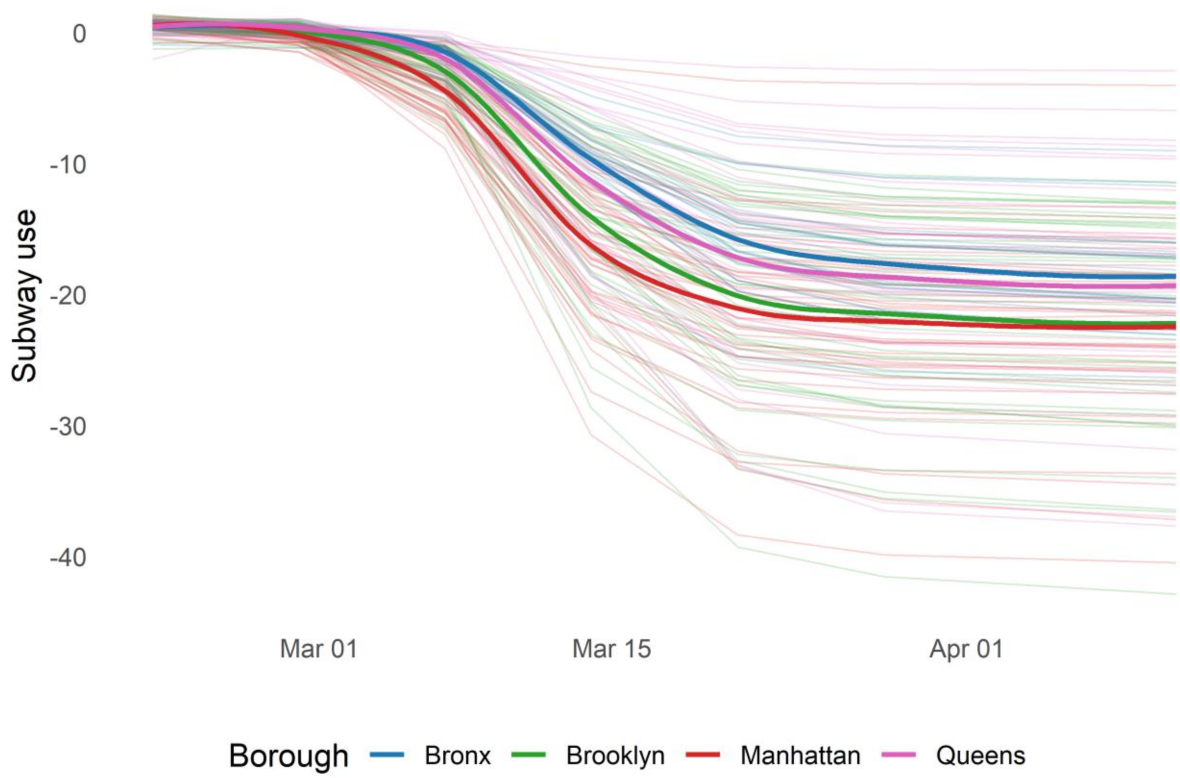
Conflict of interest: none declared.

