Living in Boston During COVID-19: The Inequitable Consequences of Vaccination Intentions

Report #7 in a Series

Boston Area Research Initiative

UMASS BOSTON

BOSTON PUBLIC HEALTH COMMISSION

Partnering Institutions

The Boston Area Research Initiative is an interuniversity partnership based at Northeastern University that convenes researchers, policymakers, practitioners, and community leaders to envision and realize the future of the city. Our primary goal is to leverage data and technology to better understand and serve cities, with a focus on enhancing equity, justice, and democracy.

The Center for Survey Research (CSR) at the University of Massachusetts Boston is a full-scale academic survey research center. CSR conducts basic and applied research that contributes to knowledge and understanding of important social issues and supports public and private agencies and university scholars in carrying out high quality policy-related research. Its projects include Beacon, a panel study on Boston neighborhoods.

Boston Public Health Commission, the country’s oldest health department, is an independent public agency providing a wide range of health services and programs. Public service and access to quality health care are the cornerstones of our mission—to protect, preserve, and promote the health and well-being of all Boston residents, particularly those who are most vulnerable.

The Team for this Report

Daniel T. O’Brien, PhD, Associate Professor in the School of Public Policy and Urban Affairs; Director, Boston Area Research Initiative; Northeastern University
Alina Ristea, PhD, Postdoctoral Associate, Boston Area Research Initiative, Northeastern University
Yanchao Wang, Doctoral Student, Civil and Environmental Engineering, Northeastern University
Ryan Qi Wang, PhD, Assistant Professor of Civil and Environmental Engineering; Associate Director of Research on Social Media, Boston Area Research Initiative; Northeastern University
Jianxi Gao, PhD, Assistant Professor of Computer Science, Rensselear Polytechnic Institute
Russell K. Schutt, PhD, Professor, Department of Sociology, University of Massachusetts Boston; Clinical Research Scientist I, Beth Israel Deaconess Medical Center, Harvard Medical School.
Lee Hargraves, PhD, Interim Director, Center for Survey Research, University of Massachusetts Boston
Dan Dooley, Director, Research and Evaluation Office, Boston Public Health Commission
Floyd (Jack) Fowler, PhD, Senior Research Fellow, Center for Survey Research, University of Massachusetts Boston
Anthony Roman, MA, Senior Research Fellow, Center for Survey Research, University of Massachusetts Boston
Mehrnaz Amiri, Research Assistant and Student in the Masters of Science in Urban Informatics, Boston Area Research Initiative, Northeastern University
Sage Gibbons, Research Assistant and Student in the Masters of Science in Urban Informatics, Boston Area Research Initiative, Northeastern University
Hannah Grabowski, Research Assistant and Student in Graduate Program in Applied Sociology, Dept. of Sociology, University of Massachusetts Boston
Nikola Kovacevic, MA, Assistant Study Director, Center for Survey Research, University of Massachusetts Boston

Funding

The survey was funded by the National Science Foundation’s Human-Environment and Geographical Sciences (HEGS) program through a grant for rapid-response research (RAPID; Award #2032384)

Executive Summary

In the Summer of 2020, the Boston Area Research Initiative (BARI) at Northeastern University, the Center for Survey Research (CSR) at University of Massachusetts Boston, and the Boston Public Health Commission (BPHC) conducted a survey that captures the experiences of 1626 Bostonians during the first months of the COVID-19 pandemic. We conducted a follow-up survey in the fall in which 864 participants provided further insights on their experiences and attitudes

In the fifth report in this series, we examined vaccination intentions across communities. Here we extend this work by simulating what these variations mean for infection rates across greater Boston during the vaccination rollout, as well as the attainability of herd immunity. To make the simulation more robust, we incorporate similar data generated by two other surveys conducted around the same time. We allowed the simulation to respond to changing circumstances by permitting those who are "on the fence" or uncertain about vaccination to be persuaded as the level of vaccination in their own community goes up. We otherwise assume that rollout will conform to the three-month timeline laid out by the government for Phase 3 and that vaccine efficacy will match the levels reported by pharmaceutical companies.

Main Findings

- **Black and Latinx respondents were less likely to state they would get vaccinated,** with approximately 20% saying they would "definitely not," relative to less than 10% of White and Asian respondents.
- **Vaccination hit a bottleneck in all communities after six weeks**, as all willing individuals had been vaccinated by then and further rollout depended on persuasion.
- **Herd immunity was eventually reached in all communities, but it took up to six months**, even with the assumption that vaccine supply is sufficient to vaccinate the entire population in three months.
- **All results reveal stark disparities by race.**
    - **Communities of color reached the bottleneck in vaccination earlier**, at which point a lower level of overall vaccination undermines persuasion, setting them back even further.
    - **Infection rates are still higher in high Black and Latinx communities after three months**, even relative to expectations in a no-vaccine scenario.
    - **The average high Black or Latinx community reached herd immunity a month-and-a-half after the average predominantly White community.**
- **Increased persuasion is critical to lowering racial inequities *but* will require convincing those who currently say that they will never get vaccinated.**

Conclusions and Next Steps

The lessons are clear. Vaccination will eventually deliver herd immunity, but there will be racial inequities along the way. Persuasion will help to narrow these gaps, but only so much. This is especially true if we take those who say they will never get vaccinated at their word. Well-crafted, compassionate communication will be needed to sway these individuals if we are to reach herd immunity and protect all of our communities—especially communities of color—from lingering vulnerability to infection.

Table of Contents

1. Living in Boston during COVID-19: A Neighborhood Survey

The NSF-Beacon survey captures the experiences of 1626 Bostonians during the first months of the COVID-19 pandemic, including: their ability and tendency to follow social distancing recommendations; attitudes toward regulations; and the economic and personal impacts of the pandemic. It provides unique insights into how these factors varied across the populations and neighborhoods of a single city—something not currently available from any other source, in Boston or otherwise. The survey was conducted over the summer as a collaboration of the Boston Area Research Initiative (BARI) at Northeastern University, the Center for Survey Research (CSR) at University of Massachusetts Boston, and the Boston Public Health Commission (BPHC). It was funded by the National Science Foundation's Human-Environment and Geographical Sciences (HEGS) program through a grant for rapid-response research (RAPID). The survey used a probability-based random sample stratified by 25 neighborhoods. A follow-up survey in September-November, which had 864 participants, included an item asking whether the respondent planned to receive the COVID-19 vaccine when it became available. Though we have an established process for weighting the data, because we segment the analysis by race, one of the main factors for weighting, we use the raw counts here. More detail on the survey methodology can be found in Appendix A.

This is the seventh in a series of reports[1] describing key insights from the survey. The series focuses especially on the racial and socioeconomic inequities that have exacerbated—and may continue to exacerbate—differential impacts of the pandemic and the associated shutdown. As we seek a thorough understanding of vaccination intentions across the region and by race, we also utilize results from two other surveys conducted in November and December, one by MassInc Polling on behalf of the Boston Museum of Science, the other by Suffolk University Polling and the Boston Globe. Including these permits a more robust set of measures for vaccination intentions by race. Although the exact wording varied between them, all three surveys permitted respondents to say that they planned to get the vaccine "as soon as possible" or "never" get the vaccination or that they were currently uncertain. The responses to this item, including the breakdown by race, were highly consistent across the three surveys.

