

Rationing safe and effective Covid-19 vaccines: allocating to states proportionate to population may undermine commitments to mitigating health disparities

Harald Schmidt,^{1*} Parag A. Pathak,^{2*} Michelle A. Williams,³ Tayfun Sönmez,⁴ M. Utku Ünver,^{4#} Lawrence O. Gostin.^{5#}

¹Department of Medical Ethics and Health Policy, Research Associate, Center for Health Incentives and Behavioral Economics, Senior Fellow, Leonard Davis Institute of Health Economics, University of Pennsylvania

²Class of 1922 Professor, Department of Economics, Massachusetts Institute of Technology, Cambridge, MA

³Harvard T.H. Chan School of Public Health, Harvard University, Boston, Massachusetts

⁴Department of Economics, Boston College, Chestnut Hill, MA

⁵O'Neill Institute for National & Global Health Law, Georgetown University, Washington, DC

*Co-first-authors, #Co-senior-authors

Introduction

A central goal in the National Academies of Science, Engineering and Medicine's (NASEM) framework for equitable COVID-19 vaccine allocation is to mitigate existing inequities, particularly those affecting economically worse-off racial and ethnic minorities.^{1,2,3} The Advisory Committee on Immunization Practice (ACIP) likewise notes that equity demands to "reduce, rather than increase, health disparities in each phase of vaccine distribution". A crucial question in this regard is how vaccines should be distributed to states. The default is to allocate proportionate to population size. However, this approach risks increasing scarcity for worse-off populations in states where they represent above-average shares. To avoid lower odds of receiving a vaccine for worse-off groups, more vaccines could be given to states with larger shares of worse-off populations, and fewer to ones with smaller shares. We show here the consequences of allocating by these two different approaches.

Methods

We simulated the NASEM allocation framework, which recommends a 10% reserve for the most-deprived population quartile, adapting our earlier study design (see Appendix).^{4,5} We compared the consequences for worse-off populations if states receive vaccines proportionate to population:

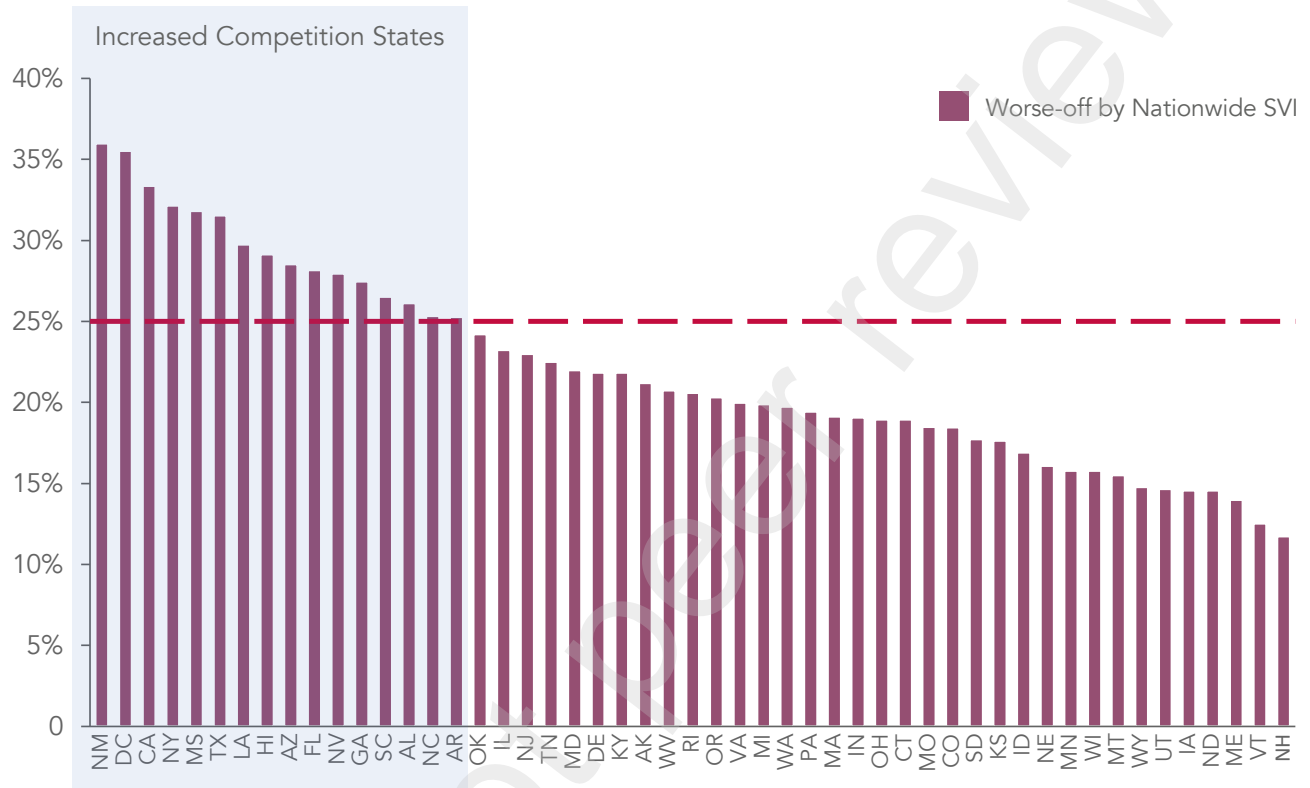
1. without any adjustments;
2. setting aside a reserve of 10% of a state's share as additional allotment for its most-deprived quartile; and
3. setting aside a reserve of 10% of the overall nationally available vaccines to allocate it for the most-deprived quartile within the nation (without additional efforts at the state level).

For the purpose of this illustration, we follow NASEM and CDC's preferences and use the CDC's Social Vulnerability Index (SVI) to determine the most-deprived quartiles;⁶ accordingly "worse-off" refers to the most deprived quartile on the nationwide SVI, and "better-off" to the remainder population. Consistent with NASEM and key analyses,^{1,7} we used the American Community Survey (ACS) 2014- 2018 5-year Public Use Microdata Sample (PUMS) 2018 Behavioral Risk Factor Surveillance System (BRFSS) to capture key population groups.

Results

Figure 1 shows what share of each state's population falls into the nation's worse-off quartile, varying from 36% (NM) to 12% (NH). In 16 'Increased Competition' states, the worse-off group accounts for more than 25% of its population.

Figure 1: Share of each state's population falling under the worse-off quartile nationally (SVI)

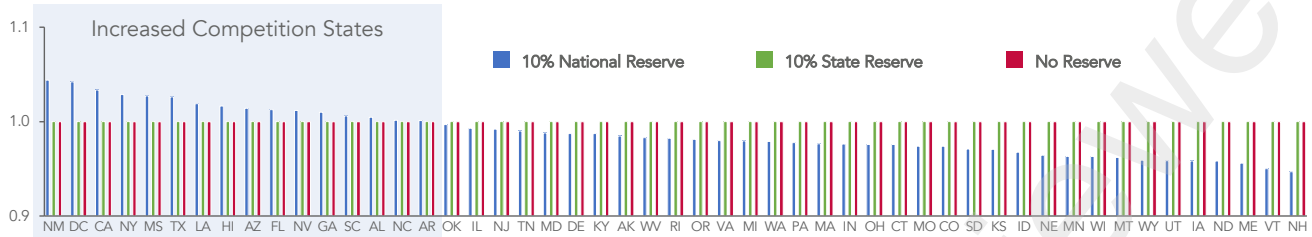


Notes: Using individualized SVI estimation approach based on the ACS PUMS data set. CDC's census-tract approach and our approach has correlation 0.924 for worse-off population shares of states.
 Sources: National Academies of Sciences, Engineering, and Medicine Framework for Equitable Allocation of Vaccine for the Novel Coronavirus and American Community Survey (ACS) 5-Year.

