Mobility Response to COVID-19-related Restrictions in New York City

Emily Chen
Platial Analysis Lab, McGill University
Montréal, Québec, Canada
emily.chen3@mail.mcgill.ca

Grant McKenzie
Platial Analysis Lab, McGill University
Montréal, Québec, Canada
grant.mckenzie@mcgill.ca

ABSTRACT

The first case of the 2019 novel coronavirus was detected in the United States in January 2020, and since then, efforts to contain the virus, such as stay-at-home policies, have greatly restricted human mobility. While stay-at-home policies and concern over the virus contributed to an increase in time spent at home, little is known as to how a change in home dwell time varied by population. The work presented in this paper seeks to understand the relationships between levels of mobility and socioeconomic and demographic characteristics of communities within New York City from February to April 2020. By analyzing the factors that contributed to changes in home dwell time, this work aims to support policymakers and inform future strategies for infection mitigation. Findings from this research reinforce the need for physical distancing policies that acknowledge the existence of socioeconomic and demographic diversity between not only geographic regions in the U.S. but also within a single city.

CCS CONCEPTS

• **Social and professional topics** → **Geographic characteristics**.

KEYWORDS

mobility, covid-19, demographic, socioeconomic, pandemic

ACM Reference Format:
Emily Chen and Grant McKenzie. 2021. Mobility Response to COVID-19-related Restrictions in New York City. In *2nd ACM SIGSPATIAL International Workshop on Spatial Computing for Epidemiology (SpatialEpi 2021) (SpatialEpi'21), November 2, 2021, Beijing, China.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3486633.3491094

1 INTRODUCTION

Since the United States detected its first case of the 2019 novel coronavirus in January 2020, efforts have been underway to contain the virus. Government policies that limited human interaction were enacted to reduce the spread of the virus and upended daily routines and momentous occasions alike. While a growing body of literature has investigated the relationship between government policy action and mobility patterns [12], few studies have investigated how the mobility of populations within a single city with a large outbreak changed between the start of the pandemic and immediately after the emergence of the virus at a fine-scale. Much of the prior literature has examined the effect of mobility and various explanatory variables on the COVID-19 case positivity growth rate [6, 11, 13]. For instance, a recent study found that decreased mobility had a positive and significant impact on reduced case growth in several U.S. counties [1]. In our work, we used a large sample of place-based mobility data to identify changes in community mobility patterns in New York City (NYC) due to the emergence of the COVID-19 pandemic. In this paper, we explore the data related to the amount of time individuals spend at home, otherwise known as *Home Dwell Time*. In essence, this variable is the inverse of mobility, and we demonstrate that on average, home dwell time increased in New York City once the initial cases of COVID-19 were disclosed. While this is not a particularly novel discovery, it does suggest that government policy actions such as lockdowns and stay-at-home orders did have an impact. However, the specific catalyst for the increase in home dwell time is not the purpose of this work. The research question we will address is: *Which socioeconomic and demographic variables had the greatest effect on a change in mobility in New York City before and after initial cases and implementation of COVID-19-related lockdown measures in March 2020?*

To address this question, we accessed mobility data from the place-based data collection platform *SafeGraph* and the American Community Survey data from the U.S. Census. Using these data, we built a spatial lag regression model using a change in home dwell time as our dependent variable and a set of socioeconomic and demographic independent variables aggregated at the spatial resolution of census tracts.

SpatialEpi'21, November 2, 2021, Beijing, China

ACM ISBN 978-1-4503-9119-1/21/11…$15.00
https://doi.org/10.1145/3486633.3491094

2 RELATED WORK

Analyzing population movement to glean human behavior patterns from aggregated smartphone data became increasingly common leading up to the outbreak of COVID-19 [4]. In the earliest months of the pandemic, several researchers advocated for the analysis of mobile phone surveillance data to predict the spread of COVID-19 and understand population mobility trends [3]. Academic and industry researchers from wide-ranging disciplines and around the world acted upon these sentiments, producing a staggering number of analyses on spatial mobility trends during COVID-19.

A number of studies have examined the effects of mobility reduction on case counts outside of the United States [10, 14]. In a comprehensive review focused on the geospatial and spatial-statistical analysis of the pandemic, Franch-Pardo et al. [7] evaluated numerous scientific articles on the subject and concluded that interdisciplinary action, proactive planning, and international solidarity were of utmost importance for controlling the virus. One notable paper by Pullano et al. [13] examined how mobility in France changed before and during lockdowns based on aggregated cellphone data.

