Title: Impacts of vaccination and Severe Acute Respiratory Syndrome Coronavirus 2 variants Alpha and Delta on Coronavirus Disease 2019 transmission dynamics in four metropolitan areas of the United States

Running Title: Impacts of COVID-19 vaccination and SARS-CoV-2 variants

One-Sentence Summary: Using a compartmental model parameterized to reproduce available reports of new Coronavirus Disease 2019 (COVID-19) cases, we quantified the impacts of vaccination and Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) variants Alpha (lineage B.1.1.7) and Delta (lineage B.1.617.2) on regional epidemics in the metropolitan statistical areas (MSAs) surrounding Dallas, Houston, New York City, and Phoenix.

Authors: Abhishek Mallela[1], Ye Chen[1], Yen Ting Lin, Ely F. Miller, Jacob Neumann, Zhili He, Kathryn E. Nelson, Richard G. Posner, William S. Hlavacek*

[1]These authors contributed equally.

Corresponding author. Email: wish@lanl.gov

Address for correspondence: William S. Hlavacek, Theoretical Biology and Biophysics Group, Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA.

Affiliations: Los Alamos National Laboratory, Los Alamos, New Mexico, USA (W.S. Hlavacek, Y.T. Lin, A. Mallela); Northern Arizona University, Flagstaff, Arizona, USA (Y. Chen, Z. He, E.F. Miller, K.E. Nelson, J. Neumann, R.G. Posner).

ABSTRACT

To characterize Coronavirus Disease 2019 (COVID-19) transmission dynamics in each of the metropolitan statistical areas (MSAs) surrounding Dallas, Houston, New York City, and Phoenix in 2020 and 2021, we extended a previously reported compartmental model accounting for effects of multiple distinct periods of non-pharmaceutical interventions by adding consideration of vaccination and Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) variants Alpha (lineage B.1.1.7) and Delta (lineage B.1.617.2). For each MSA, we found region-specific parameterizations of the model using daily reports of new COVID-19 cases available from January 21, 2020 to October 31, 2021. In the process, we obtained estimates of the relative infectiousness of Alpha and Delta as well as their takeoff times in each MSA (the times at which sustained transmission began). The estimated infectiousness of Alpha ranged from 1.1x to 1.4x that of viral strains circulating in 2020 and early 2021. The estimated relative infectiousness of Delta was higher in all cases, ranging from 1.6x to 2.1x. The estimated Alpha takeoff times ranged from February 1 to February 28, 2021. The estimated Delta takeoff times ranged from June 2 to June 26, 2021. Estimated takeoff times are consistent with genomic surveillance data.

Keywords: Coronavirus Disease 2019 (COVID-19), Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2), vaccination, SARS-CoV-2 variant Alpha (lineage B.1.1.7), SARS-CoV-2 variant Delta (lineage B.1.617.2), mathematical model, Bayesian inference

INTRODUCTION

In 2020, Coronavirus Disease 2019 (COVID-19) transmission dynamics were significantly influenced by non-pharmaceutical interventions [1–7]. In 2021, other factors arose with significant impacts on disease transmission, namely, vaccination [8–9] and emergence of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) variants [10–11].

Mass vaccination in the United States (US) began on December 14, 2020 [12], with demonstrable reduction of disease burden within vaccinated populations [13]. As the vaccination campaign progressed into March 2021, there was widespread reduction in disease incidence [14] and relaxation of state-mandated non-pharmaceutical interventions [15].

In early 2021, SARS-CoV-2 variant Alpha (lineage B.1.1.7) spread across the US and became the dominant circulating strain [16]. By the end of July 2021, the Delta variant (lineage B.1.617.2) had supplanted Alpha [17], concomitant with increases in new COVID-19 case detection [14]. Both Alpha and Delta have been estimated to be more transmissible than strains circulating earlier [18–23], and it was determined that vaccinated persons infected with Alpha and Delta were capable of transmitting disease [24, 25].

The Alpha variant was first detected in Kent, England in September 2020 [26], and was declared a variant of concern (VOC) on December 18, 2020 [26]. The Delta variant was first detected in Maharashtra, India in October 2020 [27] and was declared a VOC on May 6, 2021 [26]. In the literature, previous modeling works have investigated the relative transmissibility and timing of SARS-CoV-2 variants Alpha, Delta, and Omicron in multiple regions/countries, including England, Greece, Iran, China, and the US [28–36]. Modeling approaches applied in these studies involved deterministic [32–34] and stochastic [28, 30] compartmental models of

COVID-19 transmission. Other approaches involved statistical models and various forms of regression, e.g., multinomial logistic regression [29] and multivariable binary hyperbolastic regression [36]. Bayesian methods, including the sequential Bayesian method [32], a Bayesian evidence synthesis framework [30], a Bayesian approach to estimate the effective reproduction number R_t [31] and Markov Chain Monte Carlo (MCMC) sampling [28, 33], were predominantly used. The data considered in these studies included case data [28, 30–35] viral load data [28], seroprevalence data [30], information on deaths and hospital admissions [33], age-specific vaccine coverage data [31], sequencing data [29], and biospecimen data [35].

In earlier work, we demonstrated that new COVID-19 case detection over various periods in 2020 can be faithfully reproduced for 280 (out of the 384) metropolitan statistical areas (MSAs) in the US and all 50 states by region-specific parameterizations of a compartmental model that accounts for time-varying non-pharmaceutical interventions [5–7]. We found that the multiple surges in disease incidence seen in 2020 [14] could be explained by changes in protective disease-avoiding behaviors, which we will refer to collectively as social-distancing. However, in 2021, the model lost its ability to capture disease transmission dynamics, presumably because of the impacts of vaccination and the emergence of more transmissible SARS-CoV-2 variants, namely, Alpha and Delta. Here, to quantify the impacts of vaccination and SARS-CoV-2 variants Alpha and Delta on COVID-19 transmission dynamics, we extended the model of Lin et al. [5] by adding consideration of vaccination and variants with increased transmissibility. We then found region-specific parameterizations of the model using vaccination and surveillance case data available for the MSAs surrounding Dallas, Houston, New York City, and Phoenix.

METHODS

Data

Daily reports of new confirmed COVID-19 cases were obtained from the GitHub repository maintained by *The New York Times* newspaper [37]. Daily reports of newly completed vaccinations were obtained from the Covid Act Now database for the MSAs surrounding New York City and Phoenix [38]. Because of reporting gaps in the Covid Act Now database, we used a different source of vaccination data for the MSAs surrounding Dallas and Houston, the *Democrat and Chronicle* newspaper [39]. County-level surveillance and vaccination data were aggregated to obtain daily case and vaccination counts for the MSAs surrounding Dallas, Houston, New York City, and Phoenix. In the case of a missing daily report, we imputed the missing information as described in the Appendix.

Compartmental Model for Disease Transmission Dynamics

We used the compartmental model illustrated in Figure 1 (and Appendix Figure 1) to analyze data available for each MSA of interest. The model consists of ordinary differential equations (ODEs) describing the dynamics of 40 populations (state variables) (Appendix Equations 1–38). The state variables are each defined in Appendix Table 1. Model parameters are defined in Tables 1–3. Key features of the model are described below, and a full description of the model is provided in the Appendix.

We extended the model of Lin et al. [5] by including 15 new populations and 28 new transitions. Briefly, new parts of the model can be described as follows. Vaccination is modeled by moving susceptible persons (in the S_M and S_P populations) into the V_1 population. Another consequence of vaccination is the movement of recovered unvaccinated persons (in the R_U population) into the R_V population. The rate of vaccination is changed as needed to match the

empirical daily rate of vaccination. Recovered and susceptible persons have the same *per capita* probability of vaccination. Persons in S_M are mixing (i.e., not practicing social-distancing) and persons in S_P are practicing social-distancing (and thereby protected from infection to a degree). The series of transitions involving the populations $V_1, \dots, V_6$ was introduced to model the immune response to vaccination (i.e., the amount of time required for vaccination to induce neutralizing antibodies). With this approach, the time from vaccination to appearance of neutralizing antibodies is a random variable characterized by an Erlang distribution. Persons in $V_1, \dots, V_6$ may be infected. Persons in V_6 transition to one of the following four populations: $S_{V,1}, \dots, S_{V,4}$. These populations represent persons with varying degrees of immune protection. Persons in $S_{V,1}$ are not protected against productive infection (i.e., an infection that can be transmitted to others) by any viral strain. Persons in $S_{V,2}$ are protected against productive infection by viral strains present before the emergence of Alpha but not Alpha or Delta. Persons in $S_{V,3}$ are protected against productive infection by viral strains present before the emergence of Alpha and also Alpha but not Delta. Persons in $S_{V,4}$ are protected against productive infection by all of the viral strains considered up to October 31, 2021. Vaccinated persons who become infected move into E_V. The time spent in E_V corresponds to the length of the incubation period for vaccinated persons. The mean duration of the incubation period is taken to be the same for vaccinated and unvaccinated persons; however, as a simplification, for vaccinated persons, the time spent in the incubation period is taken to consist of a single stage and consequently is exponentially distributed (vs. Erlang distributed for unvaccinated persons). Non-quarantined exposed persons in populations E_V and $E_{i,M}$ and $E_{i,P}$ for $i = 2, \dots, 5$ are taken to be infectious. Persons exiting E_V leave the incubation period and enter the immune clearance phase of infection, during which they may be asymptomatic (A_V) or symptomatic with mild disease (I_V). Non-quarantined asymptomatic

persons in populations A_V, A_M, and A_P are taken to be infectious. Persons in A_V are assumed to eventually recover (i.e., to enter R_V). Persons with mild symptomatic disease may recover (i.e., enter R_V) or experience severe disease, at which point they move to H_V (in hospital or isolated at home). Vaccinated persons have a diminished probability of severe disease in comparison to unvaccinated persons. Persons in H_V either recover (move to R_V) or die from COVID-19 complications (move to D). For a person with severe disease, the probability of death is independent of vaccination status. We assume that vaccinated persons do not participate in social-distancing, quarantine, or self-isolation driven by symptom awareness.

The new compartments and transitions, which are highlighted in Figure 1, capture vaccination among susceptible persons, recovered persons and infected non-quarantined persons without symptoms at a time-varying *per capita* rate $\mu(t)$. The value of $\mu(t)$ changes daily for consistency with MSA-specific daily reports of completed vaccinations (Appendix Equation 37) from the COVID Act Now database and the *Democrat and Chronicle* COVID-19 vaccine tracker. The model also captures immune responses to vaccination yielding varying degrees of protection and consequences of breakthrough infection of vaccinated persons. Vaccine protection against transmissible infection was taken to be variant-dependent.

We introduced a dimensionless step function, denoted $Y_\theta(t)$ (Appendix Equation 38), which multiplies the disease-transmission rate constant β to account for m variants. In this study, $m = 2$ (see Appendix Equations 1–4, 18–22, and 24). Thus, in the new model, the quantity $Y_\theta(t)\beta$ (vs. β alone) characterizes disease transmissibility at time t. The step function $Y_\theta(t)$ was initially assigned a value of $y_0 = 1$, and the value of $Y_\theta(t)$ was allowed to increase at times $t = \theta_1$ and $t = \theta_2$ (Appendix Equation 38). The disease transmission rate constant of the Alpha variant was considered by introducing a step increase from $y_0\beta \equiv 1\beta$ to $y_1\beta$ (with $y_1 > 1$) at

time $t = \theta_1$ (the Alpha takeoff time). Similarly, the disease transmission rate constant of the Delta variant was considered by introducing a step increase from $y_1\beta$ to $y_2\beta$ (with $y_2 > y_1$) at time $t = \theta_2 > \theta_1$ (the Delta takeoff time). We will refer to y_1 and y_2 as the Alpha and Delta transmissibility factors, respectively.

As in the original model of Lin et al. [5], the extended model accounts for a series of $n + 1$ distinct social-distancing periods (an initial period and n additional periods). Social-distancing periods are characterized by two step functions: $P_\tau(t)$ and $\Lambda_\tau(t)$. The values of these functions change coordinately at a set of switch times $\tau = (\sigma, \tau_1, \ldots, \tau_n)$ (Appendix Equations 35 and 36), where σ is the start time of the initial social-distancing period and τ_i is the start time of the ith social-distancing period after the initial social-distancing period. The values of $P_\tau(t)$ and $\Lambda_\tau(t)$ are zero before time $t = \sigma$. The value of $P_\tau(t)$ defines the steady-state setpoint fraction of the susceptible population practicing social-distancing at time t, and the value of $\Lambda_\tau(t)$ defines a time scale for establishment of the steady state. The value of $\Lambda_\tau(t)$ is an eigenvalue equal to a sum of social-distancing rate constants [5]. The non-zero values of $P_\tau(t)$ and $\Lambda_\tau(t)$ are denoted $p_0, \ldots, p_n$ and $\lambda_0, \ldots, \lambda_n$. We assume that vaccinated persons do not practice social-distancing. Recall that we use the term "social-distancing" to refer to behaviors adopted to protect against infection. These behaviors are assumed to reduce the risk of infection by a factor m_b.

Parameters

As indicated in Tables 1 and 2, we used MSA-specific case reporting data available up to October 31, 2021 to infer MSA-specific values for parameters characterizing the start time of the local epidemic (t_0), local disease transmissibility of ancestral viral strains (β), local social-distancing dynamics (σ, λ_0, p_0, τ_i, λ_i and p_i, for $i = 1, \ldots, n$), local emergence of variants (θ_1,

y_1, θ_2, y_2), the local rate of new case detection (f_D), and noise in local case detection and reporting (r). Values for other parameters were fixed (Table 3); inferences are conditioned on these fixed parameter estimates. There are 18 fixed parameters taken to be applicable for all MSAs. The total regional population S_0, which is taken to be fixed, was set on the basis of census data. The real-time *per capita* vaccination rate $\mu(t)$, a piecewise linear function, was set for consistency with the current empirical *per capita* rate of vaccination [38]. We adopted the fixed parameter estimates of Lin et al. [5]. New fixed parameter estimates made in this study for m_h, f_0, f_1, f_2, and k_V are explained in the Appendix. The m_h parameter characterizes vaccine protection against severe disease, the f_0, f_1, and f_2 parameters account for differential vaccine effectiveness against the three viral strains considered in this study (ancestral, Alpha, and Delta), and the k_V parameter characterizes the waiting time between vaccination and the acquisition of vaccine-induced immunity. Our model does not account for gradual loss of immunity over time. We let $\tilde{\mu}(t) = \mu_i$ for times t throughout the ith day after January 21, 2020 (Appendix Equation 37), where μ_i is the fraction of the local population reported to complete vaccination over the 1-d surveillance period [38]. We then defined $\mu(t)$ as the piecewise linear interpolant to $\tilde{\mu}(t)$. In summary, for a given inference, the number of adjustable parameters was $2m + 3n + 4$, where m is the number of variants under consideration ($m = 2$ in this study) and n is the number of distinct social-distancing periods being considered beyond an initial social-distancing period. The setting for n was determined through a model-selection procedure described in the Appendix.

Auxiliary Measurement Model

We assumed that state variables of the compartmental model (Figure 1, Appendix Table 1) are related to the expected number of new cases reported on a given calendar date through an

auxiliary measurement model (Appendix Equations 39 and 40). The measurement model has one parameter: f_D, the region-specific fraction of new symptomatic infections detected. Thus, $f_D \in [0, 1]$. As a simplification, we considered f_D to be time-invariant. This simplification means that we assumed, for example, that case detection was neither limited nor strongly influenced by testing capacity, which varied over time. This assumption is reasonable if, for example, case detection is mainly determined by presentation for testing and, moreover, the motivations and societal factors that influence presentation remained roughly constant over the period of interest. One can also interpret f_D as the time-averaged case detection rate. The measurement-model parameter f_D was inferred jointly with the adjustable model parameters and the likelihood parameter r (see below).

Statistical Model for Noise in Case Detection and Reporting

We assumed that noise in new case detection and reporting on the ith day after January 21, 2020 is captured by a negative binomial distribution $\mathrm{NB}(r, q_i)$ centered on $I(t_i, t_{i+1})$, the expected number of new cases detected over the ith day after January 21, 2020 as given by the compartmental model and the auxiliary measurement model (Appendix Equations 1–40). These and other assumptions led to the likelihood function used in inference (Appendix Equations 41–43). We determined the probability parameter q_i in $\mathrm{NB}(r, q_i)$ using Appendix Equation 43; the dispersion parameter r was taken to be a time-invariant adjustable parameter applicable for all days of case reporting. The likelihood parameter r was inferred jointly with the adjustable model parameters and the adjustable measurement-model parameter f_D.