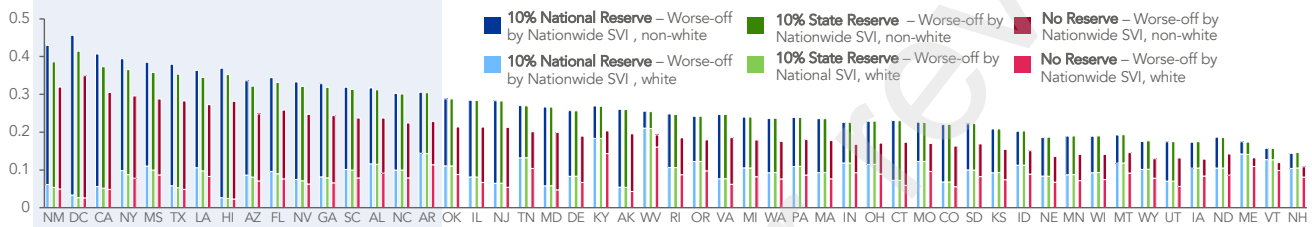
Figure 2 shows the allocation of the first 161 million vaccines under NASEM's Phase 1&2. Panel 2a shows that total allocation to states would be the same under no adjustments and a 10% state reserve (as both are proportionate to population). But under the 10% national reserve, a ratio above 1 (e.g., NM) indicates that a state obtains more units relative to population, while a ratio below 1 (e.g., NH) indicates it receives fewer units. Panel 2b reports what the three allocation methods mean for worse-off groups. Compared to no adjustments, reserving 10% at either the national or state-level benefits worse-off populations. Under either approach, the benefit is almost equal for states with <25% worse-off populations, but larger using the 10% national reserve for states with >25%. Panel 2c shows how many more doses would need to be allocated to ensure these gains (with larger states requiring higher adjustments due to their size).

Figure 2: Cumulative doses received by state by allocation scenario, 161 million doses allocated nationally

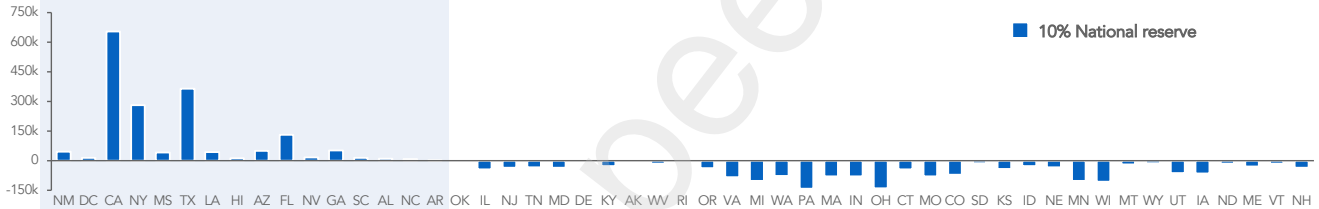
2a: Total population offered (ratio of share relative to population proportion of state)



2b: Worse-off population offered (ratio of share relative to population proportion of state)



2c: Gains and losses for each state on National Reserve relative to No Reserve/State Reserve scenarios (number of vaccines)



Sources: National Academies of Sciences, Engineering, and Medicine Framework for Equitable Allocation of Vaccine for the Novel Coronavirus and American Community Survey (ACS) 5-Year.

Discussion

The population of worse-off groups is not distributed evenly across states. If vaccines are allocated to states proportionate to population alone, worse-off groups in 16 states face greater scarcity than those in the remaining states. How the NASEM's recommendation to reserve 10% of vaccines for worse off groups is implemented, and whether it is sufficient, therefore requires close attention. To promote equity and to avoid increasing competition for scarce vaccines among worse off groups, specifying a national reserve using SVI with consideration of states' shares of worse-off populations is preferable over adjusting by SVI within states.

Acknowledgements

We thank Chetan Patel, Edwin Song, and Adam Solomon for outstanding research assistance.

Authorship note

HS wrote the first draft of the manuscript, all authors critically reviewed and edited. HS, MW and LG shaped the conceptualization of the study and visualization of data jointly with UU, TS and PP. UU, TS and PP lead all simulations.

¹ Gayle H, Foege W, Brown L, Kahn B, eds. *A Framework for Equitable Allocation of Vaccine for the Novel Coronavirus*. National Academy of Sciences, Engineering, and Medicine; August 2020. Accessed September 21, 2020. Url: <https://www.nationalacademies.org/our-work/a-framework-for-equitable-allocation-of-vaccine-for-the-novel-coronavirus>

² Pathak PA, Sönmez T, Ünver MU, Yenmez MB. Fair Allocation of Vaccines, Ventilators and Antiviral Treatments: Leaving No Ethical Value Behind in Health Care Rationing. August 2, 2020. Url: <https://arxiv.org/abs/2008.00374> [preprint]

³ Schmidt H. Vaccine rationing and the urgency of social justice in the COVID-19 response. *Hastings Cent Rep.* 2020;50(3):46-49.

⁴ Pathak PA, Schmidt H, Solomon A, Song E, Sönmez T, Ünver MU. Do Black and Indigenous Communities Receive their Fair Share of Vaccines Under the 2018 CDC Guidelines? *NBER Working Paper* No. 27817, September 2020, doi: 10.3386/w27817. Url: <https://www.nber.org/papers/w27817>

⁵ Schmidt H, Ünver MU, Williams MA, Pathak PA, Sönmez T, Gostin L. What Prioritizing Worse-Off Minority Groups for COVID-19 Vaccines Means Quantitatively: Practical, Legal and Ethical Implications. *SSRN Working Paper*, October 27 2020, url: <https://ssrn.com/abstract=3716686> or <http://dx.doi.org/10.2139/ssrn.3716686>

⁶ Dasgupta et al. CDC Covid-19 Response team. Association Between Social Vulnerability and a County's Risk for Becoming a COVID-19 Hotspot — United States, June 1–July 25, 2020. *MMWR*, October 22, 2020.

⁷ McCormack G, Avery C, Spitzer AK, Chandra A. Economic Vulnerability of Households With Essential Workers. *JAMA*. 2020;324(4):388–390. doi:10.1001/jama.2020.11366

Appendix: Simulation Procedure, Data Sources and Supplementary Graphs

to

“Rationing safe and effective Covid-19 vaccines: allocating to states proportionate to population may undermine commitments to mitigating health disparities”¹

Harald Schmidt,* Parag A. Pathak,* Michelle A. Williams, Tayfun Sönmez, M. Utku Ünver,# Lawrence Gostin.#

*Co-first-authors, #Co-senior-authors

Contents

A1. Overview 2

 A1.1. Data Sources 2

 A1.2. Reserve Allocation Procedure 3

 A1.2.1. The State Reserve Allocation Procedure 3

 A1.2.2. The Federal Reserve Allocation Procedure 4

 A1.3. Simulation Methodology 5

A2. Phase Inclusion Criteria 6

 A2.1. NASEM phases 6

 A2.2. Data issues and imperfect workarounds 6

 A2.2.1. High-risk conditions and the BRFSS 6

 A2.3. Comment on phase sizes 8

 A2.4. Phase by phase details for labelling of PUMS data 8

 A2.4.1. Phase 1a 8

 A2.4.2. Phase 1b 9

 A2.4.3. Phase 2 10

 A2.4.4. Phase 3 12

 A2.4.5. Phase 4 12

A3. Components and Calculation of SVI 12

 A3.1. Overview 12

 A3.2. Computation details 13

 A3.3. Component by component details 14

A4. Worse-off and better-off computation 16

¹ The methodology of the study itself at several places follow the technical appendix of the working paper Schmidt et al. (2020) on COVID-19 vaccine distribution using state reserves and its national consequences to worse-off communities of different racial/ethnic backgrounds. The only differences in methodology are the omission of ADI, consideration of reserves allocation procedure on national SVI, and inclusion of SVI census-tract-level worse-off quartiles using directly CDC data in addition as a robustness check. Also see Pathak et al. (2019) for study design.

A5.	Racial-Ethnic Group Definitions.....	16
A6.	Table of Program Files	17
A7.	Allocation at 48 million Doses	18
A8.	Our SVI Approach vs CDC’s Census-Tract Level Approach.....	18

A1. Overview

A1.1. Data Sources

We base our phases on the COVID-19 vaccine allocation framework outlined in the [2020 NASEM guidance](#) (Gayle et al. 2020). The US population is partitioned into 5 subphases/phases (Phase 1(a), Phase 1(b), Phase 2, Phase 3, Phase 4) based on essential or high-risk occupations and demographic factors indicating risk for spreading/contracting severe disease.

The NASEM document contains approximate estimates of the number of people within each phase based on various sources. As these numbers are based on estimates from various sources and they are not available to us, we use the American Community Survey (ACS) 2014-2018 5-year Public Use Microdata Sample (PUMS) (U.S. Census Bureau 2014-2018). The 5-year PUMS contains two datasets, one containing housing unit characteristics for a sample of housing units and another containing individual characteristics for the individuals within those same housing units. We link the person data, which has over 15 million observations, with the household data for our labelling of phases.

The PUMS person data is a weighted sample. Every observation of an individual in the PUMS person data is associated with a weight called “person’s weight for generating statistics on individuals” (PWGTP). This weight is “used to bring the characteristics of the sample more into agreement with those of the full population by compensating for differences in sampling rates across areas, differences between the full sample and the interviewed sample, and differences between the sample and independent estimates of basic demographic characteristic estimates of population characteristics”, according to the ACS [Design and Methodology report](#) (U.S. Census Bureau 2014).