Numerous studies have focused on the spread of COVID-19 in the United States. Recent work [5] sought to understand how the virus spread in ten of the largest U.S. metropolitan areas by constructing fine-grained dynamic mobility networks derived from geolocation data that mapped the hourly movements of 98 million people from neighborhoods to points of interests between March and May 2020. The authors found that their model simulating the spread of SARS-CoV-2 accurately predicted that higher infection rates occurred during the first two months of the pandemic amongst disadvantaged racial and socioeconomic groups because of only differences in mobility. Work by Badr et al. [1] investigated the effect of large-scale social distancing adherence on the spread of COVID-19 in 25 U.S. counties with the highest number of confirmed cases as of mid-April 2020. In their analysis, the authors concluded that social distancing had a significant effect on the spread of COVID-19 and that their findings could translate to other U.S. locations, given the geographical diversity of the counties in their sample set.

Within NYC, Lamb et al. [11] conducted an ecological study of residents using data for the number of daily visits to points of interest (POIs). The authors found that the proportion of the population living in households with more than three inhabitants, the proportion of uninsured 18-64-year-olds, the proportion of population self-identifying as White, and median household income were the four aggregate markers of socioeconomic status that yielded the highest R^2 value across four time periods in April 2020. Their analyses revealed that changes in mobility considered with SES markers explained 56% of the variability in case positivity through 1 April 2020, but then dropped to a rate of explanation for case positivity variability of just 18% by 30 April 2020, suggesting that after COVID-19 cases peaked on 6 April 2020 in NYC, these SES markers became less predictive due to several factors, including greater testing capacity, higher SES areas having lower case positivity due to potentially greater engagement with unwarranted testing, and lower SES areas containing a higher number of actual infections. The authors also found that increased case positivity was independently associated with greater reductions in mobility on 10 April and 20 April, but not on 1 April and 30 April, and they attributed these mixed findings to the correlation between time and a city-wide decrease in case positivity as testing capacity increased.

3 METHODOLOGY

We constrained our area of interest to New York City because it was the epicenter of the COVID-19 outbreak in the United States, with approximately 203,000 cases of laboratory-confirmed COVID-19 reported by the NYC Department of Health and Mental Hygiene between 1 March and 31 May 2020. On March 16, 2020, the NYC school system, gyms, and casinos closed and restaurants and bars were restricted to take-out and delivery services. On March 22, 2020, all non-essential businesses closed, and the *NYC on Pause Program*'s stay-at-home orders went into effect. Building off of these key dates, we identified the month of February as our *before* time period and April as our *after* time period.

Mobility patterns were ascertained via SafeGraph's *Social Distancing Metrics* dataset.[1] Specifically for our use case, we extracted the *median home dwell time* variable as our representation of mobility. SafeGraph collects data using GPS pings from 20 million anonymous cellphone devices across the US at the Census Block Group (CBG) level. To calculate a mobile device's home, SafeGraph determines the device's common nighttime location to a Geohash-7 granularity of about 153 meters by 153 meters. SafeGraph then groups devices into *Home* CBGs based on their common nighttime location and provides aggregated data, every 24-hours, for each CBG. The median home dwell time variable is reported as the sum of the "observed minutes at home across the day (whether or not these were contiguous) to get the total minutes for each device." We then aggregated these daily median home dwell times per census block group to a single value per census tract by taking the median value. This aggregation was done for the months of February and April 2020, independently.

Socioeconomic and demographic data were accessed through the 2019 5-year estimates from the American Community Survey (ACS). ACS data at the Census Tract level were the highest resolution available for our chosen selection of socioeconomic and demographic variables. These data were cleaned to remove null and erroneous values.

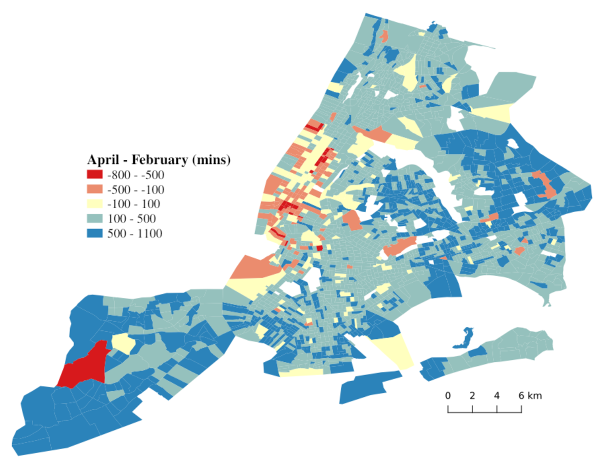

Figure 1: Change in median home dwell time by census tract between February and April 2020. February values subtracted from April values by census tract.