2. Vaccination Intentions by Race and Geography

Across the three surveys of Boston and Massachusetts residents, 49.6% of respondents said that they planned to get the COVID-19 vaccine, and 8.8% said that they would not; the remaining 41.6% were uncertain. These responses featured prominent disparities by race,

[1] https://cssh.northeastern.edu/bari/projects/covid-19-in-boston/

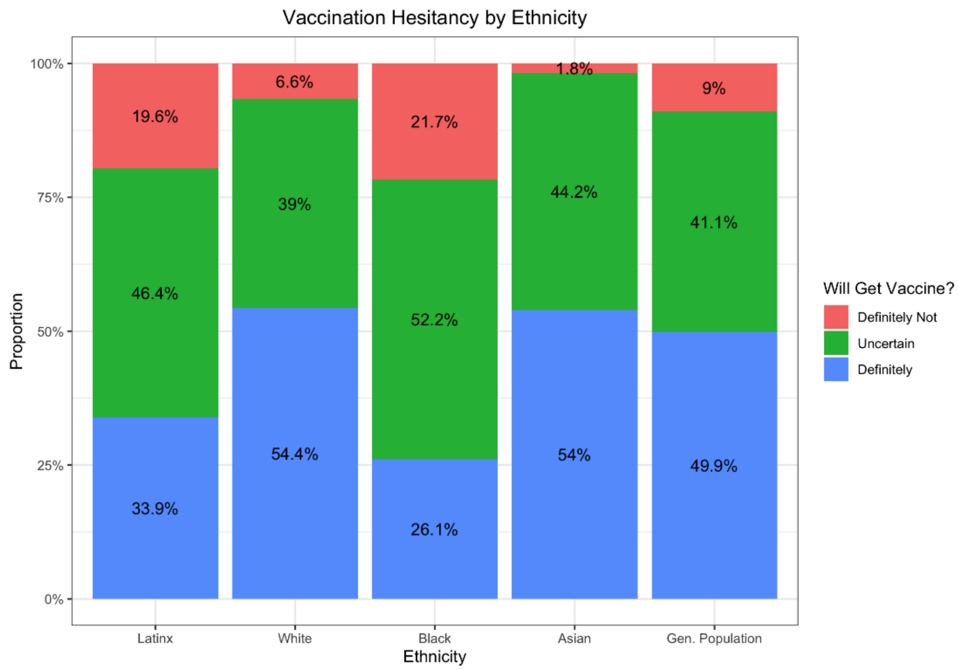

Figure 1. Vaccination intentions across the three surveys for the general population and by race.

however (see Figure 1). On the low end, 6.6% of White respondents and 1.8% of Asian respondents said they would definitely not get the vaccine, compared to 21.7% of Black and 20% of Latinx respondents.

We combined these ratios with the racial composition of communities to estimate the percentage of residents in each who plan to get vaccinated, to definitely not get vaccinated, and who were uncertain. This revealed stark differences across communities, with municipalities in the region varying between 33.9% and 54.1% of the population saying they would definitely get vaccinated, and between 5.4% and 17.2% saying they definitely would not. We saw slightly less receptivity of vaccines in Boston's ZIP codes, which have a higher Black and Latinx population than most surrounding municipalities (definitely will: 28.2% - 52.1%; definitely will not: 5.1% - 19.8%). Figure 2 illustrates these disparities geographically, showing that the percentage who planned to be vaccinated were fewer and those who definitely would not be vaccinated were greatest in Boston—especially the southern majority-minority core of the city—and in the high-minority municipalities to the north that have been hardest hit by the pandemic (e.g., Revere, Chelsea, Lynn).

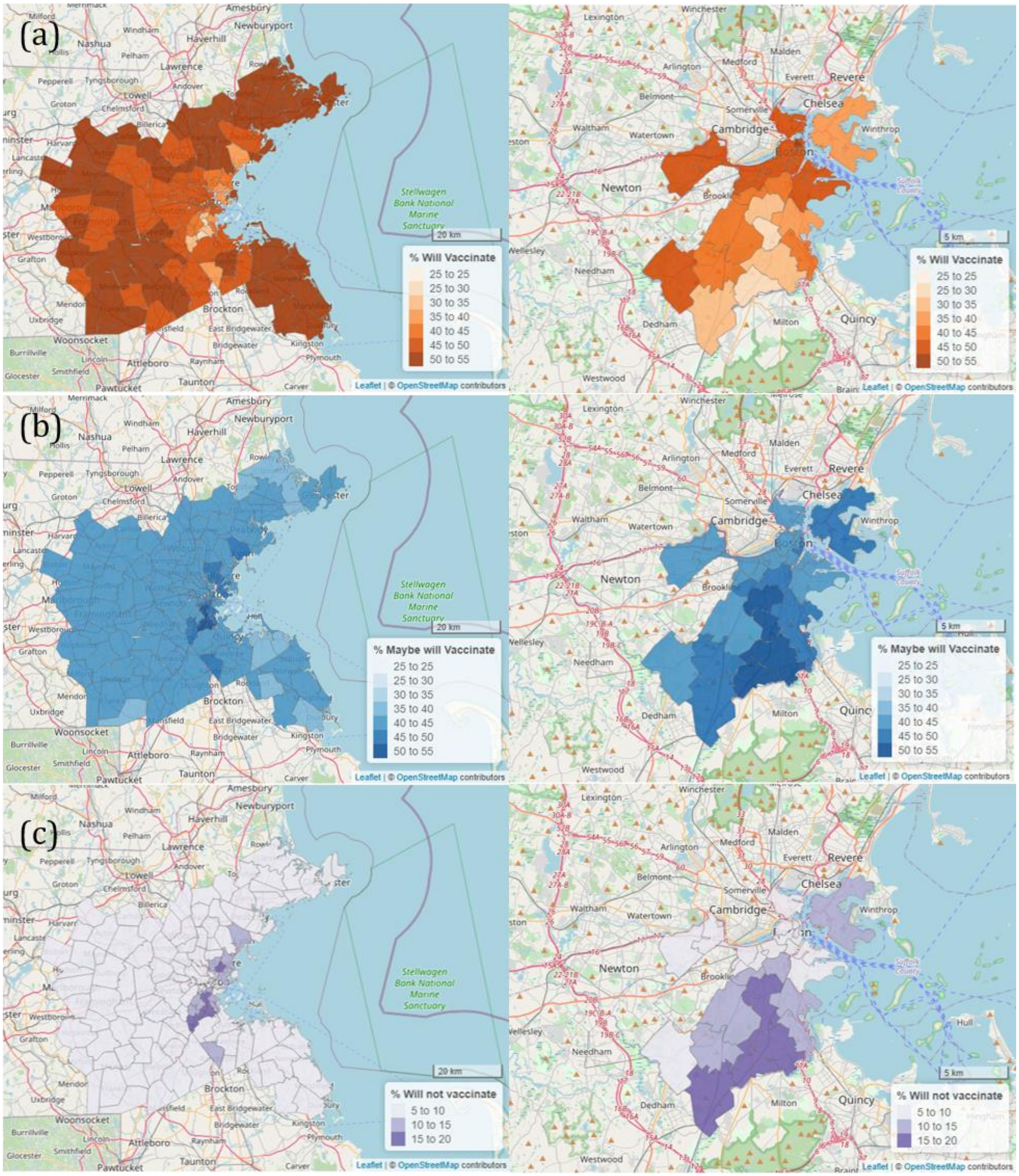

Figure 2. Estimated vaccination intentions by municipalities across the region (left panel) and Boston ZIP codes (right panel), broken down by those who (a) will definitely vaccinate, (b) will maybe vaccinate, and (c) definitely will not vaccinate, according to racial composition and survey responses.

3. Simulating Vaccination across Communities

We simulated the impact of vaccination across communities by incorporating it into a model that estimates the proportion of individuals in a community who are susceptible to infection, currently infected, and recovered. This *SIR model* (i.e., susceptible-infected-recovered) is the type of simulation that public health experts and network scientists have been using to project the trajectory of the pandemic. It also took into consideration movement between communities in approximating the exposure of susceptible individuals and their subsequent likelihood of infection. We ran this model for October-December 2020, using historical infection records and mobility data drawn from cell phone records to make a clear comparison with what would have happened with and without vaccination. The model includes a number of other assumptions and parameters that are described in Appendix B.