Computational Procedures

We determined the intervals of the step functions $Y_\theta(t)$, $P_\tau(t)$, and $\Lambda_\tau(t)$ (i.e., θ and τ) using a model-selection procedure described in the Appendix. Simulations and Bayesian inferences were performed as previously described [5–7] and in the Appendix. Files needed to reproduce inferences using the software package PyBioNetFit [40] are available online [41]. The files include case data, vaccination data, and diagnostic plots related to Bayesian inference using Markov chain Monte Carlo (MCMC) sampling, including trace plots [41]. Summary diagnostics characterizing the sampling for each MSA are given in Table 4. Briefly, we computed the stable Gelman-Rubin statistic using the methodology of Vats and Knudson [42]. An advantage of this statistic (over the original Gelman-Rubin statistic) is that only one Markov chain is required in sampling. In addition, the stable Gelman-Rubin statistic and effective sample size (ESS) have a one-to-one relationship.

RESULTS

Explanatory power of the model

As illustrated in Figures 2A, 3A, 4A, and 5A for the Dallas, Houston, New York City, and Phoenix MSAs respectively, the new model is able to explain surveillance case data over the period starting on January 21, 2020 and ending on October 31, 2021. The surveillance case data—daily reports of newly detected COVID-19 cases—for each of the 4 MSAs of interest largely lie within the 95% credible interval of the posterior predictive distribution for new case detection, which indicates that each regional model has explanatory power for the period of interest. Parameter estimates are summarized in Tables 1–3. Table 1 provides region-specific maximum *a posteriori* (MAP) estimates obtained through Bayesian inference enabled by Markov chain Monte Carlo (MCMC) sampling. Table 4 provides metrics that measure the

quality of MCMC sampling. The diagnostic results of Table 4 indicate that sampling was acceptable, because all multivariate potential scale reduction factors (PSRFs) are very close to 1 and all multivariate effective sample sizes (ESSs) are at least 100. Diagnostic plots are provided online [41].

Quantification of the impacts of non-pharmaceutical interventions and emergence of variants

Each regional model (parameterized to reproduce MSA-specific case reports) provides insight into the impacts of social-distancing behaviors and the emergence of the Alpha and Delta variants; compare panel A and the corresponding panel B in Figures 2–5. For example, as can be seen in Figure 4A, the New York City MSA experienced four notable surges in disease incidence over the period of interest. Figure 4B suggests that the first surge ended because of adoption of social-distancing behaviors, the second surge occurred because of relaxation of social-distancing behaviors, the third surge was caused by Alpha, and the fourth surge was caused by Delta. Interestingly, in other MSAs, Alpha had relatively little impact on disease incidence (compare Figures 4A and 4B to Figures 2A and 2B, 3A and 3B, and 5A and 5B). These differences are attributable to various factors, including the lower contagiousness of COVID-19 in Dallas, Houston, and Phoenix relative to New York City (see entries for β in Table 1), the later arrival of Alpha in Houston and Phoenix relative to New York City, and further progress of mass vaccination in Phoenix relative to New York City. Alpha takeoff, as well as Delta takeoff, is tracked locally in panel B (rightmost vertical axis) of Figures 2–5, and vaccination progress is tracked locally in panel C of Figures 2–5.

Impacts of vaccination

On the basis of region-specific parameterizations, we estimated the immune and susceptible fractions of each MSA population, as well as the fractions that achieved immunity through infection and/or vaccination (see panel C in Figures 2–5). Each of these panels shows the time evolution of five different populations in an MSA. In each MSA, only a minority of the population remained susceptible to infection (with Delta) on October 31, 2021, with a sizable fraction of the susceptible population being protected to a degree against severe disease by having completed vaccination.

Estimates of transmissibility factors and takeoff times of Alpha and Delta

Our inferences provide quantitative insights into the transmissibility factors of Alpha and Delta (y_1 and y_2) and their takeoff times (θ_1 and θ_2) in each of the 4 MSAs of interest. Figure 6 shows the marginal posteriors of y_1, y_2, θ_1, and θ_2, which were found on the basis of surveillance data for each MSA (daily case counts) reported between January 21, 2020 and October 31, 2021. Maximum *a posteriori* (MAP) estimates and 95% credible intervals for θ_1 and θ_2 for each MSA are also shown in Figure 6. The MAP estimates for y_1 range from 1.1 (for the Dallas MSA) to 1.4 (for the New York City MSA). The MAP estimates for y_2 range from 1.6 (for the Dallas MSA) to 2.1 (for the New York City MSA). The MAP estimates for θ_1 range from February 1, 2021 (for the Dallas MSA) to February 28, 2021 (for the Phoenix MSA). The MAP estimates for θ_2 range from June 2, 2021 (for the Dallas MSA) to June 26, 2021 (for the Houston and Phoenix MSAs). The estimated takeoff times are consistent with regional genomic surveillance data [43], which are summarized by the shaded regions of panel A in Figures 2–5.

DISCUSSION

We extended a model for COVID-19 transmission dynamics that already incorporated time-varying changes in non-pharmaceutical interventions [5] to include the effects of vaccination and new variants. This model together with its region-specific parameterizations based on case data available through October 31, 2021 provide quantitative insights into the relative infectiousness of SARS-CoV-2 variants Alpha (lineage B.1.1.7) and Delta (lineage B.1.617.2). The increased transmissibility of Alpha and Delta in comparison to ancestral strains is characterized by the marginal posteriors for the transmissibility factors y_1 and y_2 shown in Figure 6 (panels A, C, E, and G). The maximum *a posteriori* (MAP) estimates of the transmissibility factors were similar across the four metropolitan statistical areas (MSAs) of interest (centered on Dallas, Houston, New York City, and Phoenix). The averages of our y_1 and y_2 MAP estimates indicate that Alpha was 1.2x more infectious than ancestral strains, whereas Delta was 1.9x more infectious (corresponding to Delta being approximately 50% more infectious than Alpha). These estimates are consistent with estimates provided in other studies [18–23]. Using a formula for the basic reproduction number R_0 derived earlier [7] and parameter estimates of Tables 1 and 3, we obtain an R_0 estimate of 12 for Delta in the New York City MSA, which places this SARS-CoV-2 variant among the more infectious viruses known.

We also obtained estimates of precisely when sustained transmission of Alpha and Delta began in each of the four geographically distinct MSAs. In earlier work, takeoff times for Alpha and Delta were estimated only for regions outside the US (e.g., Denmark [44] and Mexico [45]), for regions in the US different from those considered here (e.g., New England [46] and two MSAs in Northern California [47]), and for the entire US [48]. The takeoff times for the four MSAs in the present study are characterized by the marginal posteriors for θ_1 and θ_2 shown in Figure 6 (panels B, D, F, and H). The estimated takeoff times are consistent with the observed

prevalence of Alpha and Delta sequences detected in regional genomic surveillance [43], as can be seen by comparing the two shaded regions in panel A of Figures 2–5 against the changes in transmissibility $Y_\theta(t)$ depicted in the corresponding panel B of Figures 2–5. It should be noted that the case data shown in Figures 2–5 are from the MSAs of interest (through aggregation of county-level data), whereas the genomic surveillance data are from larger regions [43].

Our study has notable limitations, starting with the obvious uncertainties related to model assumptions and fixed parameter estimates, which are discussed in some detail in the Appendix. For example, our model assumes a constant rate of case detection and neglects gradual loss of sterilizing immunity over time. This latter simplifying assumption is supported by a study indicating that protection against re-infection from pre-Omicron variants is significant and remains high even after 40 weeks [49]. Moreover, we caution that the inference jobs performed in this study were challenging because of the relatively high-dimensional parameter spaces involved (in comparison to typical inferences involving an ODE model-constrained likelihood function). Diagnostics indicate good sampling (Table 4, [41]), but we cannot be entirely certain that the samples obtained fully characterize the parameter posteriors of interest. It can be seen in trace plots [41] that mixing across parameters was not uniform, and moreover, in some cases, there are indications of trends (manifesting as a trace plot that lacks the appearance of a horizontally extended, fuzzy caterpillar). The likely cause of these trace plots is poor local performance of the proposal kernel, which was optimized in the adaptive sampling scheme for global performance. The consequences of poor mixing may include biased parameter and uncertainty estimates. Another concern is that the model incorporates redundant disease-incidence surge mechanisms. In the model, an increase in viral infectiousness caused by the emergence of a new variant can be mimicked, to some extent, by a relaxation of social-

distancing, and vice versa. For the MSAs considered here, inferred social-distancing levels were low at the time of Alpha and Delta emergence, so the inferred transmissibility factors probably reflect, at least mostly, changes in intrinsic viral infectiousness.

It is known that viral transmissibility does not depend only on viral features but also on population features [7]. This study looked at multiple regions to ascertain whether population features varied significantly. Our analysis indicates that population features influencing transmission did not vary dramatically across these similar urban regions because we obtained similar region-specific estimates for the Alpha and Delta transmissibility factors. The Alpha transmissibility factors for the Dallas, Houston, New York City, and Phoenix MSAs were 1.1x, 1.2x, 1.4x, and 1.3x, respectively. the Delta transmissibility factors for these MSAs were 1.6x, 2.0x, 2.1x, and 1.8x, respectively.

We show that it is possible to gain insights into variant dynamics through a mechanistic modeling approach. This study provides an example for modelers interesting in understanding the impacts of a mass vaccination campaign and emergence of variants with altered transmissibility during an epidemic of an aerosol-transmitted respiratory disease similar to COVID-19.

In summary, this report provides estimates of Alpha and Delta transmissibility for specific regions within the US as well as their takeoff times.

ACKNOWLEDGMENTS

W.S.H., Y.T.L., and A.M. were supported by the LDRD program at Los Alamos National Laboratory (project 20220268ER). Y.C., Z.H., E.F.M., J.N., K.E.N., and R.G.P. were supported by a grant from the National Institute of General Medical Sciences of the National Institutes of

Health (grant R01GM111510). A.M. was supported by the 2020 Mathematical Sciences Graduate Internship program, which is sponsored by the Division of Mathematical Sciences of the National Science Foundation, and the Center for Nonlinear Studies at Los Alamos National Laboratory. Computational resources used in this study included Northern Arizona University's Monsoon computer cluster, which is funded by Arizona's Technology and Research Initiative Fund, and the FARM computer cluster at the University of California, Davis. Y.C. thanks Song Chen (University of Wisconsin, La Crosse, Wisconsin, USA) for technical assistance.

REFERENCES

[1] Courtemanche C, Garuccio J, Le A, Pinkston J, Yelowitz A. Strong social-distancing measures in the United States reduced the COVID-19 growth rate. Health Aff (Millwood). 2020;39:1237–46. https://doi.org/10.1377/hlthaff.2020.00608

[2] Hsiang S, Allen D, Annan-Phan S, Bell K, Bolliger I, Chong T, et al. The effect of large-scale anti-contagion policies on the COVID-19 pandemic. Nature. 2020;584:262–7. https://doi.org/10.1038/s41586-020-2404-8

[3] Matrajt L, Leung T. Evaluating the effectiveness of social-distancing interventions to delay or flatten the epidemic curve of coronavirus disease. Emerg Infect Dis. 2020;26:1740–8. https://doi.org/10.3201/eid2608.201093

[4] Bo Y, Guo C, Lin C, Zeng Y, Li HB, et al. Effectiveness of non-pharmaceutical interventions on COVID-19 transmission in 190 countries from 23 January to 13 April 2020. Int J Inf Dis. 2021;102:247–53. https://doi.org/10.1016/j.ijid.2020.10.066

[5] Lin YT, Neumann J, Miller EF, Posner RG, Mallela A, Safta C, Ray J, Thakur G, Chinthavali S, Hlavacek WS. Daily forecasting of regional epidemics of Coronavirus

376 Disease with Bayesian uncertainty quantification, United States. Emerg Infect Dis.
377 2021;27:767–78. https://doi.org/10.3201/eid2703.203364
378 [6] Mallela A, Neumann J, Miller EF, Chen Y, Posner RG, Lin YT, Hlavacek WS. Bayesian
379 inference of state-level COVID-19 basic reproduction numbers across the United States,
380 Viruses. 2022;14:157. https://doi.org/10.3390/v14010157
381 [7] Mallela A, Lin YT, Hlavacek WS. Differential contagiousness of respiratory disease
382 across the United States, Epidemics. 2023;45:100718.
383 https://doi.org/10.1016/j.epidem.2023.100718
384 [8] Oliver SE, Gargano JW, Marin M, Wallace M, Curran KG, Chamberland M, et al. The
385 Advisory Committee on Immunization Practices' interim recommendation for use of
386 Pfizer-BioNTech COVID-19 vaccine—United States, December 2020. MMWR Morb
387 Mortal Wkly Rep. 2020;69:1922–4. https://doi.org/10.15585/mmwr.mm6950e2
388 [9] Oliver SE, Gargano JW, Marin M, Wallace M, Curran KG, Chamberland M, et al. The
389 Advisory Committee on Immunization Practices' interim recommendation for use of
390 Moderna COVID-19 vaccine—United States, December 2020. MMWR Morb Mortal
391 Wkly Rep. 2021;69:1653–6. https://doi.org/10.15585/mmwr.mm695152e1
392 [10] Galloway SE, Paul P, MacCannell DR, Johansson MA, Brooks JT, MacNeil A, et
393 al. Emergence of SARS-CoV-2 B.1.1.7 lineage—United States, December 29, 2020–
394 January 12, 2021. MMWR Morb Mortal Wkly Rep. 2021 / 70(3);95–99.
395 http://dx.doi.org/10.15585/mmwr.mm7003e2
396 [11] Herlihy R, Bamberg W, Burakoff A, Alden N, Severson R, Bush E, et al. Rapid
397 increase in circulation of the SARS-CoV-2 B.1.617.2 (Delta) variant—Mesa County,

398 Colorado, April–June 2021. MMWR Morb Mortal Wkly Rep. 2021;70:1084–7.
399 https://doi.org/10.15585/mmwr.mm7032e2
400 [12] Department of Health and Human Services. COVID-19 Vaccine Milestones [cited
401 2022 Dec 21] https://www.hhs.gov/coronavirus/covid-19-
402 vaccines/distribution/index.html
403 [13] Daniel W, Nivet M, Warner J, Podolsky DK. Early evidence of the effect of
404 SARS-CoV-2 vaccine at one medical center. N Engl J Med. 2021;384:1962–3.
405 https://doi.org/10.1056/NEJMc2102153
406 [14] Centers for Disease Control and Prevention. COVID data tracker. 2021 [cited
407 2021 Oct 6] https://covid.cdc.gov/covid-data-tracker/#datatracker-home
408 [15] Coronavirus Resource Center, Johns Hopkins University. Impact of opening and
409 closing decisions by state: a look at how social-distancing measures may have influenced
410 trends in COVID-19 cases and deaths. 2021 [cited 2021 Oct 6]
411 https://coronavirus.jhu.edu/data/state-timeline
412 [16] Washington NL, Gangavarapu K, Zeller M, Bolze A, Cirulli ET, Schiabor Barrett
413 KM, et al. Emergence and rapid transmission of SARS-CoV-2 B.1.1.7 in the United
414 States. Cell. 2021;184:2587–94. https://doi.org/10.1016/j.cell.2021.03.052
415 [17] Centers for Disease Control and Prevention. Covid data tracker, variant
416 proportions. 2021 [cited 2021 Oct 6] https://covid.cdc.gov/covid-data-tracker/#variant-
417 proportions
418 [18] Volz E, Mishra S, Chand M, Barrett JC, Johnson R, Geidelberg L, et al. Assessing
419 transmissibility of SARS-CoV-2 lineage B.1.1.7 in England. Nature. 2021;593:266–9.
420 https://doi.org/10.1038/s41586-021-03470-x

421 [19] Davies NG, Abbott S, Barnard RC, Jarvis CI, Kucharski AJ, Munday JD, et al.
422 Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England.
423 Science. 2021;372:eabg3055. https://doi.org/10.1126/science.abg3055
424 [20] Ito K, Piantham C, Nishiura H. Predicted dominance of variant Delta of SARS-
425 CoV-2 before Tokyo Olympic Games, Japan, July 2021. Euro Surveill.
426 2021;26:2100570. https://doi.org/10.2807/1560-7917.ES.2021.26.27.2100570
427 [21] Arav Y, Fattal E, Klausner Z. Increased transmissibility of emerging SARS-CoV-
428 2 variants is driven either by viral load or probability of infection rather than
429 environmental stability. Mathematics. 2022, 10(19), 3422.
430 https://doi.org/10.3390/math10193422
431 [22] Public Health England. SARS-CoV-2 variants of concern and variants under
432 investigation in England. Technical briefing 15, June 11, 2021. [cited 2021 Aug 24]
433 https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_
434 data/file/993879/Variants_of_Concern_VOC_Technical_Briefing_15.pdf
435 [23] Cascella M, Rajnik M, Aleem A, Dulebohn SC, Di Napoli R. Features,
436 evaluation, and treatment of coronavirus (COVID-19). July 30, 2021 update. Treasure
437 Island (FL): StatPearls Publishing; 2022.
438 https://www.ncbi.nlm.nih.gov/books/NBK554776/
439 [24] Lopez Bernal J, Andrews N, Gower C, Gallagher E, Simmons R, Thelwall S, et
440 al. Effectiveness of Covid-19 Vaccines against the B.1.617.2 (Delta) variant. N Engl J
441 Med. 2021;385:585–94. https://doi.org/10.1056/NEJMoa2108891
442 [25] Rovida F, Cassaniti I, Paolucci S, Percivalle E, Sarasini A, Piralla A, et al. SARS-
443 CoV-2 vaccine breakthrough infections with the alpha variant are asymptomatic or