REFERENCES

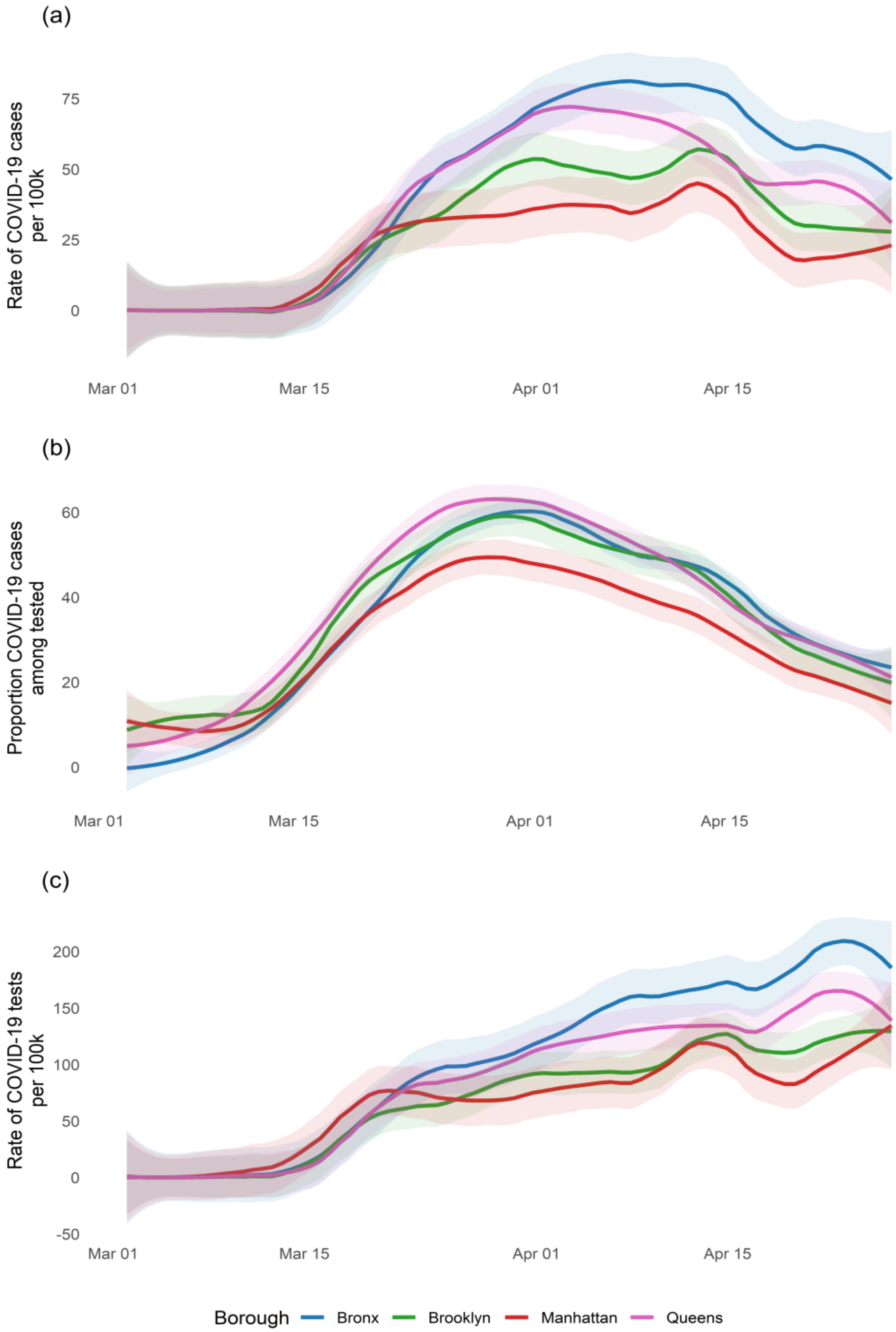
1. Stojkoski V, Utkovski Z, Jolakoski P, et al. The socio-economic determinants of the coronavirus disease (COVID-19) pandemic [preprint]. *medRxiv*. 2020. (https://doi.org/2020.04.15.20066068).
2. Wu T, Perrings C, Kinzig A, et al. Economic growth, urbanization, globalization, and the risks of emerging infectious diseases in China: a review. *Ambio*. 2017;46(1): 18–29.
3. Rader B, Scarpino SV, Nande A, et al. Crowding and the shape of COVID-19 epidemics. *Nat Med*. 2020;26(12): 1829–1834.
4. US Census Bureau. 2012–2016 American Community Survey (ACS) 5-year Estimates. 2020. https://www.census.gov/programs-surveys/acs/data/summary-file.2016.html. Accessed February 26, 2020.
5. New York City Department of Transportation. *2019 Citywide Mobility Survey Results*. New York, New York. 2020. https://www1.nyc.gov/html/dot/downloads/pdf/nycdot-citywide-mobility-survey-report-2019.pdf. Accessed May 21, 2020.
6. Metropolitan Transit Authority. Annual subway ridership. 2020. http://web.mta.info/nyct/facts/ridership/ridership_sub_annual.htm. Accessed May 11, 2020.
7. Blue VJ. Coronavirus deaths in New York increase slightly, Cuomo says. *New York Times*. May 20, 2020. https://www.nytimes.com/2020/05/02/nyregion/coronavirus-new-york-update.html. Accessed May 22, 2020.
8. Dunn CG, Kenney E, Fleischhacker SE, et al. Feeding low-income children during the COVID-19 pandemic. *N Engl J Med*. 2020;382(18):e40.