Figure 3 depicts a steady increase over the three-month period in the proportion of people who were vaccinated across the region, reaching 75%. In late November, however, the vaccination process hit a bottleneck. It had exhausted all individuals who either were willing to be vaccinated at the outset or were persuaded to that point, as indicated by the blue line reaching zero (*mean* = 53 days into simulation). As a result, vaccination from then on was dependent on those additional individuals who were persuaded in each week, which was less than the rate at which vaccination was possible. This created the kinks in the red and green lines, indicating a slowed vaccination process from that point on, which explains why the simulation did not

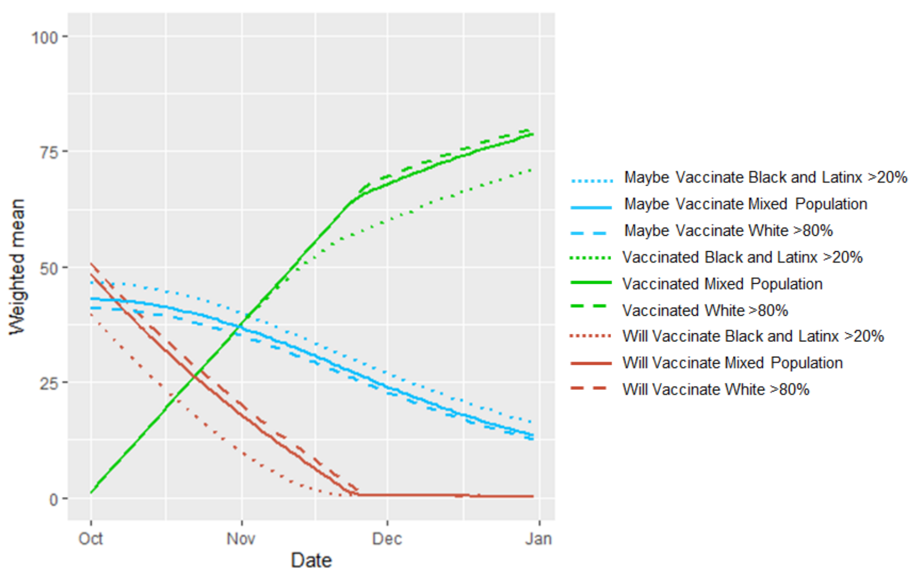

Figure 3. Percentage of residents intending to vaccinate (red line), will maybe vaccinate (blue line), and have been vaccinated (green line) across the three-month simulation, comparing communities that are predominantly White, high Black-Latinx, and other.

successfully vaccinate 100% of the population after 12 weeks, despite having the capacity to do so.

The bottleneck in vaccination did not occur at the same time in all communities. For purpose of comparison here and moving forward, we divide communities into those that are predominantly White (>80% White residents; 61% of municipalities and 18% of Boston ZIP codes), have high Black-Latinx populations (>20% Black and Latinx residents; 9% of municipalities and 61% of Boston ZIP codes), and those that are neither (30% of municipalities and 21% of Boston ZIP codes). Predominantly White communities hit the bottleneck at the very end of November (*mean* = 57th day). Meanwhile, communities with high Black-Latinx populations hit the same milestone before November 15th (*mean* = 42nd day).

Hitting the bottleneck in vaccination earlier had a sharply negative impact on high Black-Latinx communities, as is apparent from the maps in Figure 4. First, and more obviously, it resulted in vaccination slowing at an earlier time point. Second, and more subtle, it means that when this time was reached there was a lower proportion of vaccinated residents, and thus fewer individuals who could persuade their neighbors. Consequently, fewer additional people were persuaded each week thereafter in these communities than in their predominantly White counterparts. This further exacerbated disparities in cumulative vaccinations. By the end of the simulation, residents in predominantly White communities were consistently 80% vaccinated whereas those living in high Black-Latinx communities were 71% vaccinated—but with some communities seeing rates of vaccination as low as 59%.

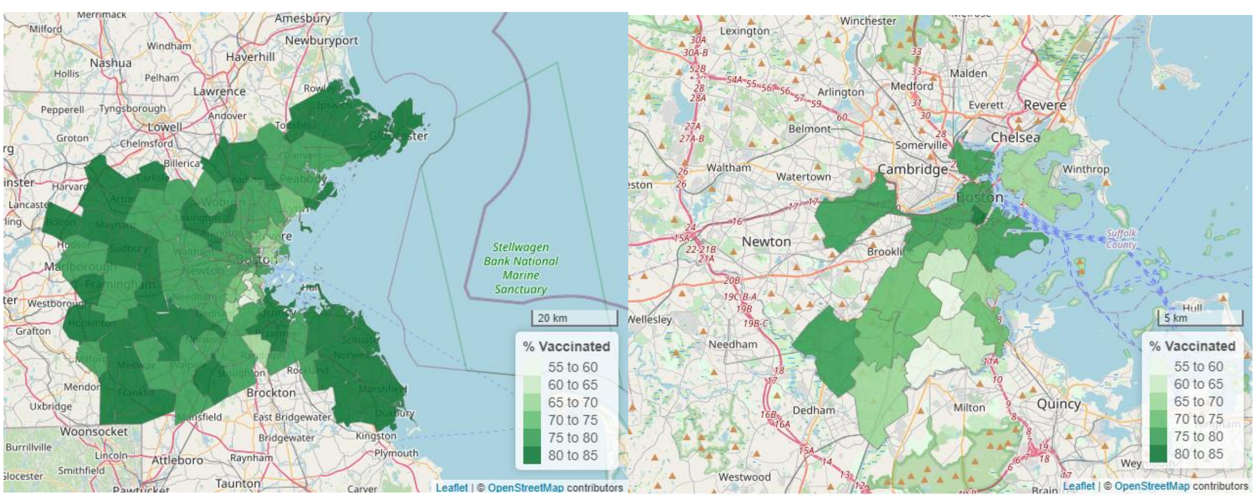

Figure 4. Estimated percentage of residents vaccinated at the end of the three-month simulation, by municipality in the region (left panel) and ZIP codes within Boston (right panel).

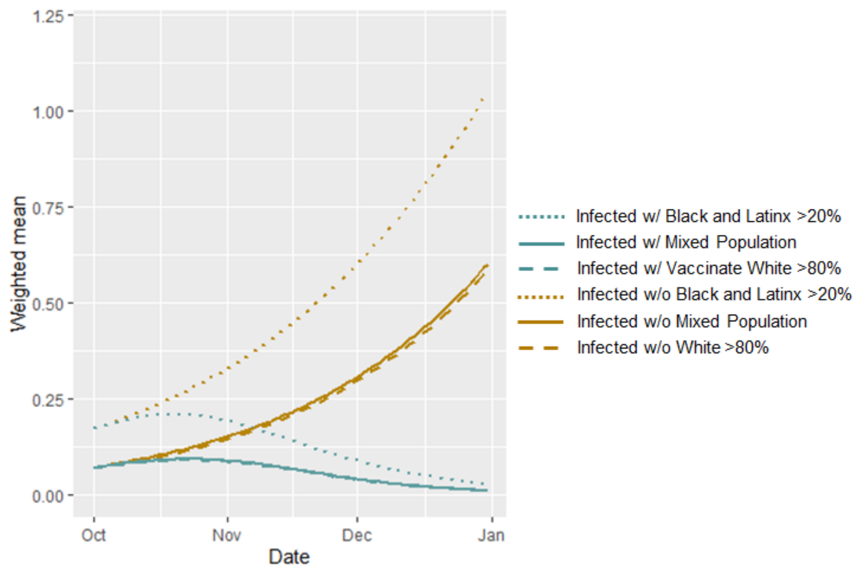

Figure 5. Evolution of infections over the course of the three-month simulation without vaccination (gold line) and with vaccination (blue line), broken out by ethnic composition of the community.