The change in mobility was determined by subtracting the median home dwell times per census tract in February from those in April (Figure 1). A resulting positive value (shown in blue) indicates a decrease in mobility (i.e. an increase in amount of time at home) over the two month time period. Resulting change in median dwell time values ranged between -739 and 1,035 minutes.

Next, we assessed spatial dependence within our dependent variable, change in median home dwell time, by first running an

[1] https://docs.safegraph.com/docs/social-distancing-metrics

ordinary least squares regression model and then using a range of diagnostics, such as *Moran's I*, which indicated strong spatial auto-correlation in the residuals, and Robust Lagrange Multiplier tests for error and lag. Given these results, we constructed both spatial lag and spatial error regression models using the change in median home dwell time as our dependent variable and our set of socioeconomic and demographic variables as our inputs. Our analysis indicated that the spatial lag model outperformed the spatial error model. Multicollinearity within our initial set of variables was assessed through a variance inflation factor (VIF) calculation. After removing *Race - percentage White alone*, all remaining variables reported a VIF under 3.

4 RESULTS

The frequency distribution and descriptive statistics for the change in home dwell time indicate that, overall, most NYC census tracts experienced an increase in home dwell time in April compared to February (mean of 705.5 minutes for median home dwell time in February versus mean of 1125.8 minutes for median home dwell time in April). A value greater than zero suggested that people in that census tract stayed home for longer periods of time in April than in February. Figure 2 shows the histogram of change in median home dwell time in NYC.

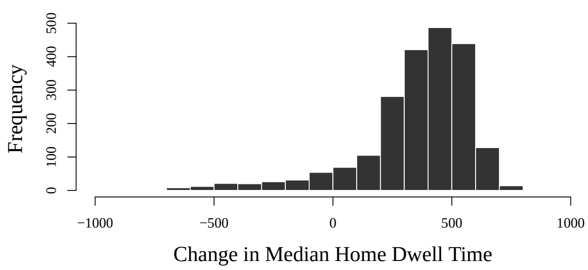

Figure 2: Histogram of the change in median home dwell time by census tract between February and April 2020. February values subtracted from April values by census tract.

The results of our spatial lag regression model are shown in Table 1 and indicate that several socioeconomic and demographic factors significantly contributed to a change in median home dwell time. While the *Nagelkerke pseudo-R squared* value indicates that our selection of variables explains only roughly 40% of the change in median home dwell time, our primary interest in this analysis is the significance and magnitude of the coefficients. Over our nine independent variables, five of them were highly significant ($p<0.01$): percent Black alone, percent Hispanic alone, median household income, percent of the population with a college degree or higher, and percent of the population that is married.

The negative direct impact values associated with the median household income variable indicate that if the income amount in Census Tract A were to increase by $1, A's change in home dwell time would decrease by the estimate of the coefficient. The positive direct impact values associated with percent Black alone, percent Hispanic alone, percent with college degree or higher, and percent married variables indicate that if these percentages were to increase by 1% in Census Tract A, A's home dwell time would also

Table 1: Results of the spatial lag regression model with the difference in home dwell time as dependent variable.

Coefficient	Estimate	Std. Error
Constant	-83.485	59.634
Percent Black alone	1.2323***	0.2161
Percent Asian alone	0.2696	0.2957
Percent Hispanic alone	1.5876***	0.2409
Median Age	-0.9745	0.7432
Median Household Income	-0.0017343***	1.5468e-04
Percent with College Degree +	7.3011***	0.9792
Percent Female	-0.5495	0.9811
Percent Married	10.362***	0.6702
Nagelkerke pseudo-R-squared:		0.3984
Log Likelihood:		-13850.9
Akaike Information Criterion:		27724

*** $p < 0.01$.

increase by the estimate of the coefficient. Notably, percent Asian alone, median age, and percent female were not significant in our regression model.