444 mildly symptomatic among health care workers. Nat Commun. 2021;12:6032.
445 https://doi.org/10.1038/s41467-021-26154-6
446 [26] Chand, M. et al. Investigation of Novel SARS-COV-2 Variant: Variant of
447 Concern 202012/01. Technical Briefing 1.
448 https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_
449 data/file/959438/Technical_Briefing_VOC_SH_NJL2_SH2.pdf (Public Health England,
450 2020).
451 [27] Andre M, Lau LS, Pokharel MD, Ramelow J, Owens F, Souchak J, et al. From
452 Alpha to Omicron: How Different Variants of Concern of the SARS-Coronavirus-2
453 Impacted the World. Biology. 2023;12:1267. https://doi.org/10.3390/biology12091267
454 [28] Davies NG, Jarvis CI, Edmunds WJ, Jewell NP, Diaz-Ordaz K, Keogh RH.
455 Increased mortality in community-tested cases of SARS-CoV-2 lineage B. 1.1. 7. Nature.
456 2021;593:270-4. https://doi.org/10.1038/s41586-021-03426-1
457 [29] Campbell F, Archer B, Laurenson-Schafer H, Jinnai Y, Konings F, Batra N, et al.
458 Increased transmissibility and global spread of SARS-CoV-2 variants of concern as at
459 June 2021. Euro Surveill. 2021;26:2100509. https://doi.org/10.2807/1560-
460 7917.es.2021.26.24.2100509
461 [30] Sonabend R, Whittles LK, Imai N, Perez-Guzman PN, Knock ES, Rawson T, et
462 al. Non-pharmaceutical interventions, vaccination, and the SARS-CoV-2 delta variant in
463 England: a mathematical modelling study. Lancet. 2021;398:1825–35.
464 https://doi.org/10.1016/s0140-6736(21)02276-5

465 [31] Cai J, Deng X, Yang J, Sun K, Liu H, Chen Z, et al. Modeling transmission of
466 SARS-CoV-2 Omicron in China. Nat Med. 2022;28:1468–75.
467 https://doi.org/10.1038/s41591-022-01855-7
468 [32] Sheikhi F, Yousefian N, Tehranipoor P, Kowsari Z. Estimation of the basic
469 reproduction number of Alpha and Delta variants of COVID-19 pandemic in Iran. PLOS
470 One. 2022;17:e0265489. https://doi.org/10.1371/journal.pone.0265489
471 [33] Barnard RC, Davies NG, Jit M, Edmunds WJ. Modeling the medium-term
472 dynamics of SARS-CoV-2 transmission in England in the Omicron era. Nat Commun.
473 2022;13:4879. https://doi.org/10.1038/s41467-022-32404-y
474 [34] Liossi S, Tsiambas E, Maipas S, Papageorgiou E, Lazaris A, Kavantzas N.
475 Mathematical modeling for Delta and Omicron variant of SARS-CoV-2 transmission
476 dynamics in Greece. Infect Dis Model. 2023;8:794–805.
477 https://doi.org/10.1016/j.idm.2023.07.002
478 [35] Donnelly MA, Chuey MR, Soto R, Schwartz NG, Chu VT, Konkle SL, et al.
479 Household transmission of severe acute respiratory syndrome coronavirus 2 (SARS-
480 CoV-2) Alpha variant–United States, 2021. Clin Infect Dis. 2022;75:e122-32.
481 10.1093/cid/ciac125
482 [36] Tabatabai M, Juarez PD, Matthews-Juarez P, Wilus DM, Ramesh A, Alcendor
483 DJ, et al. An Analysis of COVID-19 Mortality During the Dominancy of Alpha, Delta
484 and Omicron in the USA. J Prim Care Community Health. 2023;14:21501319231170164.
485 10.1177/21501319231170164
486 [37] The New York Times. Coronavirus (COVID-19) data in the United States. 2021
487 [cited 2021 Aug 24] https://github.com/nytimes/covid-19-data

488 [38] Covid Act Now. US COVID risk and vaccine tracker. 2021 [cited 2021 Sep 29]
489 https://covidactnow.org/
490 [39] Democrat and Chronicle. COVID-19 vaccine tracker. 2023 [cited 2023 Nov 11]
491 https://data.democratandchronicle.com/covid-19-vaccine-tracker
492 [40] Neumann J, Lin YT, Mallela A, Miller EF, Colvin J, Duprat AT, Chen Y,
493 Hlavacek WS, Posner RG. Implementation of a practical Markov chain Monte Carlo
494 sampling algorithm in PyBioNetFit. Bioinformatics. 2022;38:1770–2.
495 https://doi.org/10.1093/bioinformatics/btac004
496 [41] GitHub repository of supplemental materials for this study. 2023 [cited 2023 Dec
497 13] https://github.com/lanl/PyBNF/tree/master/examples/Vax_and_Variants
498 [42] Vats D, Knudson C. Revisiting the Gelman-Rubin Diagnostic. Statist. Sci.
499 2021;36:518–529. https://doi.org/10.1214/20-STS812
500 [43] Centers for Disease Control and Prevention. SARS-CoV-2 Variant Proportions.
501 2023 [cited 2023 Mar 8] https://data.cdc.gov/Laboratory-Surveillance/SARS-CoV-2-
502 Variant-Proportions/jr58-6ysp
503 [44] Michaelsen TY, Bennedbæk M, Christiansen LE, Jørgensen MS, Møller CH,
504 Sørensen EA, et al. Introduction and transmission of SARS-CoV-2 lineage B. 1.1. 7,
505 Alpha variant, in Denmark. Genome Med. 2022;14:47. https://doi.org/10.1186/s13073-
506 022-01045-7
507 [45] Zárate S, Taboada B, Muñoz-Medina JE, Iša P, Sanchez-Flores A, Boukadida C,
508 et al. The alpha variant (B. 1.1. 7) of SARS-CoV-2 failed to become dominant in Mexico.
509 Microbiol Spectr. 2022;10:e02240-21. https://doi.org/10.1128/spectrum.02240-21

510 [46] Earnest R, Uddin R, Matluk N, Renzette N, Turbett SE, Siddle KJ, et al.
511 Comparative transmissibility of SARS-CoV-2 variants delta and alpha in New England,
512 USA. Cell Rep Med. 2022;3. https://doi.org/10.1016/j.xcrm.2022.100583
513 [47] Alexander TY, Hughes B, Wolfe MK, Leon T, Duong D, Rabe A, et al.
514 Estimating relative abundance of 2 SARS-CoV-2 variants through wastewater
515 surveillance at 2 large metropolitan sites, United States. Emerg Infect Dis. 2022;28:940.
516 10.3201/eid2805.212488
517 [48] Bolze A, Luo S, White S, Cirulli ET, Wyman D, Dei Rossi A, et al. SARS-CoV-2
518 variant Delta rapidly displaced variant Alpha in the United States and led to higher viral
519 loads. Cell Rep Med. 2022;3. 10.1016/j.xcrm.2022.100564
520 [49] Stein C, Nassereldine H, Sorensen RJ, Amlag JO, Bisignano C, Byrne S, et al.
521 Past SARS-CoV-2 infection protection against re-infection: a systematic review and
522 meta-analysis. Lancet. 2023;401:833–42. https://doi.org/10.1016/S0140-6736(22)02465-
523 5
524 [50] Tregoning JS, Flight KE, Higham SL, Wang Z, Pierce BF. Progress of the
525 COVID-19 vaccine effort: viruses, vaccines and variants versus efficacy, effectiveness
526 and escape. Nat Rev Immunol. 2021;21:626-36. https://doi.org/10.1038/s41577-021-
527 00592-1

528

Table 1. Model parameter values inferred for the Dallas, Houston, New York City, and Phoenix Metropolitan Statistical Areas (MSAs) on October 31, 2021

Parameter	MAP estimate* for Dallas (Units)	MAP estimate* for Houston (Units)	MAP estimate* for New York City (Units)	MAP estimate* for Phoenix (Units)
t_0	14.8 (d)	11.2 (d)	9.6 (d)	14.0 (d)
β	0.30 (d^{-1})	0.29 (d^{-1})	0.37 (d^{-1})	0.29 (d^{-1})
σ	65 (d)	69 (d)	68 (d)	62 (d)
p_0	0.32	0.29	0.48	0.26
λ_0	9.7 (d^{-1})	7.5 (d^{-1})	9.7 (d^{-1})	6.6 (d^{-1})
τ_1	139 (d)	162 (d)	143 (d)	148 (d)
p_1	0.27	0.39	0.40	0.23
λ_1	2.8 (d^{-1})	0.1 (d^{-1})	2.7 (d^{-1})	2.5 (d^{-1})
τ_2	165 (d)	186 (d)	177 (d)	169 (d)
p_2	0.38	0.34	0.41	0.41
λ_2	0.63 (d^{-1})	0.53 (d^{-1})	0.39 (d^{-1})	0.47 (d^{-1})
τ_3	259 (d)	278 (d)	262 (d)	248 (d)
p_3	0.22	0.17	0.26	0.10
λ_3	5.1 (d^{-1})	3.2 (d^{-1})	8.1 (d^{-1})	5.2 (d^{-1})
τ_4	264 (d)	-	-	255 (d)
p_4	0.24	-	-	0.15
λ_4	1.7 (d^{-1})	-	-	3.2 (d^{-1})

θ_1	377 (d)	400 (d)	381 (d)	404 (d)
y_1	1.1	1.2	1.4	1.3
θ_2	497 (d)	521 (d)	512 (d)	522 (d)
y_2	1.6	2.0	2.1	1.8
f_D	0.48	0.36	0.39	0.54
r	3.0	3.0	9.1	3.1

*Maximum *a posteriori* (MAP) estimates are region-specific and inference-time-dependent. Here, inference was based on MSA-specific confirmed coronavirus disease case count data available in the GitHub repository maintained by *The New York Times* newspaper [37] for January 21, 2020 to October 31, 2021. Time $t = 0$ corresponds to midnight on January 21, 2020. Inferences were conditioned on the compartmental model of Appendix Equations 1–38, consideration of two viral variants (Alpha and Delta, $m = 2$), and $n + 1$ distinct social-distancing periods in total ($n = 4$ for Dallas and Phoenix; $n = 3$ for New York City and Houston), the fixed parameter estimates of Table 3, and the initial condition I_0 and S_0 at time $t = t_0$ given in Table 3. The choice of two variants and the setting for n were chosen through a model-selection procedure described in the Appendix. With $m = 2$ and $n = 3$, there are 18 adjustable model parameters: t_0, β, θ_1, y_1, θ_2, y_2, σ, p_0, λ_0, τ_1, p_1, λ_1, τ_2, p_2, λ_2, τ_3, p_3, and λ_3. With $m = 2$ and $n = 4$, there are 21 adjustable model parameters: t_0, β, θ_1, y_1, θ_2, y_2, σ, p_0, λ_0, τ_1, p_1, λ_1, τ_2, p_2, λ_2, τ_3, p_3, λ_3, τ_4, p_4, and λ_4. These parameters were jointly inferred together with f_D, the parameter of the measurement model (i.e., the fraction of new cases detected and reported) (Appendix Equation 40), and r, the dispersion parameter of the statistical model for noise in case detection and reporting (i.e., the adjustable parameter of the negative binomial

likelihood function) (Appendix Equations 41–43). We assumed a uniform proper prior, as described in the Appendix.

Table 2. Descriptions of MSA-specific adjustable model parameters.

Parameter	Description
t_0	Start of local disease transmission
β	Rate constant for disease transmission
σ	Start of first social-distancing period
p_0	Social-distancing setpoint for the period $t \in [\sigma, \tau_1)$
λ_0	Social-distancing eigenvalue paired with p_0
τ_1	Start of second social-distancing period
p_1	Social-distancing setpoint for the period $t \in [\tau_1, \tau_2)$
λ_1	Social-distancing eigenvalue paired with p_1
τ_2	Start of third social-distancing period
p_2	Social-distancing setpoint for the period $t \in [\tau_2, \tau_3)$
λ_2	Social-distancing eigenvalue paired with p_2
τ_3	Start of fourth social-distancing period
p_3	Social-distancing setpoint for the period $t \in [\tau_3, \tau_4)$
λ_3	Social-distancing eigenvalue paired with p_3
τ_4	Start of fifth social-distancing period
p_4	Social-distancing setpoint for the period $t \in [\tau_4, \infty)$
λ_4	Social-distancing eigenvalue paired with p_4
θ_1	Alpha takeoff time
y_1	Increased infectiousness of Alpha (relative to ancestral strains)

θ_2	Delta takeoff time
y_2	Increased infectiousness of Delta (relative to ancestral strains)
f_D	Fraction of cases detected and reported
r	Dispersion parameter of $\mathrm{NB}(r, q_i)$*

*The probability parameter of $\mathrm{NB}(r, q_i)$ is constrained, i.e., the value of q_i, which is reporting time-dependent, is given by Appendix Equation 43.

570 **Table 3.** Fixed parameter estimates for each region-specific compartmental model

Parameter	Estimate[1] (Units)	Description	Source
I_0	1	Number of infectious persons at time $t = t_0$	[5]
S_0	19,216,182[2]	Total population	[5]
$\mu(t)$	Empirical time-series[3] (d^{-1})	Daily *per capita* rate of vaccination	[38, 39]
m_b	0.1	Reduction in risk of infection because of social-distancing	[5]
m_h	0.04	Reduction in risk of severe disease (once symptomatic) because of vaccination	This study[4]
f_A	0.44	Fraction of all cases that are asymptomatic	[5]
f_H	0.054	Fraction of symptomatic cases that are severe (in the absence of vaccination)	[5]
f_R	0.79	Fraction of persons with severe disease who recover	[5]
$1 - f_0$	0.1	Fraction of vaccinated persons who fail to develop an immune response that protects against productive infection by ancestral strains or variants	This study[4]
$f_0 - f_1$	0.09	Fraction of vaccinated persons who develop an immune response that protects against	This study[4]

		productive infection by ancestral strains (but not the Alpha or Delta variant)	
$f_1 - f_2$	0.12	Fraction of vaccinated persons who develop an immune response that protects against productive infection by ancestral strains and the Alpha variant (but not the Delta variant)	This study[4]
f_2	0.69	Fraction of vaccinated persons who develop an immune response that protects against productive infection by ancestral strains and the Alpha and Delta variants	This study[4]
ρ_E	1.1	Relative infectiousness of persons without symptoms in the incubation period of infection	[5]
ρ_A	0.9	Relative infectiousness of persons without symptoms in the immune-clearance phase of infection	[5]
k_L	0.94 (d^{-1})	Rate constant for progression through each stage of the incubation period of infection[5]	[5]
k_Q	0.0038 (d^{-1})	Rate constant for quarantine of infected, non-vaccinated persons	[5]
j_Q	0.4 (d^{-1})	Rate constant for self-isolation of symptomatic, non-vaccinated persons	[5]

c_A	0.26 (d^{-1})	Rate constant for completion of the immune clearance phase of infection for persons without symptoms	[5]
c_I	0.12 (d^{-1})	Rate constant for completion of the immune clearance phase of infection or progression to severe disease for non-vaccinated persons with symptoms	[5]
c_H	0.17 (d^{-1})	Rate constant for recovery or progression to death for non-vaccinated persons with severe disease	[5]
k_V	0.3 (d^{-1})	Rate constant for progression through each stage of immune response to vaccination[6]	This study[4]

[1]Fixed parameter estimates are based on information external to the surveillance data used to infer the adjustable parameter values of Table 1. Estimates are applicable to all metropolitan statistical areas (MSAs) of interest except for S_0 and the vaccination function $\mu(t)$, which are region-specific. Recall that $\mu(t)$ is updated daily for consistency with vaccination data.

[2]The total population S_0 is MSA-specific. Here, S_0 is given for the New York City MSA. The initial susceptible population is taken to be the total population. We used $S_0 = 7{,}573{,}136$, 7,066,141, and 4,873,019 for the Dallas, Houston, and Phoenix MSAs, respectively.

[3]The function $\mu(t)$ is determined by region-specific vaccination data [38, 39].

[4]See the Appendix for more information about estimates of fixed parameter values made in this study.

[5]As in the study of Lin et al. [5], the incubation period is divided into 5 stages (Figure 1), each of equal duration on average.

[6]The immune response to vaccination is divided into 6 stages (Figure 1), each of equal duration on average. The choice of 6 stages is justified in the Appendix and Appendix Figure 2.