4. Attaining Herd Immunity

The impacts of vaccination, including disparities in uptake across communities, were evident in the corresponding evolution of infection rates. As shown in Figure 5, infection rates kept pace with the no-vaccination scenario until mid-October, which is when vaccination began to substantially lower the population susceptible to infection. The growth curve for infection rates under vaccination had about the same shape across communities of different racial composition, with all communities seeing substantial drops in infections relative to no vaccination. However, infections were still three times greater in high Black-Latinx communities than predominantly White communities (0.09 infections vs. 0.27 infections per 1,000 residents). These disparities were apparent when mapped across communities in Figure 6, largely matching the distribution of vaccination.

Moving beyond infection rates, did communities reach the goal of herd immunity via vaccination? We defined herd immunity as the point at which a community had 0 infections (as opposed to the proxy of what percentage of people have immunity, in which case an entrenched virus could still cause infections). We found that only 27% of communities had reached herd immunity by this definition at the end of the simulation. We thus extended the simulation for three additional months. An additional 52% of communities achieved herd immunity in the fourth month of the extended simulation (80% cumulative), and all but one remaining community achieved herd immunity in the fifth month, which reached it shortly thereafter. The average community reached herd immunity on day 103 of the simulation.

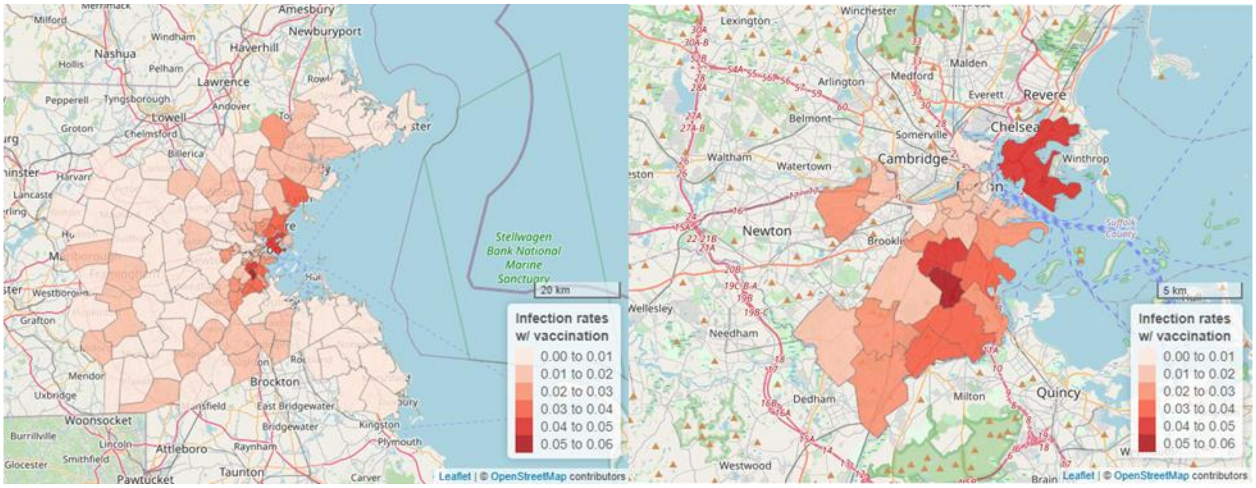

Figure 6. Estimated infection rates across municipalities in the region (left panel) and ZIP codes in Boston (right panel) at the end of the simulation.

Differences in achieving herd immunity again reflected stark disparities by race, as can also be seen in Figure 7. Only 14% of communities with high Black-Latinx populations residents saw herd immunity before the fifth month of the simulation, whereas 100% of predominantly White communities had achieved herd immunity by this time. To reiterate, *nearly all* predominantly White communities achieved herd immunity before *any* community with a high Black-Latinx population did so, in some cases by nearly two months. The average difference in achieving herd immunity between these two sets of communities was 44 days (89 days and 134 days into the simulation).

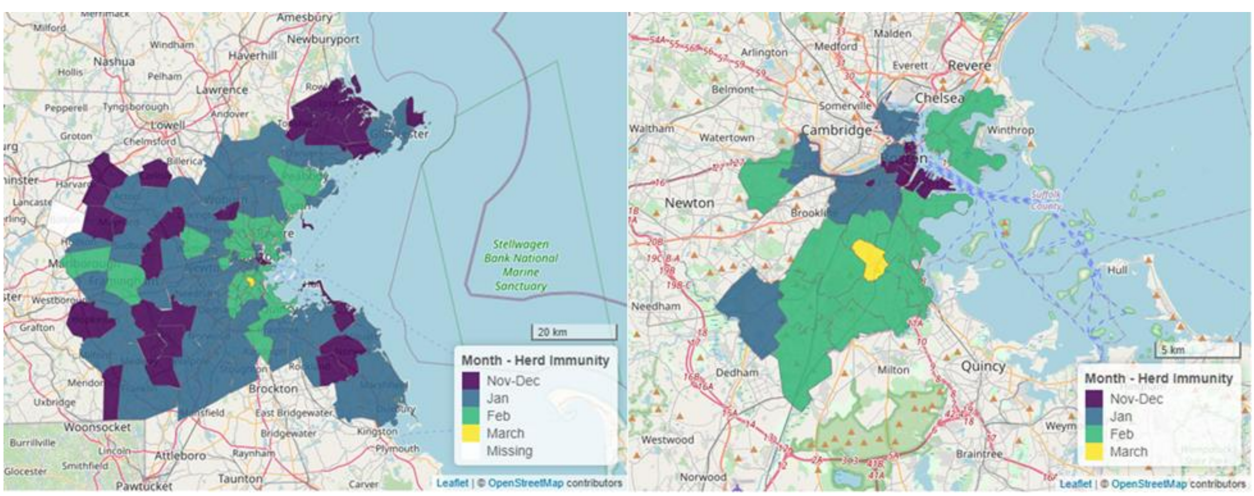

Figure 7. The month at which each municipality in the region (left panel) and each ZIP code in Boston (right panel) reached herd immunity in the simulation, defined as having zero infections.

5. What Might Happen Differently

A simulation is only as strong as its assumptions, and we have been forced to make many. To evaluate the robustness of our conclusions, we tested a variety of alterations to our model, including higher and lower levels of persuasion, as well as no persuasion; two lower levels of vaccine efficacy; and two longer timelines for vaccination roll-out. There were 36 possible combinations of these parameters, giving 36 different sets of results. Most of these alterations operated as would be expected, either extending or compressing vaccination time, or increasing or decreasing total infections. The key takeaway, though, was that disparities in infection rates and herd immunity were highly consistent across models.

Amidst these various alternative models, it is worth looking closely at the impact of increasing the rate of persuasion, being that community outreach is the focus of many in the public health community right now. We found that a substantial increase to the rate of persuasion only had marginal benefits across communities. Those with a high Black-Latinx population saw herd immunity only 10 days earlier than in the baseline model, while predominantly White communities saw an even smaller improvement of 2 days. This closed the gap between the two groups by less than 20%.

The failure of a stronger persuasion rate to substantially improve matters indicates that the proportion of people committed to being vaccinated at the outset and, conversely, the proportion committed to not being vaccinated, are highly consequential in determining the timeline of reaching herd immunity. The lessons here are two-fold. First, public messaging and outreach should be immediately amplified as much as possible in order to increase the percentage of those willing to receive the vaccine once it is their turn. Second, this messaging and outreach must reject the assumption that those who say they are “definitely not” going to get the vaccine cannot be swayed. Vaccine hesitancy and vaccine denial are different things. Hopefully, some of those who are currently rejecting the idea of being vaccinated now are strongly hesitant rather than deniers. Leaders may be able to reach these individuals with well-crafted, compassionate communication. Otherwise, we resign ourselves to the fact that 1 in 5 Black and Latinx residents in greater Boston would never be vaccinated. Even though we see widespread herd immunity in the model, in terms of communities having zero infections, this does not guarantee that infections cannot return. The middling rates of vaccination that appear possible for communities of color could still produce gaps in the protective wall of community health.