PWGTP was also scaled such that the size of any population group could be estimated by the sum of PWGTP across observations in the PUMS belonging to the group. Further details about calculation of weights in the ACS can be found in the [PUMS technical documentation](#) (U.S. Census Bureau 2020a).

The ACS PUMS does not include data on the presence of high-risk conditions, which impacts one’s phase priority. We supplement with 2018 data from the [Behavioral Risk Factor Surveillance System](#) (BRFSS) (CDC 2018b). As the observations in the BRFSS are unrelated to those in the PUMS, we rely on randomization to link the two datasets by computing the proportion with high-risk conditions for each characteristic demographic group in the BRFSS and assigning each observation in PUMS person data as high-risk using weighted coin flips. More details are in the “Phase Inclusion Criteria” section of this document.

In order to implement the NASEM guidelines, we need a methodology to identify critical workers who are at high risk of exposure to COVID-19 in the workplace. To do so, we utilize a list of occupations identified by the [U.S. Department of Homeland Security \(DHS\)](#) that are designated as Essential Critical Infrastructure Workers (DHS 2020). We then merge this list with occupation specific [Bureau of Labor Statistics \(BLS\) O*NET](#) survey results indicating how often an individual of a particular occupation is

exposed to disease at work (O*NET Online). Further details are in the “Phase Inclusion Criteria” section of this document.

We chose to run the phase assignment process only once due to the large size of the dataset and the large sizes of groups and demographic categories relative to PWGTP. This is not a Monte Carlo simulation.

We cannot assign the actual Social Vulnerability Index (SVI) (CDC 2020a) of each person in the PUMS data because SVI is computed using census tract level averages and the PUMS geographic specificity only goes down to the Public Use Microdata Area (PUMA) level. We instead approximate SVI at the individual level through a procedure described in further detail below. We do not perform any randomization in computing SVI.

The assignment of SVI to individuals involves both the person and housing unit PUMS data. We are able to link the two datasets as well as identify residents of the same housing unit with the person data, using the variable SERIALNO, and identifier unique to the housing unit that is available in both the person and housing unit data.

As a robustness check, we integrate 2018 census-tract level within-nation SVI data reported by the CDC into the national reserve simulations (ATSDR 2020). Further details on this robustness check are included in the “Our SVI Approach vs CDC’s Census-Tract Level Approach” section of this document.

A1.2. Reserve Allocation Procedure

Overall, we adopted an allocation procedure under NASEM’s phases in which reserve adjustments are designed and implemented in the most advantageous manner for the worse-off quartiles, chiefly by allowing them to proceed with the next phase when all worse-off populations of a phase have been served, and not wait until all members of the better-off population of the same phase have received their allocations. There are two different allocation procedures for state reserve and federal reserve implementations.

A1.2.1. The State Reserve Allocation Procedure

Similar to Schmidt et al. (2020), we implement a dynamic over-and-above reserve for the worse-off groups which makes maximum use of the worse-off reserve and allows jumps of the worse-off group to a further phase while the better off-group is served at an earlier phase if needed.

Suppose $r < 1$ is a reserve of the 25% (quartile) worse-off regions in each state (NASEM recommended $r = 10\%$).

1. We imagine the regions that qualify for the worse-off quartile reserve of each state are determined by
 - a. ranking the geographic regions of the state from the lowest-hit to hardest-hit based on the index used: for SVI the smallest geographic region is “census tract” and then
 - b. finding the geographic regions in the population-weighted *hardest-hit* quartile.
2. A *random lottery* determines the priority order among the individuals who belong to the same phase (if an individual qualifies for multiple phases, she is offered a vaccine in the earliest phase she qualifies for).
3. For each additional batch of vaccines allocated by the federal government, we assume that the whole batch is allocated to the states based on their population proportion.
4. Given a state s , let $\alpha_{1a}^s, \alpha_{1b}^s, \alpha_2^s, \alpha_3^s, \alpha_4^s < 1$ be the proportions of the population of individuals in the better-off regions with respect to the whole population of the state falling in prioritization

phases 1a, 1b, 2, 3, 4 according to the NASEM framework, respectively. (See Section A.2 for the explanation of who qualifies for which phase in the US population and how individuals qualifying for phases are determined in the simulation data).

5. Suppose V^s is the share of state s from the national batch of vaccines offered in a point in time. Each group of regions, the worse-off regions and better-off regions, is offered vaccines beginning from its members that qualify for Phase 1a. When the individuals in one phase end in one group, we continue to allocate the groups share to the next individual from the next phase.
 - As long as there are unserved individuals in the worse-off regions: if the current phase of the better-off regions is phase $x_s \in \{1a, 1b, 2, 3, 4\}$ we offer $V^s \cdot (1 - r) \cdot \alpha_{x_s}^s$ vaccines to the better-off regions and the remaining residents in the better-off regions are processed starting from phase x_s with respect to the priority order of the group's current phase; the remainder of the vaccines go to the worse-off regions and the remaining residents in the worse-off regions are processed starting from their current phase, which can be equal to or past phase x_s , with respect to the priority order of the group's current phase.
 - When there are no longer unserved individuals in the worse-off regions: All V^s vaccines are offered to the better-off regions and the remaining residents in the better-off regions are processed starting from phase x_s with respect to the priority order of the group's current phase

A1.2.2. The Federal Reserve Allocation Procedure

This procedure is novel in this paper. We adopted an allocation procedure under NASEM's phases in which reserve adjustments are designed and implemented in the most advantageous manner for the worse-off quartiles, chiefly by allowing them to proceed with the next phase when all worse-off populations of a phase have been served, and not wait until all members of the better-off population of the same phase have received their allocations.

We implement a dynamic over-and-above reserve for the worse-off groups which makes maximum use of the worse-off reserve and allows jumps of the worse-off group to a further phase while the better off-group is served at an earlier phase if needed.

Suppose $r < 1$ is a reserve of the 25% (quartile) worse-off population at the national level in the (NASEM recommended $r = 10$).

1. We imagine the regions that qualify for the worse-off quartile federal reserve determined by
 - a. ranking the geographic regions of the nation from the lowest-hit to hardest-hit based on nationwide SVI, where the smallest geographic region is "census tract", and then
 - b. finding the geographic regions in the population-weighted *hardest-hit* quartile in the whole nation.
2. A *random lottery* determines the priority order among the individuals who belong to the same phase (if an individual qualifies for multiple phases, she is offered a vaccine in the earliest phase she qualifies for).
3. Each state has a certain proportion of their population in the nationwide worse-off regions while the rest of the regions of state are classified as better-off. For each state s , let $\alpha_{1a}^s, \alpha_{1b}^s, \alpha_2^s, \alpha_3^s, \alpha_4^s < 1$ be the proportions of the population of individuals in the better-off regions with respect to the whole population of the state falling in prioritization phases 1a, 1b, 2, 3, 4.
4. Each group of regions, the worse-off regions and better-off regions in each state, is offered vaccines beginning from its residents that qualify for Phase 1a. When the individuals in one phase end in one group of the state, we continue to allocate the group's share to the next individual

from the next phase. We update the following at each point in time when a new batch of vaccines V becomes available for allocation at the federal level:

- c. Let S' be the set of states that have some residents who have not been offered vaccine yet.
- d. Let $S'_w \subseteq S'$ be the set of states that still have some worse-off region residents who have not been offered vaccine yet.
- e. For each $s \in S'$, let p_s be the population share of the state among states in S' relative to the overall initial populations of states.
- f. For each $s \in S'_w$, let w_s be the worse-off region population share of the state among states in S'_w relative to the overall initial worse-off region populations of states.
- g. For each $s \in S'$, let $x_s \in \{1a, 1b, 2, 3, 4\}$ be the phase of the better-off region population in the state that is about to be processed.
- h. For each $s \in S'$, the number of unreserved vaccines V_u^s is found as follows:
 - If $S'_w = \emptyset$, then $V_u^s = V \cdot p_s$, and
 - otherwise, $V_u^s = V \cdot (1 - r) \cdot p_s$.
- i. For each $s \in S'$, the number of reserved vaccines for the worse-off regions of the state is found as follows:
 - If $s \in S'_w$, then $V_r^s = V \cdot r \cdot w_s$, and
 - otherwise, $V_r^s = 0$.
- j. For each state $s \in S'_w$, $V_u^s + V_r^s$ is the whole state allotment.
 - The better-off regions are allocated $V_u^s \cdot \alpha_{x_s}^s$ units and the remaining residents in the better-off regions are processed starting from phase x_s with respect to the priority order of the group's current phase.
 - The worse-off regions are allocated $V_u^s \cdot (1 - \alpha_{x_s}^s) + V_r^s$ units and the remaining residents in the worse-off regions are processed starting from their current phase, which can be equal to or past x_s , with respect to the priority order of the group's current phase.
- k. For each state $s \in S' \setminus S'_w$, there are no worse-off region residents left to be offered vaccines, and V_u^s is the total state allotment as $V_r^s = 0$. Therefore, this allotment is fully allocated to the better-off regions and the remaining population is processed beginning from phase x_s with respect to the priority order of the group's current phase.