586 **Table 4.** Multivariate potential scale reduction factors (PSRFs) and multivariate effective sample
587 sizes (ESSs) corresponding to the stable Gelman-Rubin statistic for Markov chains generated
588 using PyBioNetFit for the Dallas, Houston, New York City, and Phoenix Metropolitan Statistical
589 Areas (MSAs)

MSA	Multivariate PSRF[1]	Multivariate ESS[1]
Dallas	1.0024	211.6
Houston	1.0026	193.2
New York City	1.0023	212.5
Phoenix	1.0033	152.9

590 [1]Computed using the methodology of Vats and Knudson [42].

591

592

593

594

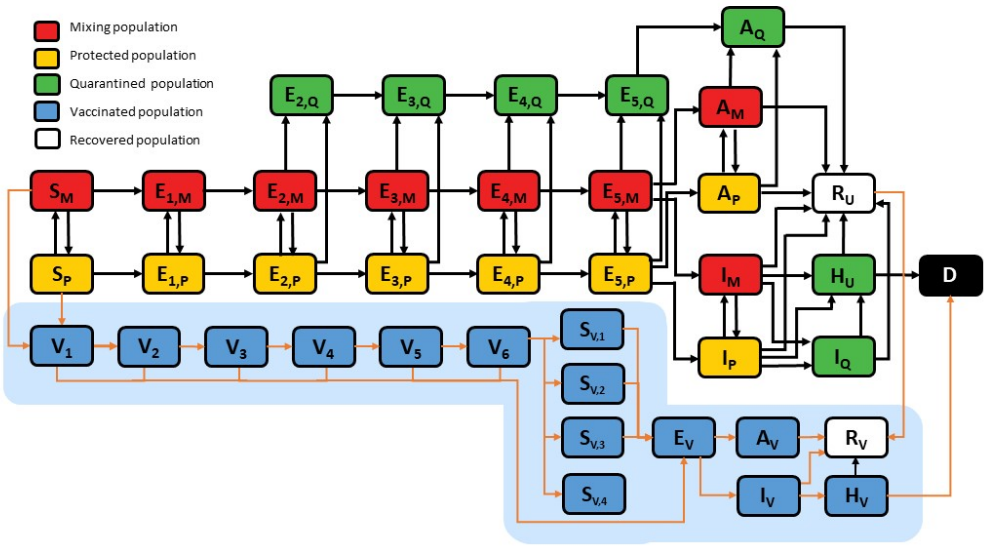
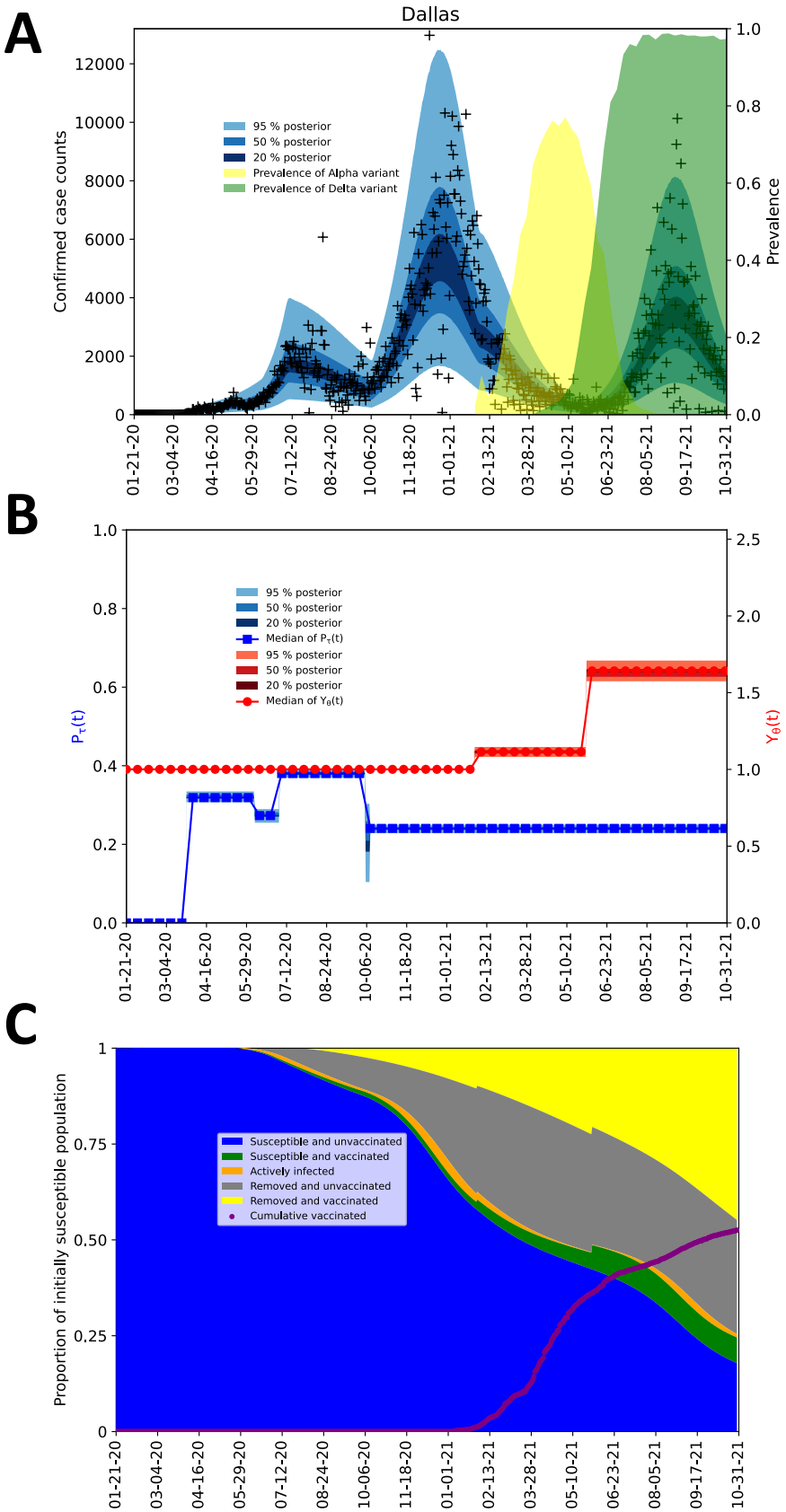


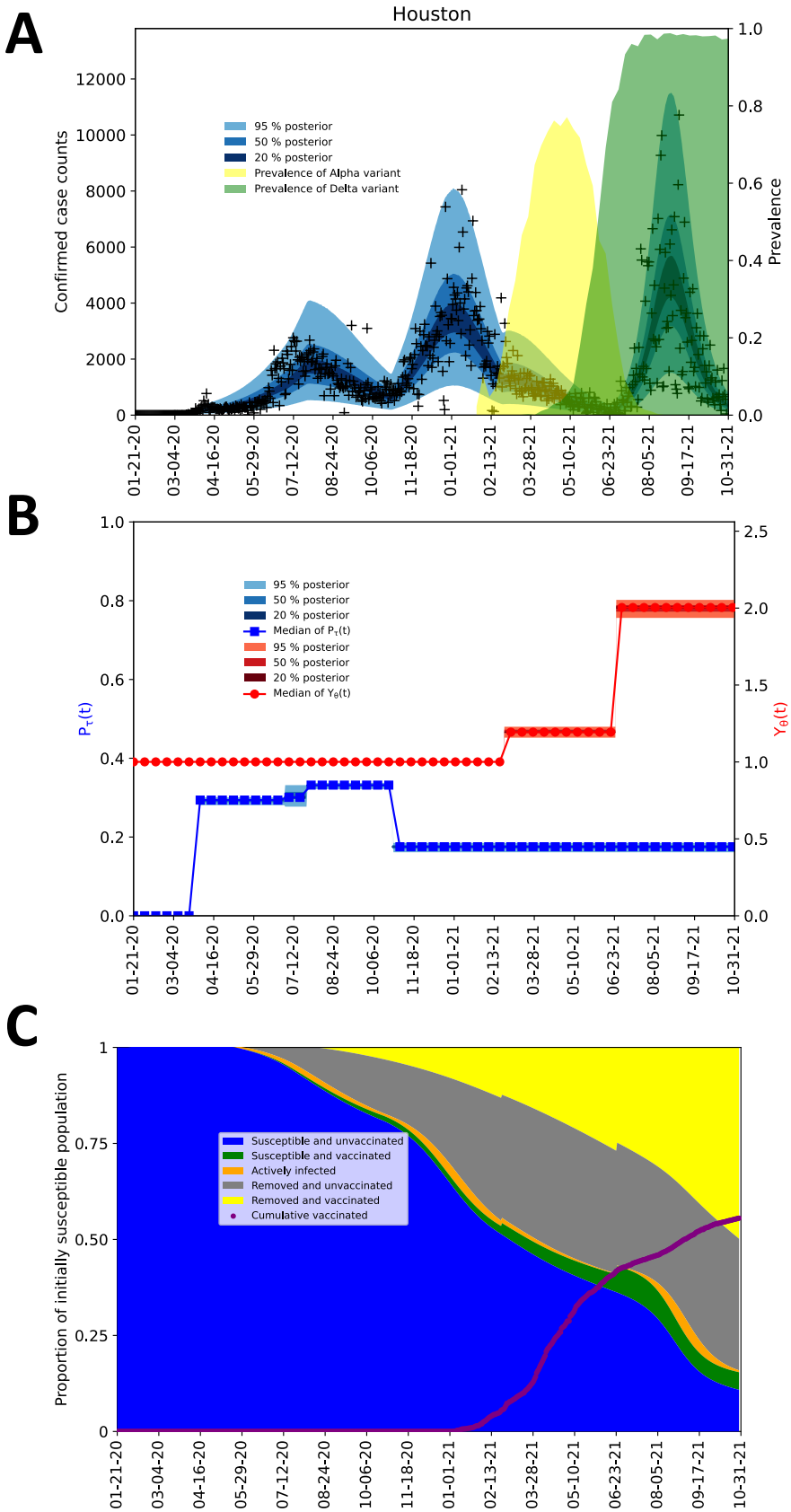
Figure 1. Illustration of compartmental model. The independent variable of the model is time t. The 40 dependent state variables of the model are populations, which are represented as boxes with rounded corners. A description of each state variable is given in Appendix Table 1. The 15 highlighted boxes (on the blue background) represent state variables introduced to capture the effects of vaccination and the Alpha and Delta variants. The other 25 boxes represent state variables considered in the model of Lin et al. [5]. Arrows connecting boxes represent transitions. Each transition represents the movement of persons from one population to another. The arrows highlighted in orange represent transitions introduced to capture the effects of vaccination and the Alpha and Delta variants. Other arrows represent transitions considered in the model of Lin et al. [5]. Each arrow is associated with one or more parameters that characterize a rate of movement; these parameters are not shown here but are shown in Appendix Figure 1. A full description of the model is given in the Appendix.


A
Dallas
Confirmed case counts
Prevalence
95 % posterior
50 % posterior
20 % posterior
Prevalence of Alpha variant
Prevalence of Delta variant
B
95 % posterior
50 % posterior
20 % posterior
Median of $P_\tau(t)$
95 % posterior
50 % posterior
20 % posterior
Median of $Y_\theta(t)$
$P_\tau(t)$
$Y_\theta(t)$
C
Proportion of initially susceptible population
Susceptible and unvaccinated
Susceptible and vaccinated
Actively infected
Removed and unvaccinated
Removed and vaccinated
Cumulative vaccinated
01-21-20
03-04-20
04-16-20
05-29-20
07-12-20
08-24-20
10-06-20
11-18-20
01-01-21
02-13-21
03-28-21
05-10-21
06-23-21
08-05-21
09-17-21
10-31-21

610

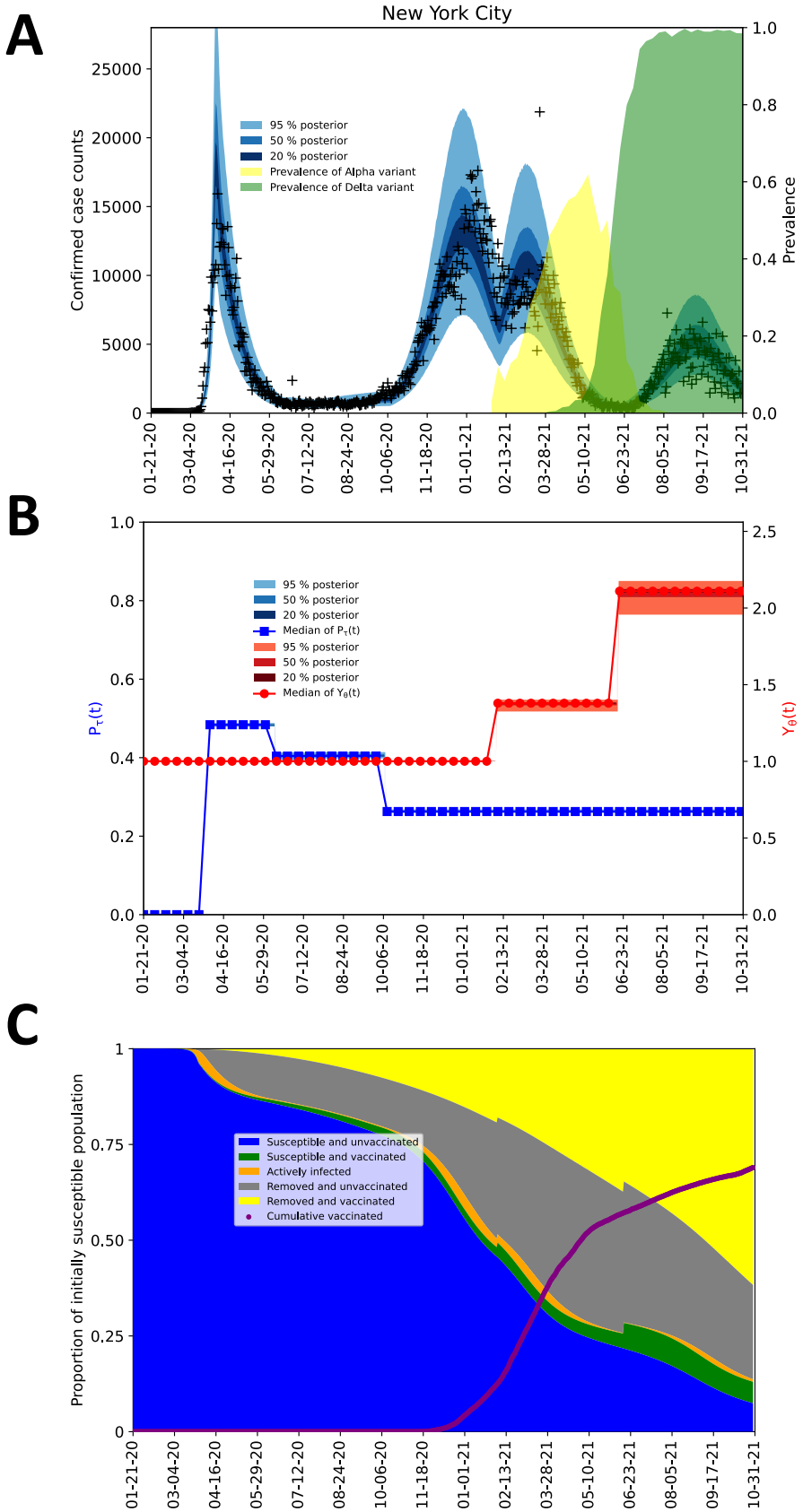
Figure 2. Inference results obtained for the MSA surrounding Dallas using regional surveillance data—daily reports of new COVID-19 cases—available for January 21, 2020 to October 31, 2021. A) Credible intervals of the time-dependent posterior predictive distribution for detected and reported new cases are shown. The monochrome bands from innermost to outermost indicate the 20%, 50%, and 95% credible intervals. This colored region all together indicates the 95% credible interval and can be expected to cover approximately 95% of the data. Empirical case reports are indicated by black symbols. Alpha prevalence is indicated with a shaded light-yellow background. Delta prevalence is indicated with a shaded light-green background. B) The social-distancing stationary setpoint is given by $P_\tau(t)$ and the variant transmissibility factor is given by $Y_\theta(t)$. Note that the values of $P_\tau(t)$ and $Y_\theta(t)$ are dimensionless. Credible intervals corresponding to 1000 samples from the time-dependent posterior predictive distributions are shown for $P_\tau(t)$ and $Y_\theta(t)$. The curve corresponding to $Y_\theta(t)$ is monotonically increasing with an initial value of 1. The curve corresponding to $P_\tau(t)$ is decreasing from left to right after the start of social-distancing. C) Inferred changes in the distribution of persons amongst five selected subpopulations over the course of the local COVID-19 epidemic. The five populations sum to a constant, S_0, the total population. Results shown here are based on the parameter values given in Tables 1 and 3. The five populations are defined as follows: the population of susceptible unvaccinated persons (blue area) is given by $S_M + S_P$, the population of susceptible vaccinated persons (green area) is given by $\sum_{i=1}^{6} V_i + S_{V,1} + U_{\theta_1}(t)S_{V,2} + U_{\theta_2}(t)S_{V,3}$, the population of actively infected persons (orange area) is given by $H_U + H_V + E_V + \sum_{i=1}^{5}\left(E_{i,M} + E_{i,P} + E_{i,Q}\right) + \sum_{x\in\{M,P,Q,V\}}(A_x + I_x)$, the population of removed unvaccinated persons (gray area) is given by $R_U + D$, and the population of removed vaccinated persons (yellow area) is given by $R_V + (1 - U_{\theta_1}(t))S_{V,2} + (1 - U_{\theta_2}(t))S_{V,3} + S_{V,4}$. Except for $U_{\theta_1}(t)$ and $U_{\theta_2}(t)$, the terms in the

above definitions refer to state variables of the compartmental model of Figure 1 and Appendix Figure 1, which are defined in Appendix Table 1. $U_{\theta_1}(t)$ and $U_{\theta_2}(t)$ are unit step functions (Appendix Equation 33 and 34), which change value from 0 to 1 at time $t = \theta_1$ and $t = \theta_2$, respectively. Recall that θ_1 and θ_2 are the Alpha and Delta takeoff times. The sum $\sum_i(\mu_i \times 1\text{ d})$ (purple dots), which is the empirical cumulative number of completed vaccinations, is shown as a function of time t.