9. Tsai J, Wilson M. COVID-19: a potential public health problem for homeless populations. *Lancet Public Health*. 2020;5(4):e186–e187.
10. Mazza M, Marano G, Lai C, et al. Danger in danger: interpersonal violence during COVID-19 quarantine. *Psychiatry Res*. 2020;289:113046.
11. US Bureau of Labor Statistics. Workers who could work at home, did work at home, and were paid for work at home, by selected characteristics, averages for the period 2017–2018. 2019. https://www.bls.gov/news.release/flex2.t01.htm. Accessed May 7, 2020.
12. Quealy K. The richest neighborhoods emptied out most as coronavirus hit New York City. *New York Times*. May 15, 2020. https://www.nytimes.com/interactive/2020/05/15/upshot/who-left-new-york-coronavirus.html. Accessed May 19, 2020.
13. Cooley P, Brown S, Cajka J, et al. The role of subway travel in an influenza epidemic: a New York City simulation. *J Urban Health*. 2011;88(5):982–995.
14. United States Census Bureau. *ZIP Code Tabuation Areas (ZCTAs)*. https://www.census.gov/programs-surveys/geography/guidance/geo-areas/zctas.html. Accessed February 26, 2020.
15. New York State Office of Information Technology Services. Fare card history for Metropolitan Transportation Authority (MTA): beginning 2010. 2020. https://data.ny.gov/Transportation/Fare-Card-History-for-Metropolitan-Transportation-/v7qc-gwpn. Accessed February 26, 2020.
16. New York City Department of Health and Mental Hygiene NYC coronavirus (COVID-19) data. 2020. https://github.com/nychealth/coronavirus-data. Accessed May 1, 2020.
17. Randhawa AK, Fisher LH, Greninger AL, et al. Changes in SARS-CoV-2 positivity rate in outpatients in Seattle and Washington State, March 1–April 16, 2020. *JAMA*. 2020;323(22):2334–2336.
18. Bernal JL, Cummins S, Gasparrini A. Interrupted time series regression for the evaluation of public health interventions: a tutorial. *Int J Epidemiol*. 2017;46(1):348–355.
19. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing; 2017.
20. Kully SA. Following orders: a timeline of state and city action on COVID. 2020; https://citylimits.org/2020/03/23/following-orders-a-timeline-of-state-and-city-action-on-covid/. Accessed April 21, 2020.
21. City of New York. Mayor de Blasio issues new guidance to New Yorkers. March 20, 2020. https://www1.nyc.gov/office-of-the-mayor/news/173-20/mayor-de-blasio-issues-new-guidance-new-yorkers.
22. Nayak A, Islam SJ, Mehta A, et al. Impact of social vulnerability on COVID-19 incidence and outcomes in the United States [preprint]. *medRxiv*. 2020. (https://doi.org/10.1101/2020.04.10.20060962). Accessed May 29, 2020.