6. Conclusion

Greater Boston as a whole, especially its public health leaders and politicians, should take three main points away from this report. First, from where we currently stand, the vaccination process will eventually lead to the elimination of infections in all communities, *but* there will be inequities in getting there. In the end, though we see all communities eventually reach zero infections, the moderate levels of vaccination in some communities of color could create gaps for long-term herd immunity. This is neither a doomsday scenario, nor is it what any mayor or governor would want for their city or region. Second, we note that vaccination in all communities appears to hit a crucial bottleneck, at which point the rollout becomes dependent entirely on persuasion and falls behind the intended pace. Policymakers and practitioners should be attuned to any signals that this point is arriving and to ramp up community messaging accordingly. If not managed properly this bottleneck can lead to the misdistribution or even wasted vaccines, as well as many more days of illness and lives lost.

Third and final, the disparities between predominantly White communities and high Black-Latinx communities are best addressed through amplified messaging, outreach, and other effective mechanisms of persuasion. These efforts should not, however, assume that all those that say they will not receive the vaccination will maintain that position. Effective messaging is required to avoid continued vulnerability to vulnerability to the virus in communities of color.

Appendix A. NSF Beacon Survey Methodology

The NSF-Beacon survey is a collaboration of the Boston Area Research Initiative (BARI) at Northeastern University, the Center for Survey Research (CSR) at University of Massachusetts Boston, and the Boston Public Health Commission (BPHC), funded by the National Science Foundation's Human-Environment and Geographical Sciences (HEGS) program through a grant for rapid-response research (RAPID) for collecting ephemeral data during or following a crisis. The survey captures the experiences of 1370 Bostonians during the first months of the COVID-19 pandemic, including ability and tendency to follow social distancing recommendations, attitudes towards regulations, and economic and personal impacts of the pandemic. The design allows for a unique observation of neighborhood-level estimates for these factors. A follow-up survey was completed by 932 of the original respondents.

I. Sample Design and Final Sample

The NSF-Beacon survey used a stratified random sample that divided the city of Boston into 25 distinct neighborhoods. The neighborhoods were defined in collaboration with members of the Mayor's Office and other experts based on social, demographic, and historical salience. They were constructed to conform to census block group boundaries, meaning that metrics associated with census geographies (including from the U.S. Census Bureau) could be linked with the data. The Marketing Systems Group (MSG) was contracted to draw a simple random sample of residential addresses from within each neighborhood. They used the most recent United States Postal Service Computerized Delivery Sequence File (CDSF) to draw Address-Based Samples (ABS) of residential addresses. Four neighborhoods with a higher proportion of Black or Latinx populations were oversampled (Hyde Park, Mattapan, Lower Roxbury, and East Boston-Eagle Hill). As shown in Table 1, there were unbalanced sample sizes and selection probabilities across neighborhoods, meaning analysis of the data requires survey weights to correct for these differences. In addition to the survey being administered to the sample obtained for the NSF-Beacon study, we also invited participants in the previously-constructed Beacon panel, which had been recruited using the same 25 neighborhood stratified sample design.

II. Data Collection Methodology

Paper copies of the survey, plus instructions for completing and returning, and a $2 cash incentive were mailed to all sampled addresses. For three neighborhoods known to have higher percentages of Hispanic households, the materials mailed, including the survey instrument, were in both English and Spanish. All recipients were also given the option of completing the survey online and an associated URL. A randomly assigned half of the mailed questionnaires had instructions for the oldest adult 18+ in the household to complete the survey while the other random half had instructions for the youngest adult 18+ to complete the survey. In this manner, an attempt was made to randomize the age of the respondent within the household completing the survey. Approximately two weeks after the initial mailing of materials, a second mailing was sent to nonrespondents, though with no additional incentive. For the follow-up survey, those who shared e-mails were invited to complete an online survey.

Table 1. NSF-Survey neighborhood sampling specifications

Neighborhood	# of Sampled Addresses	Prob. of Selection	# of Completed Surveys	Response Rate[1]
Allston	192	0.01702	51	28.81%
Back Bay	194	0.01871	53	31.36
Beacon Hill	204	0.03593	53	30.11
Brighton	187	0.00839	58	31.87
Central	198	0.06119	50	27.78
Central Northeast	196	0.02839	58	33.14
Central West	200	0.01665	55	32.35
Charlestown	190	0.02286	62	34.25
Dorchester Central	189	0.01042	39	21.08
Dorchester North	188	0.02661	42	23.86
Dorchester South	191	0.01671	60	32.97
East Boston	189	0.02501	43	24.29
East Boston-Eagle Hill	355	0.04189	93	27.84
Fenway/Kenmore	195	0.01169	39	21.91
Hyde Park	364	0.02967	59	17.10
Jamaica Plain	188	0.01138	71	39.66
Jamaica Plain-Mission Hill	191	0.02737	55	30.73
Lower Roxbury	372	0.05977	57	17.59
Mattapan	362	0.02704	61	17.58
Roslindale	188	0.01820	73	40.11
Roxbury	188	0.01511	37	20.67
Seaport	192	0.04554	40	22.47
South Boston	191	0.01150	45	24.86
South End	188	0.01070	57	32.02
West Roxbury	189	0.01407	59	32.24
Total	**5481**		**1370**	**26.88%**

[1] Response rates computed using AAPOR Method 3.

III. Data Collection Results

The final sample included 1370 completed surveys (1208 paper, 162 online; 30 were completed in Spanish). The number of completed surveys ranged from 37 in Roxbury to 93 in East Boston-Eagle Hill. Overall response rate was 26.88% and ranged from a low of 17.10% in Hyde Park to a high of 40.11% in Roslindale. Full details on each neighborhood sample are presented in Table 1. An additional 256 completed surveys were obtained from members of the previously-constructed Beacon panel, bringing the total number of completed surveys to 1626. The follow-up survey had 932 respondents.

IV. Weighting of survey data

The sample requires weighting to account for both differing probabilities of selection and response rates across neighborhoods, especially insofar as these differences create a sample that is demographically and geographically non-representative. We created two survey weights, one for sample design factors including probability of selection and number of adults in the household adjusted for nonresponse bias across neighborhoods, the other which adds a post-stratified weight to account for demographic non-representativeness. Additionally, we conducted this process twice. First, we did it only for respondents to the NSF-Beacon survey. Second, we replicated the procedures for the dataset that combined the NSF-Beacon survey responses with respondents from the previously-constructed Beacon panel (values reported in Table 2 for weighting are highly similar for the NSF-Beacon responses alone and the merged data set).

Weights for Nonresponse Bias

Weighting for nonresponse began by neighborhood with the inverse of the probabilities of selection adjusted for the response rates displayed by neighborhood according to the equation (see Table 1 for values):

W_b = (Inverse of probability of selection) / (neighborhood response rate)

The final nonresponse adjusted weight further multiplies the base weight by the number of adults 18+ in the household (capped at 4 to prevent excessively large weights). Finally, these weights are adjusted so that the percentage of the total 18+ population in Boston that belongs in each neighborhood agreed with control percentages computed from the 2014-2018 5-year American Community Survey (ACS) data from the Census Bureau. These weights sum to the ACS estimate of the total 18+ population in the city of Boston. Therefore, the final nonresponse adjusted weight can be defined as:

W_{NR} = (W_b)(number of adults in household)(ACS population adjustment factor)

Post-Stratified Weights

As shown in Table 2, even after nonresponse weights, the respondents to the survey were not demographically representative of Boston's population. Most notably, people with education beyond 4-year college degrees were overrepresented and those with a high school education or less were underrepresented. Women were also overrepresented relative to men and White non-Hispanics were overrepresented relative to Blacks and Hispanics. There was also a smaller age bias with too many 65+ people and too few 18-34. A final adjustment to the survey weights was implemented to adjust for differential survey nonresponse by age, gender, race/Hispanic origin, and education. Control percentages for these categories were computed from the 2014-2018 5-year ACS data. Post-stratification factors were then computed to match weighted survey data to citywide percentages. The final post-stratified weight can be expressed as:

$W_{PS} = (W_{NR})(\text{post-stratified factors})$

It should be noted, though, that a small amount of trimming of weights, less than one percent of all sample cases, was employed to prevent some extreme values in the post-stratified weights. As shown in Table 2, this additional adjustment process brought the weighted survey estimates much more in line with ACS citywide estimates.