A
Houston
Confirmed case counts
Prevalence
95 % posterior
50 % posterior
20 % posterior
Prevalence of Alpha variant
Prevalence of Delta variant
B
$P_\tau(t)$
$Y_\theta(t)$
Median of $P_\tau(t)$
Median of $Y_\theta(t)$
C
Proportion of initially susceptible population
Susceptible and unvaccinated
Susceptible and vaccinated
Actively infected
Removed and unvaccinated
Removed and vaccinated
Cumulative vaccinated
01-21-20
03-04-20
04-16-20
05-29-20
07-12-20
08-24-20
10-06-20
11-18-20
01-01-21
02-13-21
03-28-21
05-10-21
06-23-21
08-05-21
09-17-21
10-31-21

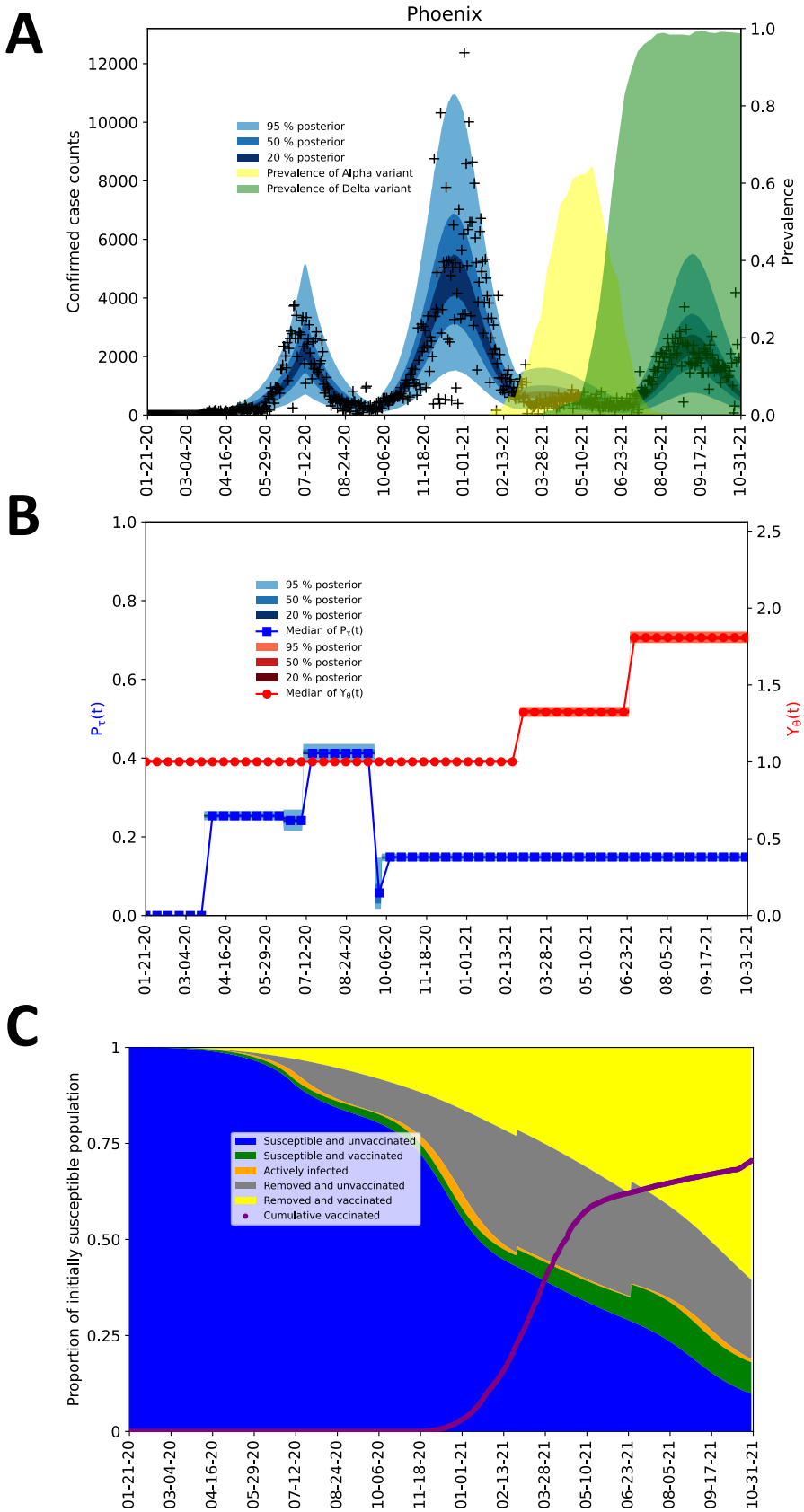
641

Figure 3. Inference results obtained for the MSA surrounding Houston using regional surveillance data—daily reports of new COVID-19 cases—available for January 21, 2020 to October 31, 2021. It should be noted that four anomalous (negative) empirical case counts are not shown in the plot. A case count of 14,300 cases on September 21, 2020 is not shown in the plot. See the caption of Figure 2 for additional information.


A
New York City
Confirmed case counts
Prevalence
95 % posterior
50 % posterior
20 % posterior
Prevalence of Alpha variant
Prevalence of Delta variant
B
$P_\tau(t)$
$Y_\theta(t)$
Median of $P_\tau(t)$
Median of $Y_\theta(t)$
C
Proportion of initially susceptible population
Susceptible and unvaccinated
Susceptible and vaccinated
Actively infected
Removed and unvaccinated
Removed and vaccinated
Cumulative vaccinated
01-21-20
03-04-20
04-16-20
05-29-20
07-12-20
08-24-20
10-06-20
11-18-20
01-01-21
02-13-21
03-28-21
05-10-21
06-23-21
08-05-21
09-17-21
10-31-21

648

Figure 4. Inference results obtained for the MSA surrounding New York City using regional surveillance data—daily reports of new COVID-19 cases—available for January 21, 2020 to October 31, 2021. It should be noted that a single anomalous (negative) empirical case count is not shown in the plot. See the caption of Figure 2 for additional information.


A
Phoenix
Confirmed case counts
Prevalence
95 % posterior
50 % posterior
20 % posterior
Prevalence of Alpha variant
Prevalence of Delta variant
B
$P_t(t)$
$Y_\theta(t)$
Median of $P_t(t)$
Median of $Y_\theta(t)$
C
Proportion of initially susceptible population
Susceptible and unvaccinated
Susceptible and vaccinated
Actively infected
Removed and unvaccinated
Removed and vaccinated
Cumulative vaccinated
01-21-20
03-04-20
04-16-20
05-29-20
07-12-20
08-24-20
10-06-20
11-18-20
01-01-21
02-13-21
03-28-21
05-10-21
06-23-21
08-05-21
09-17-21
10-31-21

662

Figure 5. Inference results obtained for the MSA surrounding Phoenix using regional surveillance data—daily reports of new COVID-19 cases—available for January 21, 2020 to October 31, 2021. It should be noted that two anomalous (negative) empirical case counts are not shown in the plot. See the caption of Figure 2 for additional information.

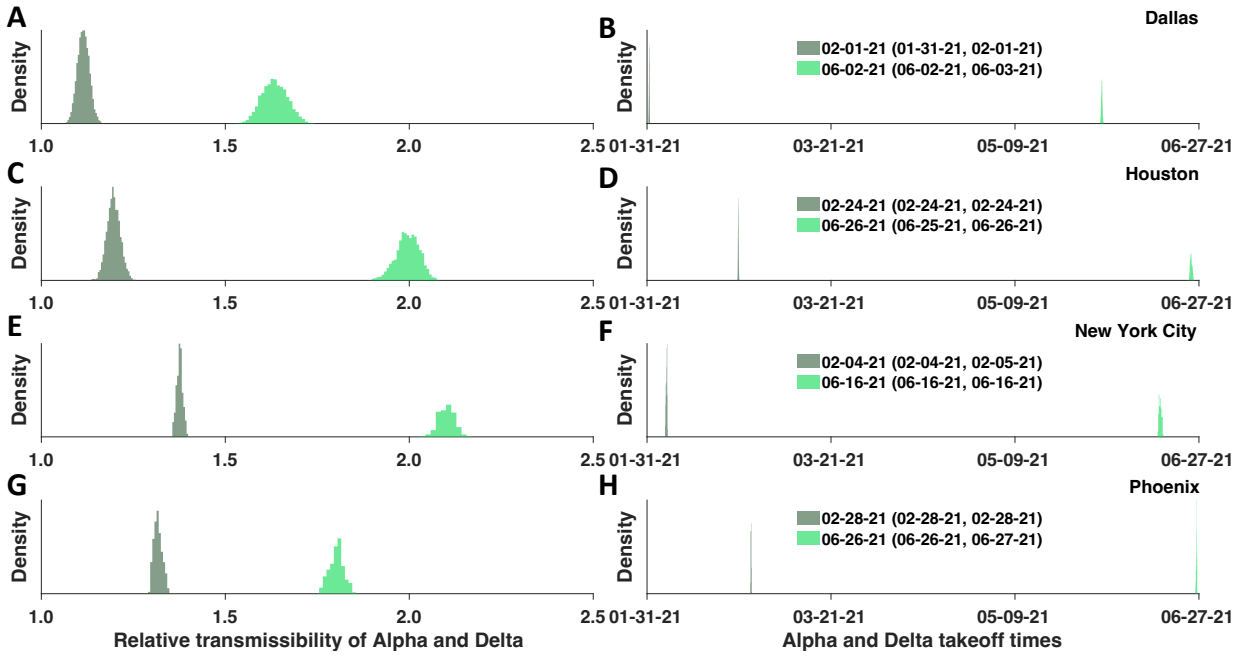


Figure 6. Marginal posterior distributions of transmissibility factors for Alpha in dark-green and Delta in light-green (left panels) and takeoff times for Alpha and Delta (right panels) in four MSAs centered on (**A**,**B**) Dallas, (**C**,**D**) Houston, (**E**,**F**) New York City, and (**G**,**H**) Phoenix. Inferences are based on daily reports of new cases from January 21, 2020 to October 31, 2021. For each of the right panels, the 95% credible intervals for Alpha and Delta takeoff times are indicated in parentheses, and the MAP estimate for a given region is indicated to the left of the credible interval.

APPENDIX

Imputation of Missing Daily Case Counts

By October 31, 2021, many regions in the US were not reporting new detected COVID-19 cases on a strictly daily basis. When one or more daily case counts were not available, we imputed daily case counts on the basis of a linear fit to the two nearest available cumulative case

counts. This approach has the effect of evenly distributing case counts across the days for which daily reports are unavailable.

Equations of the Compartmental Model

The compartmental model, which is illustrated in Figure 1 and Appendix Figure 1, consists of the following 40 ordinary differential equations (ODEs):

$$\frac{dS_M}{dt} = -\beta Y_\theta(t)\left(\frac{S_M}{S_0}\right)\left(\phi_M(t,\rho) + m_b\phi_P(t,\rho)\right) - U_\sigma(t)\Lambda_\tau(t)\left[P_\tau(t)S_M - \left(1 - P_\tau(t)\right)S_P\right] - \mu(t)S_0\left(\frac{S_M}{\phi_V(t)}\right) \tag{1}$$

$$\frac{dS_P}{dt} = -m_b\beta Y_\theta(t)\left(\frac{S_P}{S_0}\right)\left(\phi_M(t,\rho) + m_b\phi_P(t,\rho)\right) + U_\sigma(t)\Lambda_\tau(t)\left[P_\tau(t)S_M - \left(1 - P_\tau(t)\right)S_P\right] - \mu(t)S_0\left(\frac{S_P}{\phi_V(t)}\right) \tag{2}$$

$$\frac{dE_{1,M}}{dt} = \beta Y_\theta(t)\left(\frac{S_M}{S_0}\right)(\phi_M(t,\rho) + m_b\phi_P(t,\rho)) - k_L E_{1,M} - U_\sigma(t)\Lambda_\tau(t)\left[P_\tau(t)E_{1,M} - (1 - P_\tau(t))E_{1,P}\right] \tag{3}$$

$$\frac{dE_{1,P}}{dt} = m_b\beta Y_\theta(t)\left(\frac{S_P}{S_0}\right)(\phi_M(t,\rho) + m_b\phi_P(t,\rho)) - k_L E_{1,P} + U_\sigma(t)\Lambda_\tau(t)\left[P_\tau(t)E_{1,M} - (1 - P_\tau(t))E_{1,P}\right] \tag{4}$$

$$\frac{dE_{i,M}}{dt} = k_L E_{i-1,M} - k_L E_{i,M} - k_Q E_{i,M} - U_\sigma(t)\Lambda_\tau(t)\left[P_\tau(t)E_{i,M} - (1 - P_\tau(t))E_{i,P}\right], \tag{5}$$

for $i = 2, 3, 4, 5$

$$\frac{dE_{i,P}}{dt} = k_L E_{i-1,P} - k_L E_{i,P} - k_Q E_{i,P} + U_\sigma(t)\Lambda_\tau(t)\left[P_\tau(t)E_{i,M} - (1 - P_\tau(t))E_{i,P}\right], \tag{6}$$

for $i = 2, 3, 4, 5$

$$\frac{dE_{2,Q}}{dt} = k_Q(E_{2,M} + E_{2,P}) - k_L E_{2,Q} \tag{7}$$

$$\frac{dE_{i,Q}}{dt} = k_Q\left(E_{i,M} + E_{i,P}\right) + k_L E_{i-1,Q} - k_L E_{i,Q}, \text{ for } i = 3, 4, 5 \tag{8}$$

$$\frac{dA_M}{dt} = f_A k_L E_{5,M} - k_Q A_M - U_\sigma(t)\Lambda_\tau(t)[P_\tau(t)A_M - (1 - P_\tau(t))A_P] - c_A A_M \tag{9}$$

$$\frac{dA_P}{dt} = f_A k_L E_{5,P} - k_Q A_P + U_\sigma(t)\Lambda_\tau(t)[P_\tau(t)A_M - (1 - P_\tau(t))A_P] - c_A A_P \tag{10}$$

$$\frac{dA_Q}{dt} = f_A k_L E_{5,Q} + k_Q(A_M + A_P) - c_A A_Q \tag{11}$$

$$\frac{dI_M}{dt} = (1 - f_A)k_L E_{5,M} - \left(k_Q + j_Q\right)I_M - U_\sigma(t)\Lambda_\tau(t)\left[P_\tau(t)I_M - \left(1 - P_\tau(t)\right)I_P\right] - c_I I_M \tag{12}$$

$$\frac{dI_P}{dt} = (1 - f_A)k_L E_{5,P} - \left(k_Q + j_Q\right)I_P + U_\sigma(t)\Lambda_\tau(t)[P_\tau(t)I_M - \left(1 - P_\tau(t)\right)I_P] - c_I I_P \tag{13}$$

$$\frac{dI_Q}{dt} = (1 - f_A)k_L E_{5,Q} + \left(k_Q + j_Q\right)(I_M + I_P) - c_I I_Q \tag{14}$$

$$\frac{dH_U}{dt} = f_H c_I\left(I_M + I_P + I_Q\right) - c_H H_U \tag{15}$$

$$\frac{dD}{dt} = (1 - f_R)c_H H_U + (1 - f_R)c_H H_V \tag{16}$$

$$\frac{dR_U}{dt} = c_A\left(A_M + A_P + A_Q\right) + (1 - f_H)c_I\left(I_M + I_P + I_Q\right) + f_R c_H H_U - \mu(t)S_0\left(\frac{R_U}{\phi_V(t)}\right) \quad (17)$$