5 DISCUSSION & CONCLUSIONS

The estimated number of Black alone and Hispanic alone residents in a census tract correlated positively with the change in median home dwell time. This suggests that as the percentage of the Black and/or Hispanic population increases within a community, people in that community stayed at home longer after the emergence of COVID-19 than prior. This finding aligns with some of the other recent research on this topic reporting that there are inequities in how the pandemic has impacted communities [5, 9]. Some of this work found that a higher percentage of COVID-19 cases and related deaths were identified in predominantly Black and Hispanic communities [8]. In light of these findings, one might argue that it is counter-intuitive to discover that communities with higher percentages of Black and/or Hispanic populations increased their time at home with the rise of the COVID-19 pandemic. However, a reduction in time spent outside of the home does not necessarily imply a change in activities outside the home. Furthermore, higher case counts in predominantly Black and Hispanic communities also meant that more people in these communities had to self-isolate, thus increasing overall home dwell time.

Another interesting finding of this analysis is that an increase in median household income aligned with a decrease in median home dwell time after the onset of the pandemic. While the correlation is quite small, it is significant. Again, one could argue that this finding is counter-intuitive, since we might expect individuals with higher incomes to be able to work in jobs that could transition to remote work or take time off from work to reduce exposure to the virus. Existing research has already demonstrated that higher income individuals were more likely to have the opportunity to switch to working from home [15]. However, given the findings from the race and ethnicity coefficients, these results make sense, since median household income often demonstrates a negative correlation with percentage of Black and Hispanic populations [2].

The percentage of the population with a college degree or higher showed a strong significant positive relationship with an increase in home dwell time due to the emergence of COVID-19 in NYC. This finding indicates that the more educated an individual, the more likely they were to stay home longer after the initial cases were identified. These results are intuitive, since one might expected more educated populations to understand and trust the science being presented through government agencies and heed the warnings of public health experts.

The largest positive coefficient, with respect to percentage of a population, was the percent of a community that is married. With each increase in percent of a population married, the amount of time they spent at home after the arrival of COVID-19 increased substantially. This explanatory variable appears to be one of the largest factors for understanding what socioeconomic or demographic factors led to change home dwell time between February and April 2020. Notably, the median age variable, which tends to correlate with marriage, showed no significance in our model.

5.1 Limitations & Future Directions

The representativeness of SafeGraph's data is difficult to ascertain, since much of the data collection process is abstracted from the final product. Despite its exceptional size and granularity, the data came from fewer than 500,000 devices and accounted for only one-ninth of the NYC population. In order to increase the robustness of our findings, future work will involve conducting our spatial analysis using a variety of datasets from different providers. Additionally, summarizing differences in mobility datasets and providing a comprehensive evaluation of the strengths and limitations of each could help researchers choose the most appropriate datasets for their research questions. Another limitation is the potential for additional factors besides stay-at-home restrictions to influence mobility patterns. For example, warmer weather in April could have contributed to greater time spent away from home for some demographics. However, we are currently running the same methodology on data for major cities in Texas and California, both of which have milder winter climates, and preliminary findings suggest similar results to those in NYC. Lastly, aggregated mobility data collected from everyday human behavior patterns are inherently messy. The motivations of every human differ, so the findings that apply to one person might be entirely misaligned with the behavior of someone who has a similar socioeconomic and demographic profile. Thus, drawing conclusions can be difficult when using these types of data.

Future work will involve extending our methodology to data from other cities and regions. A between-city comparison will provide greater insight into how stay-at-home policies affected regions differently based on socioeconomic and demographic patterns, public transit infrastructure, or population density. In addition to comparing cities, there are other explanatory factors that could be added to the regression model. Additional data outside of our current set of variables will be employed in future work. Lastly, there are several types of datasets that could be used to cross-reference these findings and evaluate how other non-pharmaceutical interventions affected mobility. Comparing data such as case counts, vaccine acceptance, or the proportion of mask-wearers with our home dwell time data could be an important next research topic.

5.2 Conclusions

The goal of this work was to identify the ways in which a change in pandemic-induced mobility is affected by socioeconomic and demographic characteristics of a population. This goal has wide-ranging implications for policy makers exploring the varying effects of lockdown measures and designing informed strategies for infection mitigation and safe re-opening. Our findings suggest that there exist significant differences in mobility based on socioeconomic and demographic factors, particularly race, ethnicity, income, educational attainment, and marital status. These results reinforce the need for physical distancing policies that acknowledge the demographic diversity present not only between but also within cities. Future research can both confirm these findings and examine the implications of reduced mobility on the spread of COVID-19 compared with other non-pharmaceutical interventions.