23. Abedi V, Olulana O, Avula V, et al. Racial, economic and health inequality and COVID-19 infection in the United States [published online ahead of print September 1, 2020]. *J Racial Ethn Health Disparities*. https://doi.org/10.1101/2020.04.10.20060962.
24. Beckles GL, Chou CF. Disparities in the prevalence of diagnosed diabetes - United States, 1999-2002 and 2011-2014. *MMWR Morb Mortal Wkly Rep*. 2016;65(45):1265–1269.
25. Shaw KM, Theis KA, Self-Brown S, et al. Chronic disease disparities by county economic status and metropolitan classification, behavioral risk factor surveillance system, 2013. *Prev Chronic Dis*. 2016;13:E119.
26. Shahu A, Herrin J, Dhruva SS, et al. Disparities in socioeconomic context and association with blood pressure control and cardiovascular outcomes in ALLHAT. *J Am Heart Assoc*. 2019;8(15):e012277.
27. Van Lancker W, Parolin Z. COVID-19, school closures, and child poverty: a social crisis in the making. *Lancet Public Health*. 2020;5(5):e243–e244.
28. Wood LJ, Davies AP, Khan Z. COVID-19 precautions: easier said than done when patients are homeless. *Med J Aust*. 2020;212(8):384–384.e1.
29. Marques ES, Moraes CL, Hasselmann MH, et al. Violence against women, children, and adolescents during the COVID-19 pandemic: overview, contributing factors, and mitigating measures. *Cad Saude Publica*. 2020;36(4):e00074420.
30. Rader B, Astley CM, Sy KTL, et al. Geographic access to United States SARS-CoV-2 testing sites highlights healthcare disparities and may bias transmission estimates. *J Travel Med*. 2020;27(7):taaa076.
31. Souch JM, Cossman JS. A commentary on rural-urban disparities in COVID-19 testing rates per 100,000 and risk factors. *J Rural Health*. 2021;37(1):188–190.
32. Wadhera RK, Wadhera P, Gaba P, et al. Variation in COVID-19 hospitalizations and deaths across New York City boroughs. *JAMA*. 2020;323(21):2192–2195.
33. New York State, Governor's Press Office. Amid ongoing COVID-19 pandemic, Governor Cuomo announces results of state's antibody testing survey at churches in lower-income NYC communities of color show 27 percent of individuals tested positive for COVID-19 antibodies. 2020. https://www.governor.ny.gov/news/amid-ongoing-covid-19-pandemic-governor-cuomo-announces-results-states-antibody-testing-survey#:~:text=Cuomo%20today%20announced%20the%20results,New%20York%20City's%20overall%20population.
34. Kissler SM, Kishore N, Prabhu M, et al. Reductions in commuting mobility correlate with geographic differences in SARS-CoV-2 prevalence in New York City. *Nat Commun*. 2020;11(1):4674–4674.