$$\frac{dV_1}{dt} = \mu(t)S_0\left(\frac{S_M + S_P}{\phi_V(t)}\right) - k_V V_1 - \beta Y_\theta(t)\left(\frac{V_1}{S_0}\right)\left(\phi_M(t,\rho) + m_b\phi_P(t,\rho)\right) \quad (18)$$

$$\frac{dV_i}{dt} = k_V V_{i-1} - k_V V_i - \beta Y_\theta(t)\left(\frac{V_i}{S_0}\right)\left(\phi_M(t,\rho) + m_b\phi_P(t,\rho)\right), \quad (19)$$

$$\text{for } i = 2, 3, 4, 5, 6$$

$$\frac{dS_{V,1}}{dt} = (1 - f_0)k_V V_6 - \beta Y_\theta(t)\left(\frac{S_{V,1}}{S_0}\right)\left(\phi_M(t,\rho) + m_b\phi_P(t,\rho)\right) \quad (20)$$

$$\frac{dS_{V,2}}{dt} = (f_0 - f_1)k_V V_6 - U_{\theta_1}(t)\beta Y_\theta(t)\left(\frac{S_{V,2}}{S_0}\right)\left(\phi_M(t,\rho) + m_b\phi_P(t,\rho)\right) \quad (21)$$

$$\frac{dS_{V,3}}{dt} = (f_1 - f_2)k_V V_6 - U_{\theta_2}(t)\beta Y_\theta(t)\left(\frac{S_{V,3}}{S_0}\right)\left(\phi_M(t,\rho) + m_b\phi_P(t,\rho)\right) \quad (22)$$

$$\frac{dS_{V,4}}{dt} = f_2 k_V V_6 \quad (23)$$

$$\frac{dE_V}{dt} = \beta Y_\theta(t)\left(\frac{1}{S_0}\right)\left(\phi_M(t,\rho) + m_b\phi_P(t,\rho)\right) \quad (24)$$

$$\cdot \sum_{i=1}^{6}\left(V_i + S_{V,1} + U_{\theta_1}(t)S_{V,2} + U_{\theta_2}(t)S_{V,3}\right) - \left(\frac{k_L}{5}\right)E_V$$

$$\frac{dA_V}{dt} = f_A\left(\frac{k_L}{5}\right)E_V - c_A A_V \quad (25)$$

$$\frac{dI_V}{dt} = (1 - f_A)\left(\frac{k_L}{5}\right)E_V - c_I I_V \quad (26)$$

$$\frac{dH_V}{dt} = m_h f_H c_I I_V - c_H H_V \tag{27}$$

$$\frac{dR_V}{dt} = \mu(t) S_0 \left(\frac{R_U}{\phi_V(t)}\right) + c_A A_V + (1 - m_h f_H) c_I I_V + f_R c_H H_V \tag{28}$$

In these equations, the independent variable is time t, and the state variables (S_M, S_P, $E_{1,M}, \ldots, E_{5,M}$, $E_{1,P}, \ldots, E_{5,P}$, $E_{2,Q}, \ldots, E_{5,Q}$, A_M, A_P, A_Q, I_M, I_P, I_Q, H_U, D, R_U, $V_1, \ldots, V_6$, $S_{V,1}, \ldots, S_{V,4}$, E_V, A_V, I_V, H_V, and R_V) represent 40 (sub)populations (Appendix Table 1), which change over time. Thus, each ODE in Equations (1)–(28) defines the time-rate of change of a population, i.e., the time-rate of change of a state variable. Note that Equations (5), (6), (8) and (19) define 4, 4, 3, and 5 ODEs of the model, respectively. The model is formulated such that S_0, the total population, is a constant. Thus, the model does not account for birth, death for reasons other than COVID-19, immigration, or emigration.

The initial condition associated with Equations (1)–(28) is taken to be $S_M(t_0) = S_0$, $I_M(t_0) = I_0 = 1$, and all other populations (S_P, $E_{1,M}, \ldots, E_{5,M}$, $E_{1,P}, \ldots, E_{5,P}$, $E_{2,Q}, \ldots, E_{5,Q}$, A_M, A_P, A_Q, I_P, I_Q, H_U, D, R_U, $V_1, \ldots, V_6$, $S_{V,1}, \ldots, S_{V,4}$, E_V, A_V, I_V, H_V, and R_V) are equal to 0. Recall that the parameter S_0 denotes the total region-specific population size. Thus, we assume that the entire population is susceptible at the start of the local epidemic at time $t = t_0 > 0$, where time $t = 0$ corresponds to 00:00 hours (midnight) on January 21, 2020. The parameter I_0 denotes the number of infectious symptomatic persons at the start of the regional epidemic.

In the model, the parameters β, k_L, k_Q, j_Q, c_A, c_I, c_H, and k_V are positive-valued rate constants (all with units of d^{-1}), and the parameters m_b, m_h, f_A, f_H, f_R, $f_0 \geq f_1$, $f_1 \geq f_2$, and f_2 are (dimensionless) fractions. Brief definitions of parameters are given in Tables 1 and 2.

In the model, the quantities $\phi_M(t,\rho)$, $\phi_P(t,\rho)$, and $\phi_V(t)$ are functions of (time-dependent) state variables (as defined below), which represent the population of infectious persons who are mixing freely (i.e., not practicing social-distancing), the population of infectious persons who are practicing social-distancing (i.e., adopting disease-avoiding behaviors), and the population of persons eligible for vaccination, respectively. The quantities $\phi_M(t,\rho)$ and $\phi_P(t,\rho)$ are also functions of $\rho \equiv (\rho_E, \rho_A)$, where ρ_E (ρ_A) is a dimensionless ratio representing the infectiousness of persons in the incubation phase of infection (the infectiousness of asymptomatic persons in the immune clearance phase of infection) relative to the infectiousness of symptomatic persons with the same social-distancing behavior. The quantity $\phi_V(t)$ represents the population of persons eligible for vaccination.

$$\phi_M(t,\rho) = I_M + I_V + \rho_E\left(E_{2,M} + E_{3,M} + E_{4,M} + E_{5,M} + E_V\right) + \rho_A(A_M + A_V) \tag{29}$$

$$\phi_P(t,\rho) = I_P + \rho_E\left(E_{2,P} + E_{3,P} + E_{4,P} + E_{5,P}\right) + \rho_A A_P \tag{30}$$

$$\phi_V(t) = S_M + S_P + \sum_{i=1}^{5}\left(E_{i,M} + E_{i,P}\right) + A_M + A_P + R_U \tag{31}$$

The state variables that appear in these equations represent time-varying populations. Recall that state variables are defined in Appendix Table 1.

In the model, the quantities $U_\sigma(t)$, $U_{\theta_1}(t)$, and $U_{\theta_2}(t)$ are unit step functions. The values of these functions change from 0 to 1 at the times indicated by the subscripts: σ, the onset time of the initial social-distancing period; θ_1, the takeoff time of SARS-CoV-2 variant Alpha; and θ_2, the takeoff time of SARS-CoV-2 variant Delta.

$$U_{\sigma}(t) = \begin{cases} 0 & t < \sigma \\ 1 & t \geq \sigma \end{cases} \tag{32}$$

$$U_{\theta_1}(t) = \begin{cases} 0 & t < \theta_1 \\ 1 & t \geq \theta_1 \end{cases} \tag{33}$$

$$U_{\theta_2}(t) = \begin{cases} 0 & t < \theta_2 \\ 1 & t \geq \theta_2 \end{cases} \tag{34}$$

As indicated in Appendix Figure 1, transitions from S_M to S_P, for example, become possible at time $t = \sigma$, transitions from $S_{V,2}$ to E_V become possible at time $t = \theta_1$, and transitions from $S_{V,3}$ to E_V become possible at time $t = \theta_2$.

In the model, the quantities $P_\tau(t)$, and $\Lambda_\tau(t)$ are step functions that characterize changes in social-distancing. The value of $P_\tau(t)$ determines a setpoint steady-state fraction of susceptible persons who are practicing social-distancing. The value of $\Lambda_\tau(t)$ determines a time scale for approach to the setpoint steady state. Changes in the values of $P_\tau(t)$ and $\Lambda_\tau(t)$ occur coordinately. These changes occur at times $t = \sigma, \tau_1, \ldots, \tau_n$, where n is the number of distinct social-distancing periods beyond an initial social-distancing period. Initially, we took $n = 7$ (i.e., 8 total social-distancing stages). The value of n is decremented by 1 (at an inferred time) if $n \leftarrow n - 1$ is determined to be admissible by a model-selection procedure, which is described below. It should be noted that $p_0, p_1, \ldots, p_n$ are parameters of $P_\tau(t)$ and that $\lambda_0, \lambda_1, \ldots, \lambda_n$ are parameters of $\Lambda_\tau(t)$. These parameters determine the non-zero values of the step functions over different periods. For example, p_1 is the value of $P_\tau(t)$ and λ_1 is the value of $\Lambda_\tau(t)$ for the period $t \in [\tau_1, \tau_2)$.

$$P_\tau(t) = \begin{cases} p_0 & \sigma \le t < \tau_1 \\ p_1 & \tau_1 \le t < \tau_2 \\ \vdots & \vdots \\ p_n & \tau_n \le t < \infty \end{cases} \tag{35}$$

$$\Lambda_\tau(t) = \begin{cases} \lambda_0 & \sigma \le t < \tau_1 \\ \lambda_1 & \tau_1 \le t < \tau_2 \\ \vdots & \vdots \\ \lambda_n & \tau_n \le t < \infty \end{cases} \tag{36}$$

The values of $P_\tau(t)$ and $\Lambda_\tau(t)$ are 0 for $t < \sigma$.

In the model, the quantity $\mu(t)$ is a piecewise linear interpolant to a function $\tilde{\mu}(t)$ that characterizes the current rate of vaccination. The value of $\tilde{\mu}(t)$ is determined by the empirical daily rate of vaccination, and thus, can vary from day to day. Daily vaccination data were extracted from the Covid Act Now database and the *Democrat and Chronicle* newspaper [2] using the Covid Act Now Data API [1] and web scraping. We will use μ_i to refer to the value of $\tilde{\mu}(t)$ for $t \in [t_i, t_{i+1})$, where time t_i corresponds to midnight on the ith day after January 21, 2020.

$$\tilde{\mu}(t) = \mu_i \text{ for } t \in [t_i, t_{i+1}) \tag{37}$$

Settings for μ_i were made such that $\mu_i S_0 \times 1$ d is the number of vaccinations completed in the nearest past 1-d period according to Covid Act Now data and the *Democrat and Chronicle* newspaper.

In the model, the quantity $Y_\theta(t)$ is a step function that quantifies how disease transmissibility increases upon emergence of SARS-CoV-2 variants Alpha and Delta. Initially, $Y_\theta(t) = 1$. The value of $Y_\theta(t)$ is increased (by an inferred factor greater than 1 at an inferred time, θ_1 or θ_2) if the change is determined to be admissible by a model-selection procedure,

which is described below. It should be noted that y_1 and y_2 are parameters of $Y_\theta(t)$. These parameters determine the values of the step function $Y_\theta(t)$ over different periods: y_1 is the value of $Y_\theta(t)$ for the period $t \in [\theta_1, \theta_2)$ and y_2 is the value of $Y_\theta(t)$ for the remaining period of concern (with Delta as the dominant circulating viral strain).

$$Y_\theta(t) = \begin{cases} y_0 & \theta_0 \leq t < \theta_1 \\ y_1 & \theta_1 \leq t < \theta_2 \\ \vdots & \vdots \\ y_m & \theta_m \leq t < \infty \end{cases} \tag{38}$$

where m is the number of viral variants that have emerged up to the current time, $\theta_0 \equiv t_0$, and $y_0 \equiv 1$. Here, we consider $m = 2$.

Equations of the Auxiliary Measurement Model

As in the study of Lin et al. [3], we assumed that only symptomatic persons are detected in testing. The accumulation of symptomatic persons is governed by

$$\frac{dC_S}{dt} = (1 - f_A)\left[k_L\left(E_{5,M} + E_{5,P} + E_{5,Q}\right) + \left(\frac{k_L}{5}\right)E_V\right] \tag{39}$$

where $C_S(t)$ is the cumulative number of symptomatic persons (cases) at time t. Here, unlike in the study of Lin et al. [3], the expression for $C_S(t)$ accounts for exposed persons in quarantine. Initially, $C_S = 0$. We numerically integrated Appendix Equation (39) together with the ODEs of the compartmental model. From the trajectory for C_S, we derive a prediction for the expected number of new COVID-19 cases reported on calendar date $\mathcal{D}_\mathrm{i}$, $I(t_i, t_{i+1})$, using the following equation:

$$I(t_i, t_{i+1}) = f_D[C_S(t_{i+1}) - C_S(t_i)] \tag{40}$$

where f_D is an adjustable region-specific parameter characterizing the time-averaged fraction of symptomatic cases detected and reported, t_i corresponds to midnight on the ith day after January 21, 2020, and $t_{i+1} - t_i$ is the reporting interval (1 d). We compare $I(t_i, t_{i+1})$ to δC_i, the number of new cases reported on calendar date $\mathcal{D}_\mathrm{i}$.

Definition of the Likelihood Function

Bayesian inference relies on the definition of a likelihood, which here serves the purpose of assessing the compatibility of available surveillance data with adjustable (free) parameter values. Let us use $\{\delta C_i\}_{i=0}^{d}$ to denote the daily case reporting data available between 0 and d days after midnight on January 21, 2020 (the date of the first case report in the US) and let $D = \{\delta C_i\}_{i=0}^{d}$. Let us use $\theta_F(n, m)$ to denote the set of adjustable (free) parameter values. The number of adjustable parameters, $|\theta_F|$, depends on n, the number of social-distancing periods considered beyond an initial social-distancing period, and m, the number of SARS-CoV-2 variants under consideration. As in the study of Lin et al. [3], we assume that δC_i, the number of new COVID-19 cases detected over a 1-d period and reported on calendar date $\mathcal{D}_\mathrm{i}$ for a given region, is a random variable and its expected value follows a model-derived deterministic trajectory given by $I(t_i, t_{i+1})$ (Equation 40). We assume that day-to-day fluctuations in the random variable are independent and characterized by a negative binomial distribution $\mathrm{NB}(r, q_i)$, which has two parameters, $r > 0$ and $q_i \in (0,1)$. Note that $\mathbb{E}[\mathrm{NB}(r, q_i)] = r(1 - q_i)/q_i$. We assume that this distribution has the same dispersion parameter r across all case reports. With these assumptions, we arrive at the following likelihood function:

$$\mathcal{L}\left(\theta_F(n,m);\{\delta C_i\}_{i=0}^{d}\right) = \prod_{i=0}^{d} \mathcal{L}_i(\theta_F(n,m);\delta C_i) \tag{41}$$

where

$$\mathcal{L}_i(\theta_F(n,m);\delta C_i) = \mathrm{nbinom}(\delta C_i; r, q_i) = \binom{\delta C_i + r - 1}{\delta C_i - 1} q_i^r (1-q_i)^{\delta C_i} \tag{42}$$

and

$$q_i = \frac{r}{r + I(t_i, t_{i+1})}. \tag{43}$$

In these equations, i is an integer indicating the date $\mathcal{D}_i$ and period (t_i, t_{i+1}); $\mathrm{nbinom}(\delta C_i; r, q_i)$ is the probability mass function of the negative binomial distribution $\mathrm{NB}(r, q_i)$, and $\theta_F(n,m) = \{t_0, \beta, \sigma, \tau_1, \dots, \tau_n, p_0, p_{1,} \dots, p_n, \lambda_0, \lambda_1, \dots, \lambda_n, \theta_1, \dots, \theta_m, y_1, \dots, y_m, f_D, r\}$ for $n \geq 1$ and $m \geq 1$; $\theta_F(0,0) = \{t_0, \beta, p_0, \lambda_0, f_D, r\}$.

Parameters

Each model parameter is briefly described in Tables 1–3. These parameters have either fixed values or adjustable values (i.e., values inferred from surveillance data). The fixed values may be universal (i.e., applicable to all MSAs of interest) or MSA-specific. All inferred parameter values are MSA-specific. In addition, the measurement model (Appendix Equations 39 and 40) has one adjustable MSA-specific parameter, f_D, and the likelihood function (Appendix Equations 41–43) has one adjustable MSA-specific parameter, r. Values of the other likelihood parameters, $q_0, \dots, q_d$, are constrained and are determined using Appendix Equation 43.

Original model of Lin et al.