A1.3. Simulation Methodology

We consider three allocation scenarios: NASEM allocation 1) without reserve adjustment, 2) with the NASEM recommended over-and-above $r = 10\%$ worse-off state reserve, and 3) with over-and-above $r = 10\%$ worse-off federal -reserve. We follow the below procedure in our simulation analysis:

1. Approximate the SVI of every individual as explained in "Components and Calculation of SVI" section of this appendix at state level for state-level reserve and at national level for federal-level reserve.
2. Label every person in the PUMS with the highest phase group they qualify for. PUMS variables do not perfectly match up to the phase groups but can be used as a crude approximation. Our inclusion criteria for each phase is outlined in the "Phase Inclusion Criteria" section of this appendix.
3. For each individual, determine whether they fall in the most disadvantaged quartile by intrastate SVI and national SVI.
4. For each state's population in each phase, calculate the demographic averages (proportion breakdown of race, gender, age, etc.).

- Incrementally adding 100,000 vaccine units from 0 to about 323 million total doses allocated to the whole nation, use the allocation procedures explained in the previous subsection for scenarios (2) and (3) and for scenario (1) use the $r = 0$ version of either allocation procedure in the previous subsection.

We use statistical open-source R software (version 4.0.0) and Microsoft Excel in our simulation. The names of program files are given in a table at the end of this appendix. The programs and data files are available from the authors upon request.

A2. Phase Inclusion Criteria

A2.1. NASEM phases

We integrate Figure S-2 of the NASEM (2020) to arrive at the following phased approach:

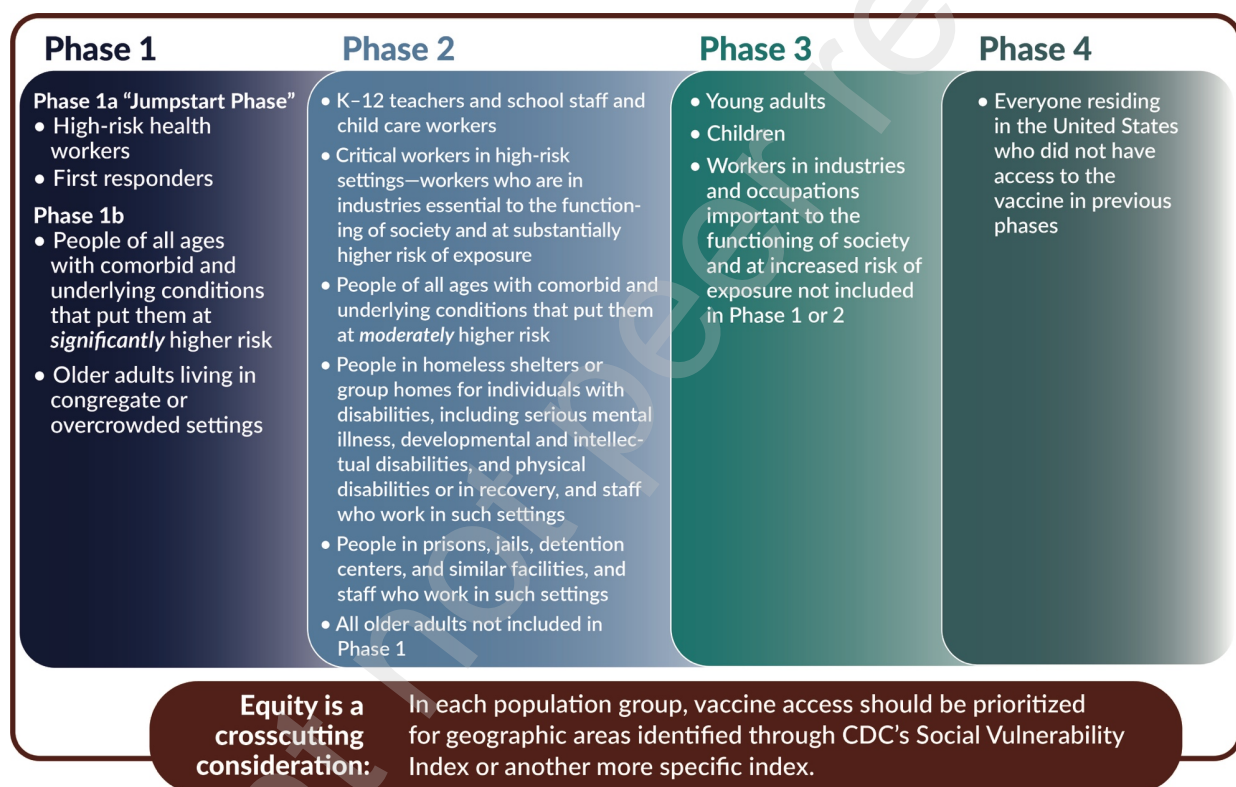


FIGURE S-2 A phased approach to vaccine allocation for COVID-19.

A2.2. Data issues and imperfect workarounds

A2.2.1. High-risk conditions and the BRFSS

The PUMS does not have data on the health conditions of the surveyed individuals. We therefore impute COVID-19 risk for each observation in the PUMS using the BRFSS data.

We label each observation in the BRFSS as significantly higher risk for COVID-19 if at least two of the following is true. We label an observation in the BRFSS as moderately higher risk for COVID-19 if at least one of the following is true:

- The individual has been “(Ever told) you had skin cancer”

- The individual has been “(Ever told) you had any other types of cancer”
- The individual has been “(Ever told) you have kidney disease” not including kidney stones, bladder infection or incontinence
- The individual has been “(Ever told) you have chronic obstructive pulmonary disease, C.O.P.D., emphysema or chronic bronchitis”
- The individual is obese, defined as having a BMI ≥ 30
- The individual has been “(Ever told) you had angina or coronary heart disease”
- The individual has been “(Ever told) you have diabetes,” not including diabetes only while pregnant, pre-diabetes, or borderline diabetes

For each of the above statements, we consider each statement false if the variable takes any value other than the one listed, including values for “don’t know/not sure,” “not asked or missing,” and “refused.” The above variable definitions and question wordings are sourced from the [2018 BRFSS codebook](#) (CDC 2018c).

The medical conditions chosen are cited by the NASEM guidelines when identifying individuals with comorbid and underlying conditions that put them at risk. The listed conditions are as follows:

- Cancer
- Chronic kidney disease
- COPD (chronic obstructive pulmonary disease)
- Immunocompromised state (weakened immune system) from solid organ transplant
- Obesity (body mass index [BMI] of 30 or higher)
- Serious heart conditions, such as heart failure, coronary artery disease, or cardiomyopathies
- Sickle cell disease
- Type 2 diabetes mellitus

Our approximation of these risk factors omits immunocompromised state from solid organ transplant and sickle cell disease due to lack of data in the BRFSS. Our risk factors also include all diabetes, not just type 2 diabetes, as the BRFSS does not distinguish between different types of diabetes.

Using the BRFSS data, labelled with whether or not there is significant or moderate COVID-19 risk, we group the data by age bin, sex, whether or not the person is Hispanic, and race/ethnicity other than Hispanic. For each interaction of those variables, we find the proportion of the population that is significant/moderate-risk. We omit any observations where one or more of these demographic variables are unknown or missing.

The age bins are the following:

18-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59, 60-64, 65-69, 70-74, 75-79, 80-84, 85+

For non-Hispanic race, those of known race are categorized into the following:

- White only
- Black or African American only
- American Indian or Alaskan Native only
- Asian Only
- Native Hawaiian or other Pacific Islander only
- Other race only
- Multiracial

Hispanic status is coded in a separate variable from non-Hispanic race.

We considered adding income bins to the interaction of demographic variables, but the introduction of income resulted in lack of data for some variable interactions. We leave income out to avoid this and for simplicity.

We output the significantly/moderately higher risk proportion of each demographic interaction in the BRFSS data. Then, we find the demographic probability of each PUMS person data observation being at significantly/moderately higher risk based on the proportions calculated from the BRFSS. We then throw a weighted coin for each individual in the PUMS to label each individual as significant/moderate risk or not. As the BRFSS does not have risk factor data for people under 18, we extrapolate the calculated risk probabilities of 18 to 24-year-olds of the same sex, race, and ethnicity for those under 18.