SUPPLEMENT



Supplementary Figure 1. Relative change in subway use during the early SARS-CoV-2 outbreak (week of February 22 to April 11, 2020). Opaque lines represent loess fitted smoothed lines for the four boroughs, and the transparent lines denote the change in subway use among zip code tabulation areas; subway use change is calculated standardized to usage during January and February.



Supplementary Figure 2. Loess smoothed line and associated 95% confidence intervals of (a) percent of positive cases among tested per 100,000 population; (b) rate of positive COVID-19 cases per 100,000 population; and (c) rate of COVID testing per 100,000 population, among boroughs.

Supplementary Table 1. Associations between sociodemographic variables, subway use and COVID-19 cases per 100,000 population. Each cell is one effect estimate for one model.

	Outcomes				
	Subway use[i]		Rate of positive cases per 100,000 population in the population[ii]		
	Unadjusted *B* (95% CI)	Adjusted *B* (95% CI)[a]	Unadjusted RR (95% CI)	Adjusted RR (95% CI)[b]	Adjusted RR (95% CI)[c]
Sociodemographic variables					
Median income ($10,000)	**-0.06* (-0.12, -0.99)**	-0.04 (-0.12, 0.05)	**0.90**** (0.88, 0.91)**	**0.91**** (0.89, 0.93)**	----[+]
Percent >75 years (10%)	0.26 (-0.24-0.75)	0.27 (-0.22, 0.75)	1.23 (0.98, 1.54)	**1.31** (1.08, 1.59)**	**1.25** (1.06, 1.48)**
Percent in essential services (10%)	**0.35* (0.01, 0.7)**	----[+]	**2.00**** (1.76, 2.26)**	**1.8**** (1.59, 2.04)**	**1.59**** (1.36, 1.86)**
Percent in healthcare essential services (10%)	**0.82** (0.2, 1.44)**	----[+]	1.19 (0.93, 1.53)	**1.36** (1.09, 1.7)**	**1.46** (1.21, 1.77)**
Percent in non-healthcare essential services (10%)	0.10 (-0.07, 0.26)	----[+]	**1.39**** (1.3, 1.48)**	**1.31**** (1.23, 1.4)**	**1.22**** (1.11, 1.33)**
Percent non-white/Hispanic (10%)	**0.06* (0.01, 0.11)**	0.04 (-0.02, 0.1)	**1.09**** (1.07, 1.11)**	**1.07* (1.05, 1.09)**	**1.04**** (1.01, 1.06)**
Percent with uninsured (10%)	0.03 (-0.32, 0.38)	-0.29 (-0.72, 0.14)	**1.34*** (1.14, 1.58)**	1.12 (0.96, 1.32)	**0.84** (0.73, 0.98)**
Percent education high school or less (10%)	0.10 (-0.02, 0.21)	0.01 (-0.18, 0.2)	**1.29**** (1.23, 1.35)**	**1.25**** (1.19, 1.31)**	**1.22**** (1.11, 1.34)**
Human mobility					
Subway use change (10 units)	----[+]	----[+]	**1.12**** (1.03, 1.23)**	**1.11** (1.03, 1.19)**	1.06 (1.00, 1.12)

pvalue * < 0.05; ** < 0.01; *** < 0.001; **** < 0.0001

[+] No adjusted estimate

[i] Linear model

[ii] Generalized linear model with a negative binomial distribution and a log link, total population as offset

[a] Adjusted for percent employed in essential services

[b] Adjusted for number of tests

[c] Adjusted for number of tests and median income

Supplementary Table 2. Associations between sociodemographic variables, subway use change and all COVID-19 outcomes at the ZCTA-level.