REFERENCES

[1] Hamada S. Badr and Lauren M. Gardner. 2021. Limitations of using mobile phone data to model COVID-19 transmission in the USA. *The Lancet Infectious Diseases* 21, 5 (2021), e113.

[2] Jason P. Block, Richard A. Scribner, and Karen B. DeSalvo. 2004. Fast food, race/ethnicity, and income: a geographic analysis. *American Journal of Preventive Medicine* 27, 3 (2004), 211–217.

[3] Caroline O. Buckee, Satchit Balsari, Jennifer Chan, Mercè Crosas, Francesca Dominici, Urs Gasser, Yonatan H. Grad, Bryan Grenfell, M. Elizabeth Halloran, Moritz U.G. Kraemer, et al. 2020. Aggregated mobility data could help fight COVID-19. *Science* 368, 6487 (2020), 145–146.

[4] Jobie Budd, Benjamin S. Miller, Erin M. Manning, Vasileios Lampos, Mengdie Zhuang, Michael Edelstein, Geraint Rees, Vincent C. Emery, Molly M. Stevens, Neil Keegan, et al. 2020. Digital technologies in the public-health response to COVID-19. *Nature Medicine* 26, 8 (2020), 1183–1192.

[5] Serina Chang, Emma Pierson, Pang Wei Koh, Jaline Gerardin, Beth Redbird, David Grusky, and Jure Leskovec. 2021. Mobility network models of COVID-19 explain inequities and inform reopening. *Nature* 589, 7840 (2021), 82–87.

[6] Yefu Chen, Junfeng Jiao, Shunhua Bai, and Josiah Lindquist. 2020. Modeling the Spatial Factors of COVID-19 in New York City. *Available at SSRN 3606719* (2020).

[7] Ivan Franch-Pardo, Brian M. Napoletano, Fernando Rosete-Verges, and Lawal Billa. 2020. Spatial analysis and GIS in the study of COVID-19. A review. *Science of The Total Environment* 739 (2020), 140033.

[8] Bradley Hardy and Trevon D. Logan. 2020. Racial economic inequality amid the COVID-19 crisis. *The Hamilton Project* (August 2020). https://www.brookings.edu/wp-content/uploads/2020/08/EA_HardyLogan_LO_8.12.pdf

[9] Jordi Honey-Rosés, Isabelle Anguelovski, Vincent K. Chireh, Carolyn Daher, Cecil Konijnendijk van den Bosch, Jill S. Litt, Vrushti Mawani, Michael K. McCall, Arturo Orellana, Emilia Oscilowicz, et al. 2020. The impact of COVID-19 on public space: an early review of the emerging questions–design, perceptions and inequities. *Cities & Health* (2020), 1–17.

[10] Moritz U.G. Kraemer, Chia-Hung Yang, Bernardo Gutierrez, Chieh-Hsi Wu, Brennan Klein, David M. Pigott, Open COVID-19 Data Working Group, Louis du Plessis, Nuno R. Faria, Ruoran Li, et al. 2020. The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science* 368, 6490 (2020), 493–497.

[11] Matthew R. Lamb, Sasikiran Kandula, and Jeffrey Shaman. 2021. Differential COVID-19 case positivity in New York City neighborhoods: Socioeconomic factors and mobility. *Influenza and Other Respiratory Viruses* 15, 2 (2021), 209–217.

[12] Grant McKenzie and Benjamin Adams. 2020. A country comparison of place-based activity response to COVID-19 policies. *Applied Geography* 125 (2020), 102363.

[13] Giulia Pullano, Eugenio Valdano, Nicola Scarpa, Stefania Rubrichi, and Vittoria Colizza. 2020. Evaluating the effect of demographic factors, socioeconomic factors, and risk aversion on mobility during the COVID-19 epidemic in France under lockdown: a population-based study. *The Lancet Digital Health* 2, 12 (2020), e638–e649.

[14] Sakranaik Saraswathi, Amita Mukhopadhyay, Hemant Shah, and TS Ranganath. 2020. Social network analysis of COVID-19 transmission in Karnataka, India. *Epidemiology & Infection* 148, e230 (2020), 1–10.

[15] Scott M. Stringer. 2021. New York City's Frontline Workers. https://comptroller.nyc.gov/reports/new-york-citys-frontline-workers [Online; accessed 15. Sep. 2021].