Weights for the second mail and web-based survey.

For the follow-up survey, where 932 of the original 1626 respondents answered questions, new post-stratification factors were developed to again match weighted survey data to the 2014-2018 5-year ACS.

Table 2. Comparison of ACS controls to nonresponse and post-stratified weights

	ACS	Nonresponse	Post-stratified
Age			
18-34	46.90%	38.40%	46.20%
35-49	21.3	20.1	21.5
50-64	18.4	22.1	18.6
65+	13.4	19.4	13.7
Gender			
Male	47.60%	38.00%	47.60%
Female	52.4	62	52.4
Education			
High School including GED or less	33.60%	16.40%	32.50%
Some college including 2-year degree	17.8	14.8	18
4-year college degree	26.5	29.3	27
Beyond 4-year college degree	22.1	39.5	22.5
Race/Hispanic origin			
White non-Hispanic	49.40%	57.50%	49.40%
Black non-Hispanic	20.6	15.8	20.6
Hispanic	16.9	12.4	16.9
Other	13.1	14.3	13.1

Appendix B. Simulation Model Construction, Inputs, and Robustness Checks

The study centers on a SIR model that uses established transmission and recovery rates in conjunction with mobility between communities to estimate daily infection rates in each community. The model was run for October-December, 2020, and the transmission and recovery rates were based on actual infection records. Mobility was also derived from historical data for the same time period. The "communities" act as nodes in the mobility model and are defined as the 100 non-Boston cities in towns in the greater Boston region (following the Metropolitan Area Planning Council's definition) and the 28 ZIP codes within Boston. This decision was made based on the availability of more granular data for Boston, as described below, as well as the large amount of between-ZIP code demographic diversity within the city, which is lower or absent in many of the surrounding municipalities.

Data and Measures

The models use four data sources: (1) population descriptors from the American Community Survey's 2014-2018 five-year estimates; (2) daily and weekly infection case counts, derived from infection records, for all towns in greater Boston and ZIP codes within Boston; (3) responses to three surveys including items on people's intentions regarding vaccination; (4) cross-community mobility records derived from cell phone records, generated by SafeGraph. The Boston Area Research Initiative's Geographical Infrastructure[2] was used to join data describing each municipality or ZIP code.

Census Indicators

We drew population descriptors from the U.S. Census' American Community Survey's 2014-2018 estimates for all census block groups in Massachusetts. This level was selected as it is the largest census geography that nests cleanly within ZIP codes in Boston and within municipal boundaries. Community indicators included total population and ethnic composition (i.e., proportion Asian, proportion Black, proportion Latinx, proportion White). All measures were aggregated from census block groups to the municipal or ZIP code level using population-weighted means.

Infection Cases

[2] https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/W0V6FX

The Commonwealth of Massachusetts' Department of Health released weekly counts of new infections for all municipalities, starting on April 14th, 2020. It also released daily counts for counties. From these two data sources we created daily town measures by: tabulating the weekly sum of infected cases in a county; calculating the percentage of a county's cases attributed to each town; estimating the daily infected cases per town as the same percentage of the daily count for the county. For Boston ZIP codes, we had case records tracked by the Boston Public Health Commission mapped to the ZIP code of residence. We tabulated these for daily counts.

Surveys

Three surveys of Massachusetts residents were conducted that included a question regarding intention to vaccinate and split these responses by race. The three surveys were conducted by the Center for Survey Research at University of Massachusetts Boston with the Boston Area Research Initiative, MassInc Polling on behalf of the Boston Museum of Science, and Suffolk University Polling and the Boston Globe. The first surveyed residents of Boston from September to November and the other two surveyed residents from throughout Massachusetts in November and December, respectively. Although the exact wording varied between them, all three surveys permitted respondents to say that they "definitely" or "definitely did not" plan to get the vaccine or that they were uncertain or undecided. The overall proportion of individuals falling in each of these three groups was consistent across the three surveys, as were the breakdowns by race (though the MassInc poll did not have enough Asian respondents to include cross-tabs for that group; see Tables B1-B3).

We summed the cross-tabs for the vaccination intention question by race across the three surveys to calculate the weighted proportion of individuals indicating "as soon as possible," "never," and something in between for each of the four major racial categories--Asian, Black, Latinx, and White. We then estimated the proportion of residents in each of these three categories regarding the vaccine for each municipality and Boston ZIP code with the following equation:

$$Y_{i,k} = \sum_{j} p_{i,j} * r_{j,k}$$

where Yi,k is the proportion of residents in community i with attitude k toward the vaccine (e.g., getting it as soon as possible), $p_{i,j}$ is the proportion of residents in community i of race j, and $r_{j,k}$ is the proportion of members of race j giving k as their response across the three surveys.

Cellphone Generated Mobility Records

We used Safegraph's daily "Social Distancing" dataset to create the mobility network. The data are generated using a panel of GPS pings from anonymous mobile devices. Each device is attributed to an estimated home census block group (CBG) based on its most common nighttime location. It also tracks all stay points of these devices within other CBGs. The published data aggregate these pieces of information to generate a mobility matrix of the daily number of visits by the assumed residents of each CBG to each other CBG.

Mobility-Driven SIR model with the Distribution of Vaccination

Model

Our model was based on a traditional SIR (susceptible-infected-recovered) model that then incorporated two additional factors: mobility, to simulate the effect of contacts brought by the mobility between communities; and vaccination, to model the effect of the adoption of vaccination across communities. The full model consists of the following differential equations, which update daily:

$$
\begin{aligned}
\partial_t j_n &= \left(\alpha_n j_n + \gamma \sum\nolimits_{m\neq n} \alpha_n w_{mn} j_m \frac{N_m}{N_n} + \gamma \sum\nolimits_{n\neq m} \alpha_m w_{nm} j_m\right) s_n - \beta j_n \\
\partial_t s_n &= -\left(\alpha_n j_n + \gamma \sum\nolimits_{m\neq n} \alpha_n w_{mn} j_m \frac{N_m}{N_n} + \gamma \sum\nolimits_{n\neq m} \alpha_m w_{nm} j_m + \mu g(p_n)\right) s_n \\
\partial_t r_n &= \beta j_n \\
\partial_t p_n &= -g(p_n) + h u_n v_n \\
\partial_t u_n &= -h u_n v_n \\
\partial_t v_n &= g(p_n)
\end{aligned}
$$

where N_n, j_n, s_n, and r_n represent the total population size, number of infected cases, susceptible individuals, and recovered cases, respectively, for a community at a given timepoint.