A2.3. Comment on phase sizes

The NASEM guidelines provide estimates of the sizes of each phase as below:

- Phase 1a (~5% of U.S. population)
- Phase 1b (~10% of U.S. population)
- Phase 2 (~30-35% of U.S. population)
- Phase 3 (~40-45% of U.S. population)
- Phase 4 (~5-15% of U.S. population)

These estimates do not take into account individual overlap between phases. Our phase by phase labeling of PUMS data as below maps each individual to the highest phase that they qualify for and thus identifies an individual's phase when taking into account phase overlap. Upon labeling individuals as in Section A2.4, the U.S. population is segmented among phases as follows:

- Phase 1a (~5% of U.S. population)
- Phase 1b (~11% of U.S. population)
- Phase 2 (~37% of U.S. population)
- Phase 3 (~39% of U.S. population)
- Phase 4 (~8% of U.S. population)

A2.4. Phase by phase details for labelling of PUMS data

A2.4.1. Phase 1a

A2.4.1.1. High-risk health workers

The NASEM guidelines define this group as frontline health care workers who are in hospitals, nursing homes, or providing home care with unavoidable risk of exposure and transmission of the virus. They also specify that morticians, funeral workers, and other death care professionals involved in handling bodies as part of this group. Additionally, NASEM includes pharmacists, public health workers, and dentists. In order to label these frontline health care workers and other individuals working in such high-potential exposure settings, we use the below NAICS (industry) codes:

- 622M: General Medical and Surgical Hospitals, and Specialty (Except Psychiatric and Substance Abuse) Hospitals
- 6231: Nursing Care Facilities (Skilled Nursing Facilities)
- 623M: Residential Care Facilities, Except Skilled Nursing Facilities

- 6216: Home Health Care Services

In order to label morticians and death care professionals, pharmacists, public health workers, and dentists, we rely on SOC (occupation) codes and include all workers with the below occupation code:

- 394031: Morticians, Undertakers, and Funeral Arrangers
- 3940XX: Embalmers, Crematory Operators, and Funeral Attendants
- 292052: Pharmacy Technicians
- 291051: Pharmacists
- 211022: Healthcare Social Workers
- 21109X: Other Community and Social Service Specialists
- 291020: Dentist
- 391091: Dental Assistants

Our labeling of public health workers is approximated by two occupation codes identifying Healthcare Social Workers and Other Community and Social Service Specialists (includes Community Health Workers). This is due to the lack of a clear occupation code in the PUMS data to indicate whether an individual is a public health worker.

The NASEM guidelines reference those frontline health care workers with unavoidable risk of exposure and transmission. In order to approximate this, we use the fraction of workers exposed to disease more than one time per month as calculated by Baker et al. (2020). Specifically, Baker et al. note that ~96.1% of Healthcare Support (2-digit SOC code of 31) and ~91.5% of Healthcare Practitioners and Technical (2-digit SOC code of 29) workers are exposed to disease at work more than once per month. As a crude approximation, we randomly sample our set of frontline health workers defined by NAICS and SOC code above to randomly assign 93.8% as health care workers at high-risk of exposure or transmission.

A2.4.1.2. First responders

NASEM guidelines identify this group as emergency medical services (EMS) personnel, police, and firefighters. We use occupation codes to assign individuals with an occupation code as one of the below as a first responder:

- 292042: Emergency Medical Technicians
- 292043: Paramedics
- 533011: Ambulance Drivers and Attendants, Except Emergency Medical Technicians
- 331011: First-Line Supervisors of Police and Detectives
- 333050: Police Officers
- 331021: First-Line Supervisors of Fire Fighting and Prevention Workers
- 332011: Firefighters

A2.4.2. Phase 1b

A2.4.2.1. People of all ages with comorbid and underlying conditions that put them at significantly higher risk

NASEM guidelines define individuals at significantly higher risk as individuals with two or more of the following conditions: cancer, chronic kidney disease, chronic obstructive pulmonary disease, immunocompromised state from solid organ transplant, obesity (body mass index [BMI] greater than or equal to 30), serious heart conditions, sickle cell disease, and type 2 diabetes mellitus. We assign this group using the procedure outlined in the Section A2.2.1.

A2.4.2.2. Older adults living in congregate or overcrowded settings

The NASEM guidelines identify these individuals as the older population living in congregate and overcrowded situations. NASEM does not propose an age cutoff and so we refer to 2018 CDC Vaccine Allocation during an Influenza Guidelines (CDC 2018a) and their definition of older as at least 65 years old. We label individuals as living in congregate and overcrowded situations as those who live in multigenerational housing or institutional group quarters (e.g., correctional facilities, nursing homes, or mental hospitals) (CDC 2018d).

A2.4.3. Phase 2

A2.4.3.1. K-12 teachers and school staff and child care workers

We rely on SOC codes to label K-12 teachers, school staff, and child care workers by including the below occupation codes:

- 252010: Preschool and Kindergarten Teachers
- 252020: Elementary and Middle School Teachers
- 252030: Secondary School Teachers
- 252050: Special Education Teachers
- 2530XX: Other Teachers and Instructors
- 259040: Teaching Assistants
- 2590XX: Other Educational Instruction and Library Workers
- 193034: School Psychologists
- 339094: School Bus Monitors
- 533051: Bus Drivers, School
- 119030: Education and Childcare Administrators
- 211021: Child, Family, and School Social Workers
- 211012: Educational, Guidance, and Career Counselors and Advisors
- 399011: Child care Workers

We note that the Teaching Assistant occupation code may include post-secondary Teaching Assistants, but we lack the granularity of occupation in the PUMS data to adjust for this.

A2.4.3.2. Critical workers in high-risk settings – workers who are in industries essential to the functioning of society and at substantially higher risk of exposure

The NASEM guidelines note that there is no single list of all workers who should be included in this phase. They reference the U.S. Department of Homeland Security categories of Essential Critical Infrastructure Workers as a list of critical workers that may have differing levels of exposure risk at work.

As an approximation, we crosswalk a list of DHS-defined critical infrastructure worker occupation codes to the occupation codes in the PUMS data. To the best of our knowledge, the version of the DHS list released on March 29, 2020 is the most recent list of critical infrastructure worker that explicitly labels SOC codes as critical infrastructure or not (DHS 2020). The PUMS data granularity of SOC codes is less detailed than the level of granularity of the DHS-defined critical infrastructure worker SOC codes. We define a PUMS SOC code as a critical infrastructure occupation if at least one of the more granular DHS-labeled SOC codes that map to the PUMS SOC code is a critical infrastructure occupation.

In order to identify critical infrastructure workers who are at a high level of disease exposure, we rely on BLS O*NET survey results that label SOC codes with an index corresponding to how often the workers surveyed are exposed to disease in the workplace (O*NET Online). The index is from 0 (Never) to 100 (Every Day) and a 50 corresponds to “Once a month or more but not every week”. As with critical infrastructure worker labeling, the PUMS SOC codes are less granular than the O*NET SOC codes. We define a PUMS SOC code’s disease exposure as the average disease exposure of the granular O*NET SOC codes that map to the PUMS SOC code.

Finally, to label individuals as critical workers with high-risk of exposure we take all individuals in a critical infrastructure occupation that have a disease exposure index of at least 50 where, as mentioned above, 50 corresponds to an exposure of “Once a month or more, but not every week”.

A2.4.3.3. People of all ages with comorbid and underlying conditions that put them at moderately higher risk

As in Phase 1b, NASEM defines comorbid conditions as the following: cancer, chronic kidney disease, chronic obstructive pulmonary disease, immunocompromised state from solid organ transplant, obesity (body mass index [BMI] greater than or equal to 30), serious heart conditions, sickle cell disease, and type 2 diabetes mellitus. NASEM guidelines define individuals with one or more of these conditions as at moderately higher risk. We assign this group using the procedure outlined in Section A2.2.1.

A2.4.3.4. People in homeless shelters or group homes for individuals with disabilities, including serious mental illness, developmental and intellectual, and physical disabilities or in recovery, and staff who work in such settings

To label individuals in homeless shelters or group homes, we first consider individuals who live in non-institutional group quarters defined to include college dormitories, military barracks, group homes, missions, and shelters (U.S. Census Bureau 2018). Unfortunately, the PUMS data granularity does not allow for labeling individuals as a particular type of non-institutional group quarters. However as a partial workaround, we are able to exclude undergraduate college students and active members of the military. We use the SOC code 119151, Social and Community Service Managers, to approximate the staff that work in such settings.