The model shares $19 + 3n$ parameters with the model of Lin et al. [3], including parameters that define the initial condition (t_0, I_0, and S_0). (Recall that n is the number of social-distancing periods being considered beyond the initial social-distancing period.) The shared parameters are t_0, I_0, S_0, β, σ, $\tau_1, \ldots, \tau_n$, $p_0, \ldots, p_n$, $\lambda_0, \ldots, \lambda_n$, ρ_A, ρ_E, m_b, f_A, f_H, f_R, k_L, k_Q, j_Q, c_A, c_H, and c_I. As in the study of Lin et al. [3], we inferred MSA-specific values for the following parameters: t_0, β, σ, $p_0, \ldots, p_n$, and $\lambda_0, \ldots, \lambda_n$. We also inferred MSA-specific values for $\tau_1, \ldots, \tau_n$ provided that $n \geq 1$. As in the study of Lin et al. [3], the remaining 14 parameters shared between the old and new models (I_0, S_0, ρ_A, ρ_E, m_b, f_A, f_H, f_R, k_L, k_Q, j_Q, c_A, c_H, and c_I) were taken to have fixed values, and we adopted the settings of Lin et al. [3] for these parameters (Table 3). These settings are universal except for the setting for S_0, the total population, which is MSA-specific.

Extension of the model of Lin et al.

Our extension of the model of Lin et al. [3] introduces $5 + 2(m + 1) + (d + 1)$ parameters, where m (= 0, 1 or 2) is the number of SARS-CoV-2 variants being considered and d is the number of days since January 21, 2020: $\theta_0, \ldots, \theta_m$, $y_0, \ldots, y_m$, m_h, f_0, f_1, f_2, k_V, and $\mu_0, \ldots, \mu_d$. The θ and y parameters are variant takeover times and transmissibility factors, respectively, except that the value of θ_0 is defined as t_0 and the value of y_0 is defined as 1. The Alpha transmissibility factor y_1, the Alpha takeoff time θ_1, the Delta transmissibility factor y_2, and the Delta takeoff time θ_2 were inferred for each MSA with $m = 2$ (cf. Table 1). The transmissibility factors were each constrained to be greater than or equal to 1. The settings for $\mu_0, \ldots, \mu_d$ are empirical and MSA-specific. Each μ_i is set such that $\mu_i S_0 \times 1$ d is the number of

833 vaccinations completed over the past 1-d period nearest to the ith day after January 21, 2020. As
834 noted earlier, the number of completed vaccinations was obtained for each MSA from Covid Act
835 Now and the *Democrat and Chronicle* newspaper [2] using the Covid Act Now Data API [1] and
836 web scraping. In the spreadsheet accessed daily from Covid Act Now, the
837 'metrics.vaccinationsCompletedRatio' column gives the percentage of the total population that
838 has received the recommended number of doses: one dose for Ad26.CoV2.S (Janssen, Johnson
839 & Johnson) or two doses for mRNA-1273 (Moderna) and BNT162b2 (Pfizer-BioNTech). As a
840 simplification, we considered all completed vaccinations to be equivalent. The parameters m_h,
841 f_0, f_1, f_2, and k_V were assigned fixed universal estimates (Table 3). Each of these estimates is
842 explained below.

843 **Estimation of k_V**

844 The rate constant k_V characterizes the rate of transition out of compartment V_i for $i =$
845 $1, \dots, n_V$. Recall that, in the model, susceptible persons enter V_1 upon vaccination (Figure 1,
846 Appendix Figure 1). The values of n_V $(= 6)$ and k_V $(= 0.3\ \mathrm{d}^{-1})$ were selected so that the time a
847 person spends in $V_1, \dots, V_{n_V}$, which we will denote as t_V, is distributed approximately the same as
848 $\tilde{t}_V$, the waiting time between vaccination of a previously uninfected person and detection of
849 vaccine-induced SARS-CoV-2-specific IgG antibodies [4] (Appendix Figure 2). According to
850 the model, the time that a person spends in $V_1, \dots, V_{n_V}$ is distributed according to the probability
851 density function $f(t_V; n_V, k_V) = k_V^{n_V} t_V^{n_V - 1} e^{-k_V t_V} / (n_V - 1)!$, i.e., t_V is Erlang distributed with
852 shape parameter $n_V = 6$ and rate parameter $k_V = 0.3\ \mathrm{d}^{-1}$. As can be seen in Appendix Figure 2,
853 the cumulative distribution function of this Erlang distribution reasonably captures the empirical
854 cumulative distribution of waiting times observed in the longitudinal study of Korodi et al. [4].

855 Thus, in the model, passage through $V_1, \ldots, V_6$ with rate constant $k_V = 0.3\ \mathrm{d}^{-1}$ accounts for the
856 variable and significantly non-zero amount of time required for development of a protective
857 antibody response after vaccination.

858 **Estimation of f_0, f_1, and f_2**

859 The parameters $f_0 > f_1$, $f_1 > f_2$, and f_2 are fractions that characterize the average effectiveness
860 of vaccines used in the US and that determine the sizes of (mutually exclusive) subpopulations of
861 vaccinated persons having different susceptibilities to productive infection (i.e., an infection that
862 can be transmitted to others): $S_{V,1}$, $S_{V,2}$, $S_{V,3}$, and $S_{V,4}$ (Figure 1, Appendix Figure 1). We assume
863 that persons in the $S_{V,1}$ subpopulation are susceptible to productive infection by any of the viral
864 strains considered, and in contrast, we assume that persons in the $S_{V,4}$ subpopulation are
865 susceptible to productive infection by none of the viral strains considered. Persons in the $S_{V,2}$
866 subpopulation are taken to be susceptible to productive infection by the Alpha and Delta variants
867 but not viral strains in circulation before the emergence of Alpha. Persons in the $S_{V,3}$
868 subpopulation are taken to be susceptible to productive infection by the Delta variant but not the
869 Alpha variant or viral strains in circulation before the emergence of Alpha. The quantity $1 - f_0$
870 defines the fraction of vaccinated persons who enter the $S_{V,1}$ subpopulation after exiting V_6, the
871 quantity $f_0 - f_1$ defines the fraction of vaccinated persons who enter the $S_{V,2}$ subpopulation after
872 exiting V_6, the quantity $f_1 - f_2$ defines the fraction of vaccinated persons who enter the $S_{V,3}$
873 subpopulation after exiting V_6, and f_2 defines the fraction of vaccinated persons who enter the
874 $S_{V,4}$ subpopulation after exiting V_6. We take f_0 to characterize vaccine effectiveness before the
875 emergence of Alpha. According to Thompson et al. [5], vaccine effectiveness was initially 90%.
876 Thus, we set $f_0 = 0.9$. We take f_1 to characterize vaccine effectiveness after the emergence of

877 Alpha but before the emergence of Delta. According to Puranik et al. [6], in May 2021, vaccine
878 effectiveness was 81%. Thus, we set $f_1 = 0.81$. We take f_2 to characterize vaccine effectiveness
879 after the emergence of Delta. According to Tang et al. [7], the effectiveness of two doses of the
880 Pfizer-BioNTech vaccine (BNT162b2) against Delta is 53.5% and the effectiveness of two doses
881 of the Moderna vaccine (mRNA-1273) against Delta is 84.8%. Taking the average of these
882 figures, we set $f_2 = 0.69$.

883 **Estimation of m_h**

884 The parameter m_h characterizes the reduced risk of severe disease for a vaccinated
885 person in the case of a breakthrough infection. We set $m_h = 0.04$, i.e., we assumed that there is a
886 25-fold reduction in the risk of severe disease for infected persons who have been vaccinated,
887 which is consistent with the observations of Lopez Bernal et al. [8].

888 **Notable New Modeling Assumptions**

889 It should be noted that we treat the incubation period for newly infected (exposed)
890 vaccinated persons differently than for newly infected (exposed) unvaccinated persons (Figure 1,
891 Appendix Figure 1). For unvaccinated persons, as in the study of Lin et al. [3], we divide
892 exposed persons in the incubation period of infection into five subpopulations: $E_{1,M}, \dots, E_{5,M}$ for
893 exposed persons who are mixing (i.e., persons who are not practicing social-distancing),
894 $E_{1,P}, \dots, E_{5,P}$ for exposed persons who are practicing social-distancing, and $E_{1,Q}, \dots, E_{5,Q}$ for
895 exposed quarantined persons. Persons move through the five stages of the incubation period
896 sequentially. In contrast, as a simplification, for vaccinated persons, we consider only a single
897 exposed population: E_V. We take persons to exit E_V with rate constant $k_L/5$ (Appendix Figure
898 1). With this choice, the duration of the incubation period of infection is the same, on average,

for both vaccinated and unvaccinated persons. The average duration is $5/k_L$ (about 5 d) in both cases. The difference is that the duration of the incubation period is Erlang distributed for unvaccinated persons, as discussed by Lin et al. [3], but exponentially distributed for vaccinated persons.

As indicated in Equation (29), we take vaccinated persons with productive infections to be equally as infectious as unvaccinated persons.

As noted earlier, we take all vaccinated persons to be mixing (i.e., to not be practicing social-distancing). Thus, populations of infected vaccinated persons (E_V, I_V, and A_V) contribute to $\phi_M(t)$ (Appendix Equation 29) but not $\phi_P(t)$ (Appendix Equation 30).

As indicated in Appendix Equation 31, we consider pre-symptomatic exposed and asymptomatic unvaccinated persons to be eligible for vaccination and, thus, these persons contribute to the consumption of vaccine doses (i.e., these persons account for a portion of the number of completed vaccinations on a given day i, $\mu_i S_0 \times 1$ d). However, we do not move these persons to vaccinated compartments. The reason is that exposed and asymptomatic persons are expected to develop immunity faster through recovery from infection (i.e., movement to R_U) than from vaccination.

As indicated in Appendix Equation 31, we do not consider symptomatic, quarantined, severely ill and hospitalized/isolated-at-home, and deceased persons to be eligible for vaccination.

Inference Approach

Recall that θ_F denotes the set of all adjustable parameters. As in the study of Lin et al. [3], for each MSA, we inferred MSA-specific adjustable parameter values θ_F using all MSA-specific surveillance data available up to a specified day of inference $\mathcal{D}_\mathrm{d}$ (i.e., the dth day after January 21, 2020). We took a Bayesian inference approach, meaning that, for a given dataset, we generated parameter posterior samples (a collection of θ_F's) through Markov chain Monte Carlo (MCMC) sampling. The parameter posterior samples provide a probabilistic characterization of the adjustable parameter values consistent with the dataset used in inference. By drawing samples from the parametric posterior distribution, we generated a posterior predictive distribution for $I(t_i, t_{i+1})$ for each i of interest. We considered all days from January 21, 2020 to October 31, 2021. In other words, for each i of interest, a prediction for $I(t_i, t_{i+1})$ was generated for each of many θ_F's drawn randomly (with uniform probability) from the parametric posterior distribution. The resulting distribution of $I(t_i, t_{i+1})$ values is the posterior predictive distribution for $I(t_i, t_{i+1})$. Recall that $I(t_i, t_{i+1})$ is given by Appendix Equation 40 and corresponds to $\mathbb{E}[\delta C_i]$, the expected number of new COVID-19 cases detected over a 1-d surveillance interval and reported for the ith day after January 21, 2020. Observation noise was injected into the posterior predictive distributions by replacing each sampled value for $I(t_i, t_{i+1})$ with $X_i \sim \mathrm{NB}(r, q_i)$, where r is a member of the sampled set of parameter values θ_F used to generate the prediction of $I(t_i, t_{i+1})$ and q_i is given by Equation 43.

According to Bayes' theorem, given surveillance data $D = \{\delta C_i\}_{i=0}^{d}$, the parametric posterior is given as

$$\mathbb{P}\{\theta_F | D\} = \frac{\mathbb{P}\{D | \theta_F\}\, \mathbb{P}\{\theta_F\}}{Z} \tag{44}$$

where $\mathbb{P}\{\theta_F\}$ is the prior (which is formulated to capture knowledge of θ_F external to D or to express lack of such knowledge), $\mathbb{P}\{D|\theta_F\}$ is the likelihood defined by Appendix Equations 41–43, and Z is a normalizing constant. We assumed a proper uniform prior, i.e., for each adjustable parameter, we assumed that all values between specified lower and upper bounds are equally likely before consideration of D. We used the same bounds as in the study of Lin et al. [3]. Then, we used an adaptive MCMC (aMCMC) algorithm [9] to generate samples from $\mathbb{P}\{D|\theta_F\}\ \mathbb{P}\{\theta_F\}$, which is proportional to $\mathbb{P}\{\theta_F|D\}$. Thus, the relative probabilities of parameter sets θ_F according to $\mathbb{P}\{\theta_F|D\}$ are correctly represented by the samples.

Specifically, the adaptive MCMC algorithm [9] generates samples from the multivariate parametric posterior for adjustable model parameters (t_0, β, and parameters for variant emergence and social-distancing), the measurement model parameter f_D, and the likelihood parameter r (Tables 1 and 2). This algorithm is available within the PyBioNetFit software package [10]. Use of the algorithm was performed as described by Lin et al. [3]. The report of Neumann et al. [10] includes helpful general usage advice, which was followed in this study. Inference jobs were executed on a computer cluster. For each inference job, a total of 25 chains were generated, and the chain with the best mixing and convergence properties was selected for subsequent analyses.

Each inference was conditioned on the compartmental model of Figure 1 (Appendix Equations 1–38), settings for the structural parameters m (the number of SARS-CoV-2 variants under consideration) and n (the number of social-distancing periods under consideration beyond an initial social-distancing period), the measurement model (Appendix Equations 39 and 40), and fixed parameter estimates (Tables 1 and 2), including empirical daily *per capita* vaccination rates (i.e., the settings for μ_i in Appendix Equation 37). We assumed a proper uniform prior for each

adjustable parameter [3] and a negative binomial likelihood function (Appendix Equations 41–43). Use of proper uniform priors means that MAP estimates are maximum likelihood estimates (MLEs). In each inference, the data entering the likelihood function, $D = \{\delta C_i\}_{i=0}^{d}$ (Appendix Equation 41), were MSA-specific daily reports of newly detected COVID-19 cases available up to the date of inference $\mathcal{D}_{\mathrm{d}}$ (i.e., the d-th day after January 21, 2020). Thus, all inferences are region-specific and time-dependent.

Use of Model Selection to Determine Intervals of Step Functions

Variant takeover times, $\theta = (\theta_1, \theta_2)$, and start times of social-distancing periods, $\tau = (\sigma, \tau_1, \dots, \tau_n)$, were inferred from data; however, changes of the associated time-dependent step functions, $Y_\theta(t)$, $P_\tau(t)$, and $\Lambda_\tau(t)$, were introduced only when an increase in model complexity was deemed to be justified. Each decision to introduce variant takeover or start of a new social-distancing period (beyond the initial period) was made using a model-selection procedure, which is described below. It should be noted that y_1 and y_2 are parameters of $Y_\theta(t)$, $p_0, p_1, \dots, p_n$ are parameters of $P_\tau(t)$, and $\lambda_0, \lambda_1, \dots, \lambda_n$ are parameters of $\Lambda_\tau(t)$. These parameters determine the values of the step functions over different periods. For example, for $n \geq 1$, p_1 is the value of $P_\tau(t)$ and λ_1 is the value of $\Lambda_\tau(t)$ for the period $t \in [\tau_1, \tau_2)$. Similarly, y_1 is the value of $Y_\theta(t)$ for the period $t \in [\theta_1, \theta_2)$, and y_2 is the value of $Y_\theta(t)$ for the period $t \in [\theta_2, t_{\mathrm{final}})$, where t_{final} corresponds to the date of inference (October 31, 2021).

We started with a setting of $n = 7$ for each MSA of interest (i.e., 8 total social-distancing stages). To determine if n could be reduced, we conducted parsimony checks. In a parsimony check, 100 MLE curves, each constituting a fit to the data, were generated via optimization jobs. In these jobs, the total number of social-distancing stages in the model $(n + 1)$ was set at 1 less

than for the current proposed best fit. Each of the 100 fits was visually inspected to determine whether or not the following criteria held: 1) the quality of fit is acceptable (i.e., comparable to what is obtained with the current proposed best fit); 2) Alpha and Delta surges (identified by sequencing data) are explained at least partly by increased transmissibility; 3) social-distancing setpoint parameter values are feasible (i.e., each setpoint parameter takes on a value between 0.2 and 0.8); and 4) social-distancing changes proximal to an Alpha or Delta surge (if any) occur only after an increase in transmissibility. If one or more of these conditions was not satisfied, we accepted the proposed best fit as the most parsimonious fit to the data. If all conditions held, we updated the proposed best fit, and the parsimony check was repeated for a model with one less social-distancing stage. In the Appendix, we show examples of violations of parsimony-check criteria for the MSA surrounding Houston (Appendix Figure 3).

Simulations

After specification of parameter values (Tables 1–3), we used the SciPy (https://www.scipy.org) interface to LSODA [12] to numerically integrate the system of coupled ODEs consisting of the 40 ODEs of the compartmental model and the 1 ODE of the measurement model (Appendix Equations 1–39). The initial condition was defined by settings for t_0, I_0, and S_0 (Tables 1–3). Integration combined with use of Appendix Equation 40 yielded a prediction of the expected number of new cases detected for each 1-d surveillance period of interest in the past or future: $I(t_i, t_{i+1})$, where t_i corresponds to midnight on the ith day after January 21, 2020. To account for randomness in case detection and reporting, we replaced $I(t_i, t_{i+1})$ with $X_i \sim \mathrm{NB}(r, q_i)$, where q_i is given by Appendix Equation 43.