The equations, which model simultaneous change over time in infected, susceptible, and recovered individuals, rely on four main components. α_n is the growth rate of infections in a community (i.e., the expected number of new cases from existing cases). It is based on a global α_0 in combination with a sigmoid function that accounts for fluctuation effects between communities when j_n is less than a threshold ε i.e. $\alpha_n = \alpha_0 \cdot \frac{\left(\frac{j_n}{\varepsilon}\right)^4}{1+\left(\frac{j_n}{\varepsilon}\right)^4}$.[3] The second component

[3] D. Brockmann, D. Helbing, The hidden geometry of complex, network-driven contagion phenomena. *Science* **342**, 1337-1342 (2013).

pertains to mobility, including two operands: $r_I = \gamma \sum_{m \neq n} \alpha_n w_{mn} j_m \frac{N_m}{N_n}$ calculates the possibility of an infected person in other communities (m) visiting the community and infecting susceptible in community n; and $r_R = \gamma \sum_{n \neq m} \alpha_m w_{nm} j_m$ is the possibility of a susceptible person from the community visiting another community (m) and becoming infected. $w_{mn}=F_{mn}/N_m$ where F_{mn} indicates the total number of visits from community m to community n, as captured by the cell-phone generated mobility data. The average mobility rate gamma is defined as $\gamma = \frac{\sum_{m \in G} F_m}{\sum_{m \in G} N_m}$. The third component is the rate of recovery of infected individuals, represented by β, which is consistent across communities. All individuals who have been infected and recovered are permanently removed from the susceptible population.

The fourth component of the model is vaccination adoption, wherein p_n is the proportion of people who will definitely receive the vaccine, u_n is the proportion of people uncertain about getting the vaccine, and v_n is the proportion of people who have already been vaccinated. μ reflects the effectiveness of the vaccine and $\mu * v_n$ are treated as part of r_n as they have been removed from the susceptible population. h is the persuasion rate at which someone uncertain about getting the vaccine will be persuaded to do so, whose strength is contingent on the proportion of residents in the community who have already been vaccinated (i.e., hu_nv_n). We assume that h is consistent across communities. Only those who were uncertain about the vaccine could be persuaded, not those who stated they would never get the vaccine. Actual vaccination is represented by the function $g(p_n)$, of the form:

$$g(p_n) = \begin{cases} c & p_n \geq c \\ p_n & p_n < c \end{cases}$$

in which c is the maximum capacity of the vaccination rollout for a day.

Fixed Parameters

Multiple parameters were established in advance of estimating the final model. α and β, the transmission and recovery rate, were estimated by running the simulation without vaccination on historical mobility and infection data for September 30th through December 22nd. Grid search identified a local optimum for α=0.096 and β=0.072, which translate to the more familiar R_0=α/β=1.33. ε was calculated as $\varepsilon = \frac{M}{\sum_{m \in G} N_m} = 3.826 \times 10^{-5}$, where M is the number of communities. μ, or the vaccine effectiveness, was set to 0.95, per the Pfizer and Moderna trial results. The rate of vaccine rollout was initially set to a 12-week (approx. 3-month) rollout period, meaning c=1/(7*12)=.0119, or 1.19% of the population could be expect to be vaccinated daily.

For *h*, we assume that the rate at which individuals uncertain about taking the vaccine are persuaded to do so is dependent on the number of people in their community who have received the vaccine. The number of people a person knows in their neighborhood and the number of people they need to know who have been vaccinated to be persuaded both vary by individual. In terms of likelihood to be persuaded, we use an additional item from the Mass Inc-Museum of Science survey regarding when people would be likely to get vaccinated. We note two groups: those who would like to see a few people get the vaccine before they do, and those who would like to see many other people get it. We use the cross-tabs from this question with the initial vaccination intention question to distribute them within the sample (see Supplementary Online Materials), and differentiate between these functionally in the model based on how many members of their neighborhood need to be vaccinated for them to be persuaded. Numerous neighborhood surveys have found that people say they are friends with or personally know very few of those living in their neighborhood (e.g., 5 people on one's street, less than 10 friends in the neighborhood; *37*). Based on these numbers, we estimate that the average individual knows approximately 1% of the residents of his or her neighborhood well enough to know and relate to their vaccination experiences. We further estimate that if about 25% of these people (i.e., 0.25% of the population) were vaccinated, the average person would know at least "a few" people who had been vaccinated. Approximately 37% of people who were uncertain about vaccination said they would do so once "a few people they knew" were, the remainder when "many people" were. From this, we estimate that when 25% of the residents of a neighborhood are vaccinated, 18.5% of those who started out as uncertain will have been persuaded (that is, half of 37%, as only the average person would have enough exposure at that time). Based on the same results, we believe that the asymptote for persuasion is at 90%. We solved for these established points in a sigmoid function and found $h = .026$.

Robustness Tests

We ran iterations of the model to test for robustness, varying three elements: persuasion, vaccine efficacy, and timeline of vaccine roll-out. We tested persuasion as 50% stronger ($h = .039$) and 50% weaker ($h = .013$), as well as the absence of persuasion altogether ($h = 0$). We tested vaccine efficacy at the lower points of 85% and 75%. We tested the timelines for vaccination roll-out at four and six months ($c = .0089$ and $.0056$, respectively). We re-ran the simulation with all possible combinations, making for 36 sets of results (4x3x3).

Table B1. Responses by race to the question "Will you get vaccinated?" in the Center for Survey Research-Boston Area Research Initiative survey.

	Asian	**Black**	**Latinx**	**White**	**Total**
Definitely	44 (.49)	21 (.25)	24 (.34)	350 (.56)	**439 (.51)**
Probably	34 (.38)	26 (.31)	27 (.39)	207 (.33)	**294 (.34)**
Probably Not	9 (.10)	21 (.25)	11 (.16)	50 (.08)	**91 (.11)**
Definitely Not	2 (.02)	16 (.19)	8 (.11)	14 (.02)	**40 (.05)**
Total	**89**	**84**	**70**	**621**	**864**

Table B2. Responses by race to the question "When an FDA-approved vaccine for COVID is made available, how likely will you be to take it?" in the Mass Inc.-Museum of Science survey.

	Black	**Latinx**	**White**	**Total**
Very likely	27 (.33)	36 (.31)	404 (.46)	**467 (.44)**
Somewhat likely	23 (.28)	28 (.24)	246 (.28)	**297 (.28)**
Not too likely	12 (.15)	17 (.15)	88 (.10)	**117 (.11)**
Not at all likely	15 (.18)	23 (.20)	70 (.08)	**108 (.10)**
Unsure	4 (.06)	11 (.09)	70 (.07)	**85 (.08)**
Total	**81**	**115**	**878**	**1073**

Table B3. Responses by race to the item "When a federally-approved COVID-19 vaccine is available to you, will you…" in the Suffolk-Boston Globe survey.

	Asian	**Black**	**Latinx**	**White**	**Total**
Take it as soon as you can?	17 (.71)	5 (.11)	16 (.32)	219 (.59)	**261 (.52)**
Wait awhile until others have taken it?	7 (.29)	24 (.55)	21 (.42)	107 (.29)	**164 (.33)**
Not take the vaccine?	0 (.00)	13 (.30)	13 (.26)	35 (.09)	**62 (.12)**
Undecided	0 (.00)	2 (.05)	0 (.00)	10 (.03)	**12 (.02)**
Total	**24**	**44**	**50**	**372**	**500**

Table B4. Cross-tabs between the items "When an FDA-approved vaccine for COVID is made available, how likely will you be to take it?" and "When an FDA-approved vaccine for COVID is made available, when do you think you will be most likely to take it?" in the Mass Inc.-Museum of Science survey.in the Suffolk-Boston Globe survey.

	Very Likely	*Somewhat likely*	*Not too likely*	*Not at all likely*
As soon as possible	.72	.13	.03	.01
After a few people I know have taken it	.18	.35	.10	.01
After many other people have taken it	.08	.45	,70	.20
Never	.00	.00	.05	.59
Unsure	.02	.07	.12	.19

Table B5. Comparisons of disparities across all simulations for herd immunity, including percentage of predominantly White and high Black-Latinx communities that reach herd immunity and the average date of reaching herd immunity.