A2.4.3.5. People in prisons, jails, detention centers, and similar facilities, and staff who work in such setting

In order to label individuals as prisoners, we include all people living in institutionalized group quarters. This includes those in correctional facilities, nursing homes, and mental hospitals (U.S. Census Bureau 2018) and unfortunately the PUMS data is not granular enough to distinguish

between the three. We note that it is likely that the majority of nursing home residents are above the age of 65 and would be included in Phase 1b meaning that our imprecision in this phase is mostly due to the presence of mental hospitals among institutional group quarters. We label the staff of correctional facilities by including individuals with SOC codes as below:

- 331011: First-Line Supervisors of Correctional Officers
- 331012: Correctional Officers and Jailers

A2.4.3.6. All older adults not included in Phase 1

As discussed, the NASEM guidelines do not specify an age cutoff for defining older adults. As before, we rely on the 2018 CDC Vaccine Allocation Guidelines (CDC 2018a) and define an older adult as anyone at least 65 years old.

A2.4.4. Phase 3

A2.4.4.1. Young adults

Per NASEM guidelines, young adults are identified as any individual at least 18 years old and at most 30 years old.

A2.4.4.2. Children

Children are defined as any individual under the age of 18 years old.

A2.4.4.3. Workers in industries and occupations important to the functioning of society and at increased risk of exposure not included in Phase 1 or 2

We use the same process as above in Section 2.4.3.2 to identify critical infrastructure workers as defined by the DHS. The NASEM guidelines note that ideally, Phase 3 would vaccinate the remaining critical infrastructure workers that have not been vaccinated in a previous phase. We therefore include any individual in a critical risk occupation in this group.

A2.4.5. Phase 4

A2.4.5.1. Everyone residing in the United States who did not have access to the vaccine in previous phases

We include any remaining individuals who have not been assigned a phase in Phase 4.

A3. Components and Calculation of SVI

A3.1. Overview

We compute the CDC Social Vulnerability Index (SVI) at the intrastate and national level using the individual level weighting each of its 15 components equally. SVI is generally computed at the census tract or county level (CDC 2020a) but given the limitations of the geographic granularity of the PUMS data, we compute each of the 15 SVI components at the individual level.

A3.2. Computation details

Most of the SVI components become binary variables at the individual level. These binary components present a challenge when following the SVI procedure of computing, summing, and then again computing percentile ranks. In order to approximate the percentile rank of a now-binary SVI component for an individual within the individual's state using the PUMS data, we follow the below procedure:

- Suppose that we have a set of individuals indexed by i for which we wish to compute SVI component j within some state s
- Suppose that this SVI component j takes on n discrete values (in the binary case, $n=2$)
- We first compute the histogram representing the proportion of individuals in s that take on each of the n discrete values of j ; such proportions are denoted x_1, \dots, x_n where the set of x_i are ordered from the most advantaged to least advantaged discrete value of j
- For any individual i , suppose that individual i is assigned the k^{th} value of SVI component j
- Then, the value of component j for individual i is the k^{th} midpoint of the histogram as below:

$$SVI_{i,j} = (x_{k-1} - x_k) / 2$$

Throughout this document, the above will be referred to as the “percentile-rank midpoint procedure”.

Some of the SVI components are household-level variables – for example whether or not a housing unit is a mobile home. In this case, we use the percentile-rank midpoint procedure using the household histogram rather than the individual histogram. The decision as to whether we compute a histogram at the household or individual level is noted for each component in Section A4.3 as *level of summarization*.

There are instances in which the histogram step in the percentile-rank midpoint procedure is not conducted for every individual or every household. For example, one of the SVI components is whether an individual is unemployed or not. For this component, we do not include individuals who are not in the labor force during the histogram computation. For these individuals, the SVI component is set to whatever the percentile-rank midpoint procedure computes for an advantaged individual. This is done because we have no reason to believe these omitted individuals are disadvantaged with respect to the SVI component and they will need to be assigned some value for the summed SVI component percentile rank to be possible. The population included as disadvantaged in the histogram step is noted for each component in Section A4.3 as *numerator of proportion computation*. The population excluding the omitted individuals discussed above is noted as *denominator of proportion computation*.

Once we have computed each SVI component's intrastate or national percentile rank, we sum the computed percentile ranks across all 15. We then perform the percentile-rank midpoint procedure on the summed components to obtain a final intrastate and national individual-level SVI measure.

A3.3. Component by component details

We compute individual-level intrastate and national percentile ranks for each of the SVI components (CDC 2020a) as below:

1. Below poverty
 - a. Numerator of proportion computation: Population with family income-to-poverty ratio below 100%. Augment with a comparison of unadjusted individual income to [poverty thresholds of the sample's year](#) (U.S. Census Bureau 2020b) if income-to-poverty income is missing but individual income is not.
 - b. Denominator of proportion computation: All individuals with non-missing augmented income-to-poverty ratio
 - c. Level of summarization: Individual
2. Unemployed
 - a. Numerator of proportion computation: Population unemployed
 - b. Denominator of proportion computation: Population in the labor force
 - c. Level of summarization: Individual
3. Income
 - a. Individual income adjusted to constant dollars. Even though income is not binary, we use the percentile-rank midpoint procedure to set the percentile rank to account for multiple individuals with the same income. If individual income is missing, set to the individual's average household income component percentile rank. If both are missing, set the component to 0.
4. No high school diploma
 - a. Numerator of proportion computation: Population aged 25+ with less than a high school diploma
 - b. Denominator of proportion computation: Population aged 25+
 - c. Level of summarization: Individual
5. Aged 65 or older
 - a. Numerator of proportion computation: Population aged 65+
 - b. Denominator of proportion computation: Full population
 - c. Level of summarization: Individual
6. Aged 17 or younger
 - a. Numerator of proportion computation: Population aged 17 or younger
 - b. Denominator of proportion computation: Full population
 - c. Level of summarization: Individual
7. Older than age 5 with a disability
 - a. Numerator of proportion computation: Civilian non-institutionalized population with a disability over the age of 5

- b. Denominator of proportion computation: Civilian non-institutionalized population over the age of 5
 - c. Level of summarization: Individual
- 8. Single-parent household
 - a. Numerator of proportion computation: Single parent household with related children under the age of 18
 - b. Denominator of proportion computation: All households (excludes group quarters)
 - c. Level of summarization: Household
- 9. Minority
 - a. Numerator of proportion computation: All persons except white, non-Hispanic
 - b. Denominator of proportion computation: All persons
 - c. Level of summarization: Individual
- 10. Speaks English “Less than Well”
 - a. Numerator of proportion computation: Persons age 5+ who speak English “Not well” or “Not at all”
 - b. Denominator of proportion computation: Population aged 5+
 - c. Level of summarization: Individual
- 11. Multi-unit structures
 - a. Numerator of proportion computation: Households with 10 or more units
 - b. Denominator of proportion computation: All households (excludes group quarters)
 - c. Level of summarization: Household
- 12. Mobile homes
 - a. Numerator of proportion computation: Households that are a mobile home or trailer unit
 - b. Denominator of proportion computation: All households (excludes group quarters)
 - c. Level of summarization: Household
- 13. Crowding
 - a. Numerator of proportion computation: Households for which number of people in household / number of rooms in the household is greater than 1
 - b. Denominator of proportion computation: All occupied households (excludes group quarters)
 - c. Level of summarization: Household
- 14. No vehicle
 - a. Numerator of proportion computation: Households with no vehicle
 - b. Denominator of proportion computation: All occupied households
 - c. Level of summarization: Household (excludes group quarters)
- 15. Group quarters

- a. Numerator of proportion computation: Individuals living in group quarters
- b. Denominator of proportion computation: Full population
- c. Level of summarization: Individual

A4. Worse-off and better-off computation

Once we have assigned an individual an SVI at the intrastate and national level, we are able to define intrastate *worse-off* and *better-off* populations for the state reserve scenario and national *worse-off* and *better-off* populations for the federal reserve scenario. A higher value indicates in either case greater relative disadvantage. As discussed above, individual SVI in their final form are computed to be intrastate / national percentile ranks. We define the individuals with the highest 25% national SVI to be *worse-off*. Then this gives for the national reserve, what percentage and who in each state falls in the national worse-off group, while for the state reserve, intrastate SVI ranking of each state tells us who are the 25% most disadvantaged people in each state. All other individuals in each state are defined to be *better-off* (which is 75% of the population for the state reserve and an idiosyncratic percentage of population depending on the fraction of worse-off individuals in the state for the state reserve).