	Outcomes								
	Rate of COVID-19 cases per 100,000 population in the population[i]			**Proportion COVID-19 among tested**[ii]		**Rate of tests per 100,000 population in the population**[i]		**Subway use**[iii]	
	Unadjusted RR (95% CI)	Adjusted[a] RR (95% CI)	Adjusted[b] RR (95% CI)	Unadjusted OR (95% CI)	Adjusted[c] OR (95% CI)	Unadjusted RR (95% CI)	Adjusted[c] RR (95% CI)	Unadjusted *B* (95% CI)	Adjusted[d] *B* (95% CI)
Sociodemographic variables									
Median income ($10,000)	**0.90**** (0.88, 0.91)**	**0.91**** (0.89-0.93)**	----[+]	**0.89**** (0.88-0.89)**	----[+]	**0.94**** (0.93-0.96)**	----[+]	**-0.06* (-0.12, 0.99)**	-0.04 (-0.12, 0.05)
Percent >75 years (10%)	1.23 (0.98, 1.54)	**1.31** (1.08, 1.59)**	**1.25** (1.06, 1.48)**	**0.69**** (0.67, 0.71)**	**0.84**** (0.81, 0.86)**	**1.33*** (1.13, 1.57)**	**1.32*** (1.14, 1.52)**	0.26 (-0.24, 0.75)	0.27 (-0.22, 0.75)
Percent in essential services (10%)	**2.00**** (1.76, 2.26)**	**1.80**** (1.59, 2.04)**	**1.59**** (1.36, 1.86)**	**1.69**** (1.65, 1.73)**	**1.34**** (1.3, 1.37)**	**1.50**** (1.35, 1.67)**	**1.39**** (1.2, 1.61)**	**0.35* (0, 0.7)**	----[+]
Percent in healthcare essential services (10%)	1.19 (0.93, 1.53)	**1.36** (1.09, 1.7)**	**1.46** (1.21, 1.77)**	**0.73**** (0.7, 0.75)**	0.97 (0.94, 1)	**1.35** (1.11, 1.63)**	**1.51**** (1.28, 1.8)**	**0.82** (0.2, 1.44)**	----[+]
Percent in non-healthcare essential services (10%)	**1.39**** (1.3, 1.48)**	**1.31**** (1.23, 1.4)**	**1.22**** (1.11, 1.33)**	**1.32**** (1.31, 1.34)**	**1.19**** (1.18, 1.21)**	**1.19**** (1.13, 1.26)**	**1.11* (1.02, 1.21)**	0.10 (-0.07, 0.26)	----[+]
Percent non-white/Hispanic (10%)	**1.09**** (1.07, 1.11)**	**1.07* (1.05, 1.09)**	**1.04**** (1.01, 1.06)**	**1.07**** (1.07, 1.07)**	**1.03**** (1.02, 1.03)**	**1.05**** (1.03, 1.06)**	1.02 (0.99, 1.04)	**0.06* (0.01, 0.11)**	0.04 (-0.02, 0.1)
Percent uninsured (10%)	**1.34*** (1.14, 1.58)**	1.12 (0.96, 1.32)	**0.84** (0.73, 0.98)**	**1.57**** (1.54, 1.59)**	**1.29**** (1.26, 1.31)**	1.06 (0.93, 1.2)	**0.79** (0.7, 0.91)**	0.03 (-0.32, 0.38)	-0.29 (-0.72, 0.14)
Percent education high school or less (10%)	**1.29**** (1.23, 1.35)**	**1.25**** (1.19, 1.31)**	**1.22**** (1.11, 1.34)**	**1.24**** (1.23, 1.25)**	**1.18**** (1.16, 1.2)**	**1.14**** (1.1, 1.19)**	**1.12** (1.03, 1.22)**	0.10 (-0.02, 0.21)	0.01 (-0.18, 0.2)
Human mobility									
Subway use change (10 units)[a]	**1.12**** (1.03, 1.23)**	**1.11** (1.03, 1.19)**	1.06 (1.00, 1.12)	**1.03**** (1.02-1.04)**	**0.97**** (0.96, 0.98)**	**1.1** (1.03, 1.18)**	**1.07* (1.01, 1.14)**		

pvalue * < 0.05; ** < 0.01; *** < 0.001; **** < 0.0001
[+] No adjusted estimate
[a] Adjusted for number of tests
[b] Adjusted for number of tests and median income
[c] Adjusted for median income
[d] Adjusted for percent employed in essential services
[i] Generalized linear model with a binomial distribution and a logit link, total number of tests as weights
[ii] Generalized linear model with a negative binomial distribution and a log link, total population as offset
[iii] Linear model

Author contributions: K. Sy, M. Martinez, L. White contributed to conceptualization. K. Sy and M. Martinez contributed to data acquisition. K. Sy and B. Rader contributed to data analysis. All authors contributed to interpretation of results and manuscript writing.

Sources of funding: This work was supported by NSF RAPID "Transmission and Immunology of COVID-19 in the Pandemic and Post-Pandemic Phase: Real-time Assessment of Social Distancing & Protective Immunity" PI M. Martinez [2029421]. B. Rader acknowledges support of Google.org and the Tides Foundation [TF2003-089662]. M. Martinez acknowledges support by the NIH Office of the Director [DP5OD023100]. L. White was supported by NIH R01 GM122876. The funding bodies had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication. All authors have seen and approved the manuscript.

Competing interests: The authors have declared no conflicts of interest.

Acknowledgements: The authors would like to thank Kathryn Cordiano and Hanbin Wang Jr. for their thoughtful contribution to the manuscript.