	No Persuasion (κ = 0)	**Low Persuasion (κ = .013)**	**Medium Persuasion (κ = .026)**	**High Persuasion (κ = .039)**
Distribution Time = 3 months				
Vaccine Efficacy = 75%	34% vs. 0% 139 days vs. NA	100% vs. 18% 129 days vs. 178 days	100% vs. 77% 115 days vs. 163 days	100% vs. 100% 109 days vs. 155 days
Vaccine Efficacy = 85%	80% vs. 0% 138 days vs. NA	100% vs. 64% 111 days vs. 164 days	100% vs. 100% 101 days vs. 149 days	100% vs. 100% 97 days vs. 138 days
Vaccine Efficacy = 95%	97% vs. 0% 121 days vs. NA	100% vs. 86% 99 days vs. 153 days	100% vs. 100% 91 days vs. 134 days	100% vs. 100% 88 days vs. 125 days
Distribution Time = 4 months				
Vaccine Efficacy = 75%	28% vs. 0% 152 days vs. NA	97% vs. 0% 142 days vs. NA	100% vs. 59% 131 days vs. 171 days	100% vs. 86% 127 days vs. 166 days
Vaccine Efficacy = 85%	57% vs. 0% 141 days vs. NA	100% vs. 50% 125 days vs. 171 days	100% vs. 100% 116 days vs. 159 days	100% vs. 100% 114 days vs. 150 days
Vaccine Efficacy = 95%	95% vs. 0% 135 days vs. NA	100% vs. 82% 112 days vs. 160 days	100% vs. 100% 106 days vs. 144 days	100% vs. 100% 104 days vs. 137 days
Distribution Time = 6 months				
Vaccine Efficacy = 75%	5% vs. 0% 169 days vs. NA	45% vs. 0% 161 days vs. NA	60% vs. 0% 162 days vs. NA	65% vs. 0% 163 days vs. NA
Vaccine Efficacy = 85%	28% vs. 0% 158 days vs. NA	88% vs. 0% 157 days vs. NA	95% vs. 18% 155 days vs. 179 days	95% vs. 23% 155 days vs. 177 days
Vaccine Efficacy = 95%	54% vs. 0% 150 days vs. NA	98% vs. 32% 145 days vs. 178 days	100% vs. 91% 144 days vs. 173 days	100% vs. 95% 144 days vs. 171 days

Table B6. Comparisons of disparities across all simulations for infection rates, including the regression parameters for percentage Black and percentage Latinx, with and without controlling for infection rates under the no-vaccine scenario.

	No Persuasion (κ = 0)	**Low Persuasion (κ = .013)**	**Medium Persuasion (κ = .026)**	**High Persuasion (κ = .039)**
Distribution Time = 3 months				
Vaccine Efficacy = 75%	% Black = 1.44 % Latinx = 2.63 % Black (cont. inf.) = 1.25 % Latinx (cont. inf.) = 1.22	% Black = 0.90 % Latinx = 1.70 % Black (cont. inf.) = 0.77 % Latinx (cont. inf.) = 0.76	% Black = 0.54 % Latinx = 1.18 % Black (cont. inf.) = 0.45 % Latinx (cont. inf.) = 0.47	% Black = 0.33 % Latinx = 0.91 % Black (cont. inf.) = 0.25 % Latinx (cont. inf.) = 0.30
Vaccine Efficacy = 85%	% Black = 1.21 % Latinx = 2.00 % Black (cont. inf.) = 1.08 % Latinx (cont. inf.) = 1.01	% Black = 0.72 % Latinx = 1.21 % Black (cont. inf.) = 0.63 % Latinx (cont. inf.) = 0.60	% Black = 0.41 % Latinx = 0.80 % Black (cont. inf.) = 0.35 % Latinx (cont. inf.) = 0.35	% Black = 0.24 % Latinx = 0.60 % Black (cont. inf.) = 0.19 % Latinx (cont. inf.) = 0.22
Vaccine Efficacy = 95%	% Black = 1.01 % Latinx = 1.52 % Black (cont. inf.) = 0.92 % Latinx (cont. inf.) = 0.83	% Black = 0.56 % Latinx = 0.87 % Black (cont. inf.) = 0.51 % Latinx (cont. inf.) = 0.46	% Black = 0.30 % Latinx = 0.54 % Black (cont. inf.) = 0.26 % Latinx (cont. inf.) = 0.26	% Black = 0.17 % Latinx = 0.39 % Black (cont. inf.) = 0.14 % Latinx (cont. inf.) = 0.16
Distribution Time = 4 months				
Vaccine Efficacy = 75%	% Black = 1.43 % Latinx = 2.99 % Black (cont. inf.) = 1.19 % Latinx (cont. inf.) = 1.18	% Black = 0.90 % Latinx = 2.09 % Black (cont. inf.) = 0.71 % Latinx (cont. inf.) = 0.71	% Black = 0.51 % Latinx = 1.60 % Black (cont. inf.) = 0.34 % Latinx (cont. inf.) = 0.38	% Black = 0.28 % Latinx = 1.39 % Black (cont. inf.) = 0.12 % Latinx (cont. inf.) = 0.19
Vaccine Efficacy = 85%	% Black = 1.23 % Latinx = 2.32 % Black (cont. inf.) = 1.05 % Latinx (cont. inf.) = 1.00	% Black = 0.73 % Latinx = 1.55 % Black (cont. inf.) = 0.60 % Latinx (cont. inf.) = 0.58	% Black = 0.39 % Latinx = 1.14 % Black (cont. inf.) = 0.28 % Latinx (cont. inf.) = 0.30	% Black = 0.20 % Latinx = 0.97 % Black (cont. inf.) = 0.09 % Latinx (cont. inf.) = 0.15

Vaccine Efficacy = 95%	% Black = 1.04 % Latinx = 1.79 % Black (cont. inf.) = 0.91 % Latinx (cont. inf.) = 0.84	% Black = 0.59 % Latinx = 1.14 % Black (cont. inf.) = 0.50 % Latinx (cont. inf.) = 0.47	% Black = 0.30 % Latinx = 0.81 % Black (cont. inf.) = 0.23 % Latinx (cont. inf.) = 0.23	% Black = 0.15 % Latinx = 0.68 % Black (cont. inf.) = 0.07 % Latinx (cont. inf.) = 0.11
Distribution Time = 6 months				
Vaccine Efficacy = 75%	% Black = 1.20 % Latinx = 4.00 % Black (cont. inf.) = 0.75 % Latinx (cont. inf.) = 0.64	% Black = 0.62 % Latinx = 3.61 % Black (cont. inf.) = 0.16 % Latinx (cont. inf.) = 0.23	% Black = 0.43 % Latinx = 3.63 % Black (cont. inf.) = -0.03 % Latinx (cont. inf.) = 0.23	% Black = 0.43 % Latinx = 3.63 % Black (cont. inf.) = -0.03 % Latinx (cont. inf.) = 0.23
Vaccine Efficacy = 85%	% Black = 1.05 % Latinx = 3.25 % Black (cont. inf.) = 0.69 % Latinx (cont. inf.) = 0.58	% Black = 0.51 % Latinx = 2.90 % Black (cont. inf.) = 0.15 % Latinx (cont. inf.) = 0.21	% Black = 0.34 % Latinx = 2.91 % Black (cont. inf.) = -0.03 % Latinx (cont. inf.) = 0.21	% Black = 0.34 % Latinx = 2.91 % Black (cont. inf.) = -0.03 % Latinx (cont. inf.) = 0.21
Vaccine Efficacy = 95%	% Black = 0.92 % Latinx = 2.64 % Black (cont. inf.) = 0.63 % Latinx (cont. inf.) = 0.52	% Black = 0.42 % Latinx = 2.32 % Black (cont. inf.) = 0.13 % Latinx (cont. inf.) = 0.18	% Black = 0.27 % Latinx = 2.33 % Black (cont. inf.) = -0.02 % Latinx (cont. inf.) = 0.18	% Black = 0.27 % Latinx = 2.33 % Black (cont. inf.) = -0.02 % Latinx (cont. inf.) = 0.18