A5. Racial-Ethnic Group Definitions

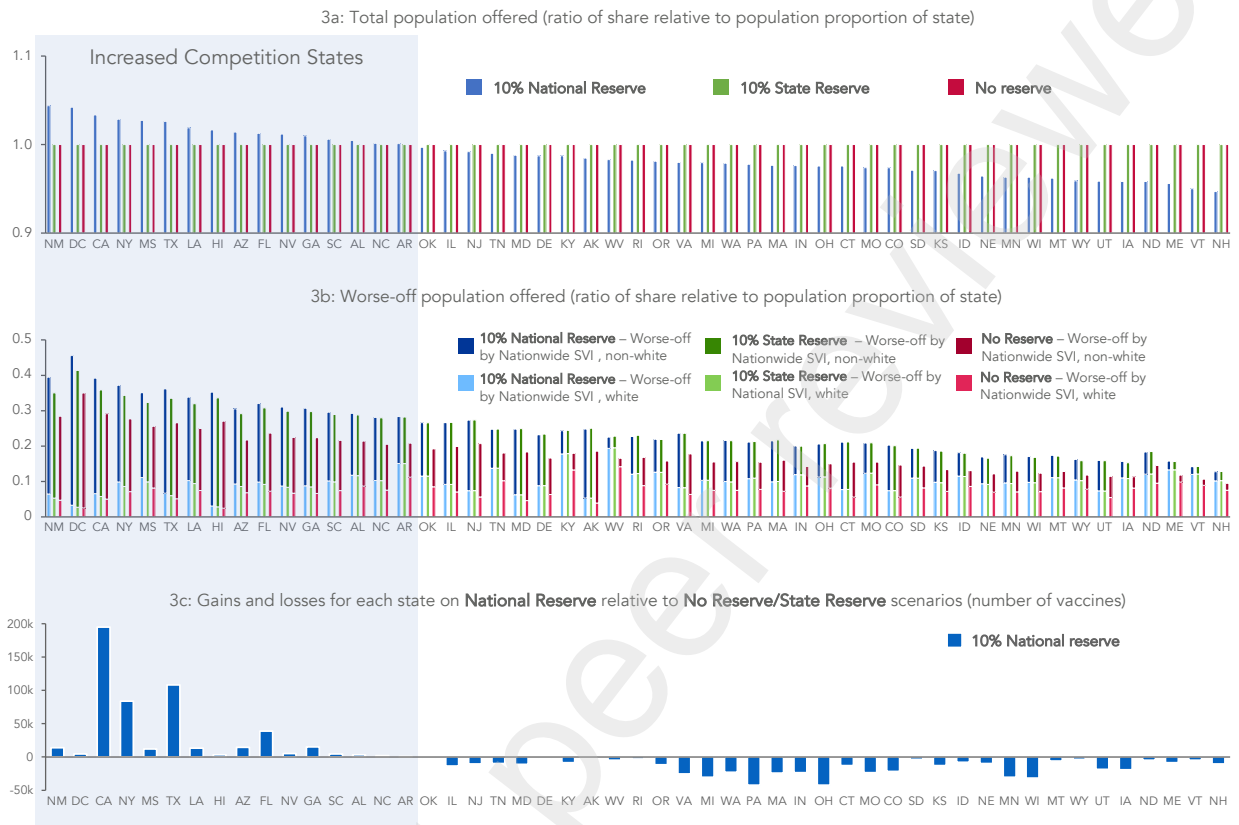
In our reporting we consider six different racial-ethnic groups. We use the label “white” to specify non-Hispanic white, “Black” to denote non-Hispanic Black, and “Hispanic” to denote Hispanic individuals. We include both non-Hispanic Asian and non-Hispanic Native Hawaiian and Other Pacific Islander when using the label “Asian”. We include non-Hispanic American Indian, non-Hispanic Alaska Native, and non-Hispanic American Indian and Alaska Native to denote “Indigenous”. Finally, we include non-Hispanic, some other race and individuals with two or more races in the “other” label. Thus, “non-white” labels everybody who are not labeled as “white”.

A6. Table of Program Files

File name	Description
brfss.R	Calculates proportion of each demographic group that has a significant/moderate-risk condition, using the BRFSS.
pums_nasem_phase_labelling.R	Assigns highest qualifying NASEM phase to PUMS data.
calculate_svi.R	Assigns SVI to PUMS data.
aggregate_cleaned_dataset.R	Consolidate SVI and NASEM phase data into one dataset.
build_master_soc_xwalk.R	Creates a crosswalk between critical infrastructure occupations, occupation disease exposure, and PUMS SOC codes.
allocate_vaccines.R	Iteratively perform vaccine allocation procedure.
compute_demographic_matrix.R	Create matrix of demographic proportions by state by phase by worse-off/better-off status.
compute_sim_descriptives.R	Outputs demographic proportions by number of doses allocated.
compare_with_actual_svi.R	Calculates share of state worse-off by 2018 CDC tract-level SVI measures.
compute_state_incrementals.R	Compute incremental units allocated to each state relative to no reserve allocation by reserve scenario.

A7. Allocation at 48 million Doses

Figure 3: Cumulative doses received by state by allocation scenario, 48 million doses allocated nationally

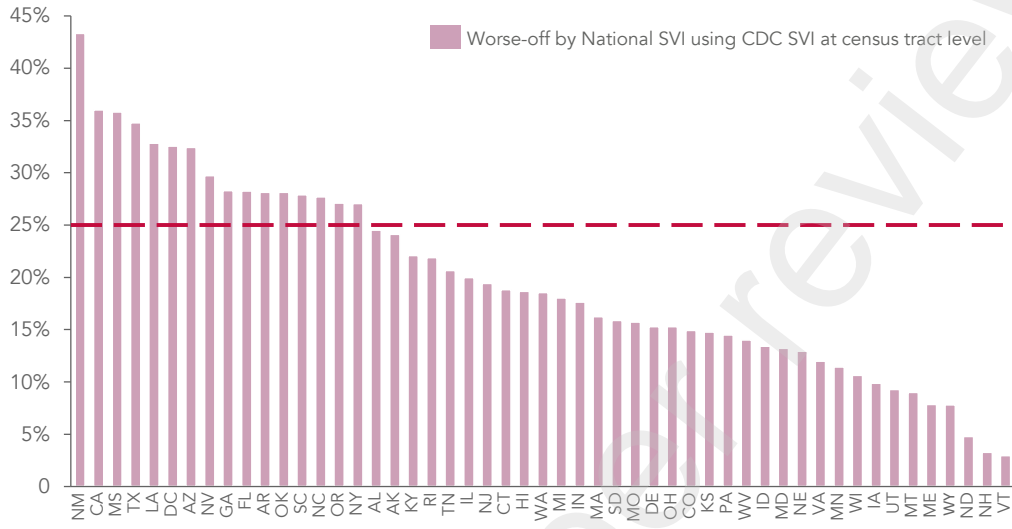


A8. Our SVI Approach vs CDC's Census-Tract Level Approach

In determining the worse-off individuals, as mentioned above, we used individual-based SVI rankings. The traditional approach taken by the CDC is using census-tract level rankings. Moreover, this is the highest geographic resolution data available at ATSDR (2020) and most likely this will be the dataset that will be used at the federal and state government levels.

Thus, we also construct the set of census tracts (regions) that would be qualified as worse-off at the national level and construct a version of Figure 1 in the main text using this measure for the states. Figure A1 is this corresponding figure. When we inspect the correlation of the figure with our Figure 1 we find a high correlation of 0.924.

Figure A1: Share of each state's population falling under the worse-off quartile nationally (SVI)



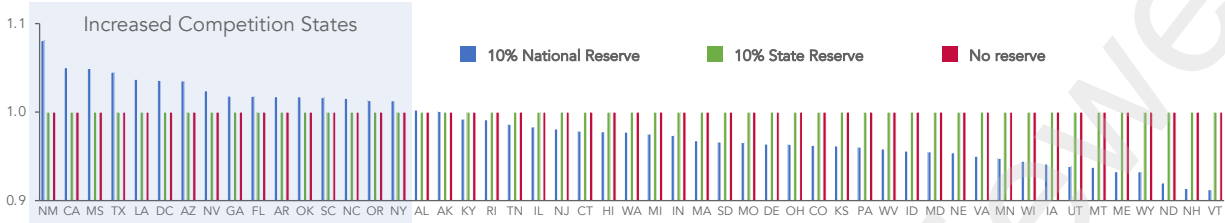
Correlation between state share worse-off by within-nation CDC SVI tract-level index and SVI index computed at individual level: **~0.924**

Notes: Worse-off is defined as a census tract in the top 25% of within-nation SVI rank. Census tracts with missing SVI within-nation rank are labeled as better-off. Sources: 2018 Center for Disease Control (CDC) Social Vulnerability Index (SVI).

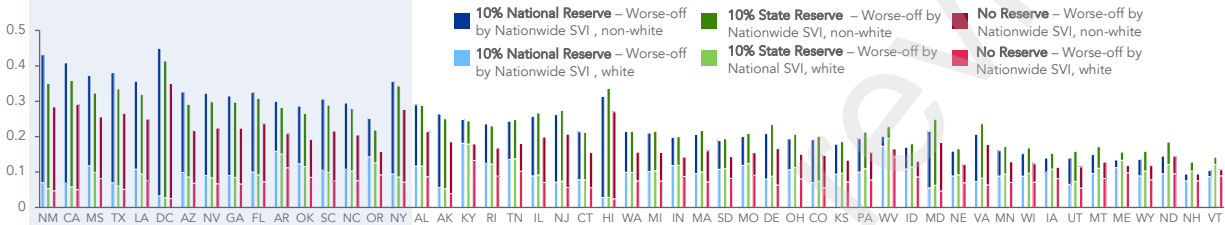
We also consider how we would have done allocation if we used these worse-off population fractions at the federal level and use our construction of individual-level SVI to choose who would be eligible to be in these worse-off groups in each state (as we do not have a way to link PUMS data to SVI census-tract addresses). The results are given below in Figure A2 for 48 million cumulative vaccines to compare with Figure 3 above.

Figure A2: Cumulative doses received by state by allocation scenario, 48 million doses allocated nationally, tract-level SVI to set reserve sizes in national reserve

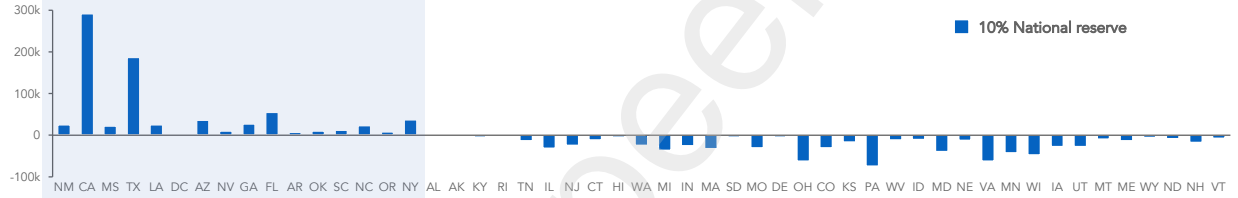
A3a: Total population offered (ratio of share relative to population proportion of state)



A3b: Worse-off population offered (ratio of share relative to population proportion of state)



A3c: Gains and losses for each state on National Reserve relative to No Reserve/State Reserve scenarios (number of vaccines)



Sources: National Academies of Sciences, Engineering, and Medicine Framework for Equitable Allocation of Vaccine for the Novel Coronavirus, American Community Survey (ACS) 5-Year, and 2018 Center for Disease Control (CDC) Social Vulnerability Index (SVI).

References for the Appendix

Agency for Toxic Substances and Disease Registry (ATSDR). 2020. “CDC SVI Data and Documentation Download.” Atlanta, Georgia: U.S. Department of Health and Human Services, Agency for Toxic Substances and Disease Registry. Retrieved on November 4, 2020 from https://www.atsdr.cdc.gov/placeandhealth/svi/data_documentation_download.html

Baker MG, Peckham TK, Seixas NS (2020) Estimating the burden of United States workers exposed to infection or disease: A key factor in containing risk of COVID-19 infection. PLoS ONE 15(4): e0232452. <https://doi.org/10.1371/journal.pone.0232452>

Centers for Disease Control and Prevention (CDC). 2018a. “Interim Updated Planning Guidance on Allocating and Targeting Pandemic Influenza Vaccine During an Influenza Pandemic.” Atlanta, Georgia: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention. Retrieved on August 11, 2020 from <https://www.cdc.gov/flu/pandemic-resources/pdf/2018-Influenza-Guidance.pdf>

Centers for Disease Control and Prevention (CDC). 2018b. Behavioral Risk Factor Surveillance System Survey Data. Atlanta, Georgia: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention

Centers for Disease Control and Prevention (CDC). 2018c. “LLCP 2018 Codebook Report” Atlanta, Georgia: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention. Retrieved on August 11, 2020 from https://www.cdc.gov/brfss/annual_data/2018/pdf/codebook18_llcp-v2-508.pdf

Centers for Disease Control and Prevention (CDC). 2018d. “Group Quarters Residence Rules” Atlanta, Georgia: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention. Retrieved on October 21, 2020 from <https://www.census.gov/topics/income-poverty/poverty/guidance/group-quarters.html>

Centers for Disease Control and Prevention (CDC). 2020a. “CDC SVI 2018 Documentation” January 31, Atlanta, Georgia: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention. Retrieved in September, 2020 from https://svi.cdc.gov/Documents/Data/2018_SVI_Data/SVI2018Documentation.pdf

Centers for Disease Control and Prevention (CDC). 2020b. “People with Certain Medical Conditions.” Atlanta, Georgia: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention. Retrieved on August 11, 2020 from https://www.cdc.gov/coronavirus/2019-ncov/need-extra-precautions/people-with-medical-conditions.html?CDC_AA_refVal=https%3A%2F%2Fwww.cdc.gov%2Fcoronavirus%2F2019-ncov%2Fneed-extra-precautions%2Fgroups-at-higher-risk.html

Centers for Disease Control and Prevention (CDC). 2020c. “COVID-19 Response. COVID-19 Case Surveillance Public Data Access, Summary, and Limitations”, Atlanta, Georgia: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention. Retrieved on August 14, 2020 from <https://data.cdc.gov/Case-Surveillance/COVID-19-Case-Surveillance-Public-Use-Data/vbim-akqf>

Department of Homeland Security Cybersecurity and Infrastructure Security Agency (DHS). 2020. “Guidance on the Essential Critical Infrastructure Workforce”. Retrieved on September 21, 2020 from

<https://www.lmiontheweb.org/more-than-half-of-u-s-workers-in-critical-occupations-in-the-fight-against-covid-19/>

Gayle H, Foegen W, Brown L, Kahn B, eds. 2020. *A Framework for Equitable Allocation of Vaccine for the Novel Coronavirus*. National Academy of Sciences, Engineering, and Medicine; August. Accessed September 21, 2020. <https://www.nationalacademies.org/our-work/a-framework-for-equitable-allocation-of-vaccine-for-the-novel-coronavirus>

Jukic AM, Baird DD, Weinberg CR, McConnaughey DR, Wilcox AJ. 2013. “Length of human pregnancy and contributors to its natural variation.” *Human Reproduction*, 28(10): 2848–2855.

Pathak PA, Schmidt H, Solomon A, Song E, Sönmez T, Ünver MU. 2020. Do Black and Indigenous Communities Receive their Fair Share of Vaccines Under the 2018 CDC Guidelines? *NBER Working Paper* No. 27817, September. doi: 10.3386/w27817, url: <https://www.nber.org/papers/w27817>

O*NET Online. Work Context—Exposed to Disease or Infections. Retrieved on September 21, 2020 from <https://www.onetonline.org/find/descriptor/result/4.C.2.c.1.b?s=1&a=1>

Schmidt H, Ünver MU, Williams MA, Pathak PA, Sönmez T, Gostin L. 2020. What Prioritizing Worse-Off Minority Groups for COVID-19 Vaccines Means Quantitatively: Practical, Legal and Ethical Implications. *SSRN Working Paper*, url: <https://ssrn.com/abstract=3716686> or <http://dx.doi.org/10.2139/ssrn.3716686>

U.S. Census Bureau. 2014. “American Community Survey Design and Methodology, Chapter 11: Weighting and Estimation.” Washington, D.C. :U.S. Census Bureau. Retrieved on August 11, 2020 from https://www2.census.gov/programs-surveys/acs/methodology/design_and_methodology/acs_design_methodology_ch11_2014.pdf

U.S. Census Bureau. 2014-2018. ACS Public Use Microdata Sample (PUMS). Washington, D.C. : U.S. Census Bureau.

U.S. Census Bureau. 2017. “American Community Survey Group Quarters Data Collection.” U.S. Census Bureau. Retrieved on September 5, 2020 at https://www.census.gov/content/dam/Census/programs-surveys/acs/Library/OutreachMaterials/ACSFlyers/2017%20Group%20Quarters%20flyer_508.pdf

U.S. Census Bureau. 2018. “Group Quarters/Residence Rules.” Retrieved on September 21, 2020 from <https://www.census.gov/topics/income-poverty/poverty/guidance/group-quarters.html>

U.S. Census Bureau. 2020a. “About PUMS.” Washington, D.C. : U.S. Census Bureau. Retrieved on September 4, 2020 from <https://www.census.gov/programs-surveys/acs/technical-documentation/pums/about.html>

U.S. Census Bureau. 2020b. Poverty Thresholds Data. January 22, 2020 Update. Washington, D.C. : U.S. Census Bureau. Retrieved on August 11, 2020 from <https://www.census.gov/data/tables/time-series/demo/income-poverty/historical-poverty-thresholds.html>