Appendix References

1006 [1] Covid Act Now. US COVID risk and vaccine tracker. 2021 [cited 2021 Sep 29]
1007 https://covidactnow.org/
1008 [2] Democrat and Chronicle. COVID-19 vaccine tracker. 2023 [cited 2023 Nov 11]
1009 https://data.democratandchronicle.com/covid-19-vaccine-tracker
1010 [3] Lin YT, Neumann J, Miller EF, Posner RG, Mallela A, Safta C, Ray J, Thakur G,
1011 Chinthavali S, Hlavacek WS. Daily forecasting of regional epidemics of Coronavirus
1012 Disease with Bayesian uncertainty quantification, United States. Emerg Inf Dis.
1013 2021;27:767–78. https://doi.org/10.3201/eid2703.203364
1014 [4] Korodi M, Rákosi K, Jenei Z, Hudák G, Horváth I, Kákes M, et al. Longitudinal
1015 determination of mRNA-vaccination induced strongly binding SARS-CoV-2 IgG
1016 antibodies in a cohort of healthcare workers with and without prior exposure to the novel
1017 coronavirus. Vaccine. 2022;40:37: 5445–5451.
1018 https://doi.org/10.1016/j.vaccine.2022.07.040
1019 [5] Thompson MG, Burgess JL, Naleway AL, Tyner HL, Yoon SK, Meece J, et al. Interim
1020 estimates of vaccine effectiveness of BNT162b2 and mRNA-1273 COVID-19 vaccines
1021 in preventing SARS-CoV-2 infection among health care personnel, first responders, and
1022 other essential and frontline workers—eight U.S. locations, December 2020–March 2021.
1023 MMWR Morb Mortal Wkly Rep. 2021;70:495–500.
1024 http://dx.doi.org/10.15585/mmwr.mm7013e3
1025 [6] Puranik A, Lenehan PJ, Silvert E, Niesen MJM, Corchado-Garcia J, O'Horo JC, et al.
1026 Comparison of two highly-effective mRNA vaccines for COVID-19 during periods of
1027 Alpha and Delta variant prevalence. [cited 2021 Oct 8]
1028 https://www.medrxiv.org/content/10.1101/2021.08.06.21261707v3

[7] Tang P, Hasan MR, Chemaitelly H, Yassine HM, Benslimane FM, Al Khatib HA, et al. BNT162b2 and mRNA-1273 COVID-19 vaccine effectiveness against the SARS-CoV-2 Delta variant in Qatar. Nat Med. 2021;27:2136–43. https://www.nature.com/articles/s41591-021-01583-4

[8] Lopez Bernal J, Andrews N, Gower C, Gallagher E, Simmons R, Thelwall S, et al. Effectiveness of Covid-19 Vaccines against the B.1.617.2 (Delta) variant. N Engl J Med. 2021;385:585–94. https://doi.org/10.1056/NEJMoa2108891

[9] Andrieu C, Thoms J. A tutorial on adaptive MCMC. Stat Comput. 2008;18:343–73. https://doi.org/10.1007/s11222-008-9110-y

[10] Neumann J, Lin YT, Mallela A, Miller EF, Colvin J, Duprat AT, Chen Y, Hlavacek WS, Posner RG. Implementation of a practical Markov chain Monte Carlo sampling algorithm in PyBioNetFit. Bioinformatics. 2022;38:1770–2. https://doi.org/10.1093/bioinformatics/btac004

[11] Burnham KP, Anderson DR. Multimodal inference: understanding AIC and BIC in model selection. Sociol Methods Res. 2016;33:261–304. https://doi.org/10.1177/0049124104268644

[12] Hindmarsh AC. ODEPACK, a systematized collection of ODE solvers. In Stepleman RS, editor. Scientific computing: applications of mathematics and computing to the physical sciences. Amsterdam: North-Holland Publishing Company; 1983. p. 55–64.

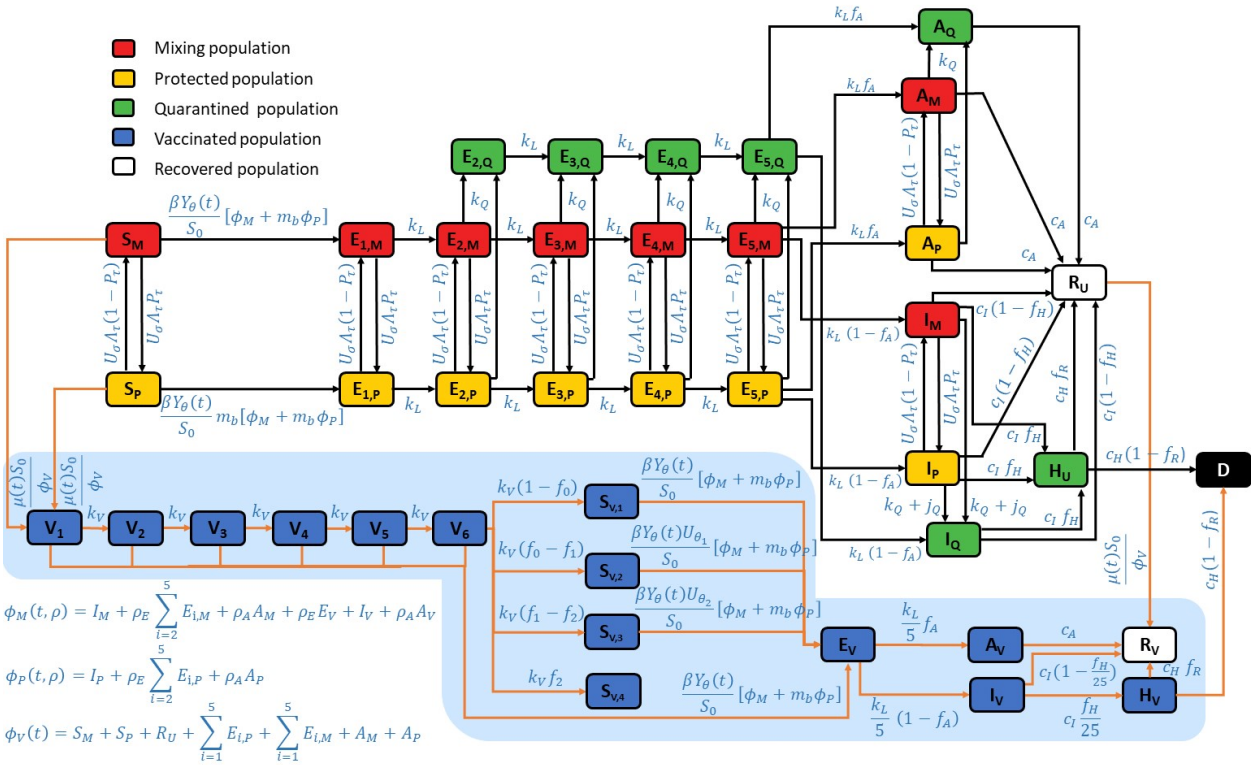
1050

Appendix Table 1. State variables of the compartmental model

State variable (population)	Description
S_M	Population of susceptible unvaccinated persons who are mixing (i.e., not practicing social-distancing)
S_P	Population of susceptible unvaccinated persons who are practicing social-distancing
$S_{V,1}$	Population of vaccinated unexposed persons who developed an immune response to vaccination that does not protect against productive infection by ancestral strains or the Alpha and Delta variants
$S_{V,2}$	Population of vaccinated unexposed persons who developed an immune response to vaccination that protects against productive infection by ancestral strains but not the Alpha or Delta variants
$S_{V,3}$	Population of vaccinated unexposed persons who developed an immune response to vaccination that protects against productive infection by ancestral strains and the Alpha variant but not the Delta variant
$S_{V,4}$	Population of vaccinated unexposed persons who developed an immune response to vaccination that protects against productive infection by ancestral strains and the Alpha and Delta variants
V_i $(i = 1, \dots, 6)$	Population of vaccinated persons in the ith stage of immune response to vaccination
$E_{i,M}$ $(i = 1, \dots, 5)$	Population of exposed unvaccinated persons in the ith stage of the incubation period of infection and who are mixing
$E_{i,P}$ $(i = 1, \dots, 5)$	Population of exposed unvaccinated persons in the ith stage of the incubation period of infection and who are practicing social-distancing
$E_{i,Q}$ $(i = 2, \dots, 5)$	Population of exposed unvaccinated persons in the ith stage of the incubation period of infection and who are quarantined
E_V	Population of vaccinated persons in the incubation period of a productive infection (i.e., an infection that can be transmitted to others)

A_M	Population of asymptomatic unvaccinated persons who are in the immune clearance phase of infection and who are mixing
A_P	Population of asymptomatic unvaccinated persons who are in the immune clearance phase of infection and who are practicing social-distancing
A_Q	Population of asymptomatic unvaccinated persons who are in the immune clearance phase of infection and who are quarantined
A_V	Population of asymptomatic vaccinated persons who are in the immune clearance phase of a productive infection (i.e., an infection that can be transmitted to others)
I_M	Population of infectious, symptomatic, and unvaccinated persons with mild disease who are mixing
I_P	Population of infectious, symptomatic, non-vaccinated, and infectious persons with mild disease who are practicing social-distancing
I_Q	Population of infectious, symptomatic, and unvaccinated persons with mild disease who are quarantined
I_V	Population of infectious, symptomatic, and vaccinated persons with mild disease
R_U	Population of recovered unvaccinated persons
R_V	Population of recovered vaccinated persons
H_U	Population of unvaccinated persons with severe disease who are hospitalized or isolated at home
H_V	Population of vaccinated persons with severe disease who are hospitalized or isolated at home
D	Population of deceased persons

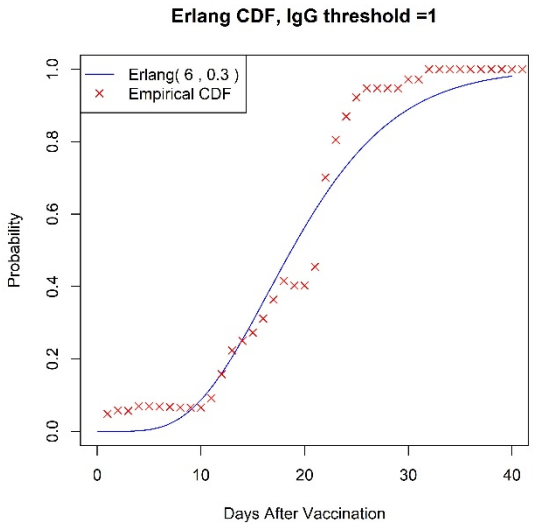
1051



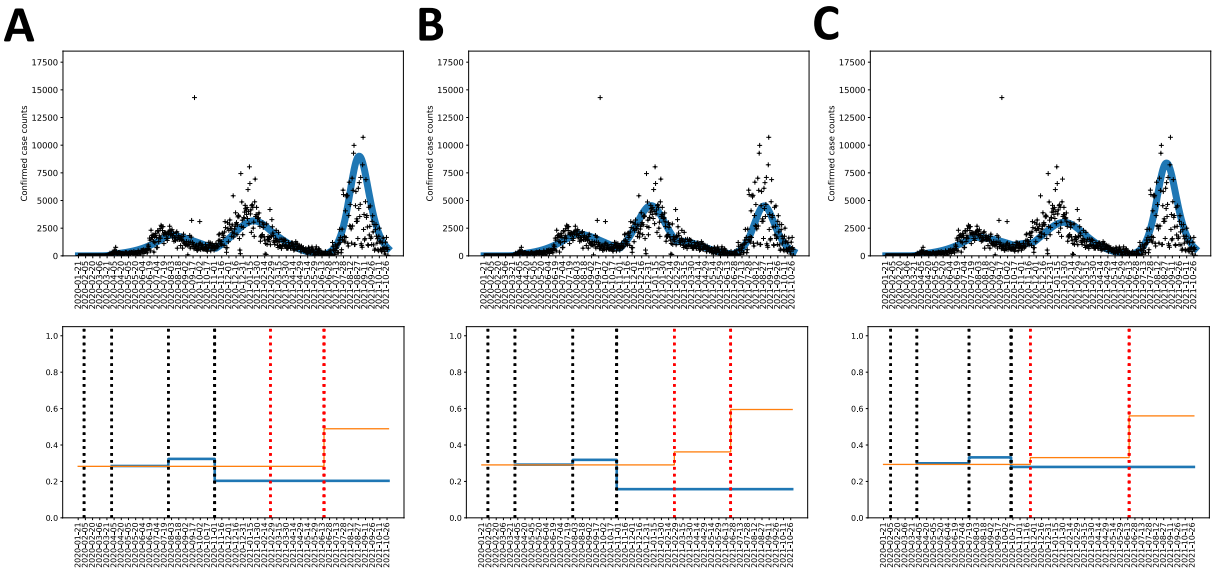
Appendix Figure 1. Expanded illustration of the new compartmental model. In the extended model, vaccination was considered by allowing susceptible and recovered persons to transition into a vaccinated compartment, either V_1 and R_V. Susceptible (recovered) persons who have completed vaccination move into the V_1 (R_V) compartment. The susceptible persons who move into V_1 are drawn from S_M (populated by susceptible persons who are mixing and unprotected by social-distancing) and from S_P (populated by susceptible persons who are protected by social-distancing). After susceptible persons enter V_1, they can move through a series of additional compartments (V_2 through V_6), which are included to capture the time needed for immunity to develop after completion of vaccination. We estimate that the time needed to acquire immunity after vaccination is approximately three weeks based on longitudinal studies of anti-spike protein IgG levels [4]. Persons who exit the V_6 compartment without becoming infected enter one of the following compartments: $S_{V,1}$, $S_{V,2}$, $S_{V,3}$, or $S_{V,4}$. Persons in $S_{V,1}$ are taken to remain susceptible

to productive infection by all SARS-CoV-2 strains of interest (Alpha, Delta, and ancestral strains). Persons in $S_{V,2}$ are taken to be susceptible to SARS-CoV-2 Alpha and Delta variants. Persons in $S_{V,3}$ are taken to be susceptible to Delta. Persons in $S_{V,4}$ are taken to be protected against all strains of interest. Infection of persons in $S_{V,3}$ is only allowed if Delta is present, i.e., at times $t > \theta_2$. Infection of persons in $S_{V,2}$ is only allowed if Alpha or Delta is present, i.e., at times $t > \theta_1$. Vaccinated persons in compartments V_1 through V_6 and compartment $S_{V,1}$ are allowed to become infected at any time, at which point they transition to compartment E_V, consisting of vaccinated persons who were exposed before development of vaccine-induced immunity. Persons in compartment $S_{V,2}$ are allowed to become infected if $t \geq \theta_1$. Similarly, persons in compartment $S_{V,3}$ are allowed to become infected if $t \geq \theta_2$. Possible outcomes for persons in E_V are taken to be the same as those for unvaccinated exposed persons; however, the incubation period is taken to be distinct. Persons in E_V can experience asymptomatic disease (upon entering A_V) or they can become symptomatic (upon entering I_V). Persons in A_V eventually recover, entering compartment R_V. Persons in I_V can progress to severe disease (upon entering H_V) or recover (upon entering R_V). Persons in H_V either recover (moving into R_V) or die (moving into D). Persons who have recovered from infection, in the R_U compartment, move directly into the R_V compartment upon vaccination. Persons in the R_U and R_V compartments are taken to have full immunity. The vaccination rate at which susceptible and recovered persons move into vaccinated compartments is updated daily for consistency with the empirical overall rate of vaccination, which we extract daily from the COVID Act Now Data API [1] and the *Democrat and Chronicle* newspaper [2]. The relative values of the vaccination rate are set such that each person eligible for vaccination has the same probability of being vaccinated. All unvaccinated persons are taken to be eligible for vaccination except symptomatic persons (in

compartments I_M and I_P), persons who are hospitalized or severely ill at home (in compartment H), quarantined persons (in the various compartments labeled with a Q subscript), and deceased persons (in compartment D). It should be noted that asymptomatic, non-quarantined persons (in compartments A_M and A_P) and presymptomatic, non-quarantined persons (in the E compartments) are taken to be eligible (and to influence the vaccination rate constants) but, as a simplification, these persons are not explicitly tracked as vaccinated or unvaccinated because each of these persons will eventually enter either the D compartment or the R_U compartment, at which point they will have immunity. In the model, the effects of SARS-CoV-2 variants are captured by a time-dependent dimensionless multiplier $Y_\theta(t)$ of the rate constant β. This rate constant, which appears in Appendix Equations 1–4, 18–22, and 24, determines the rate of disease transmission within the subpopulation unprotected by social-distancing behaviors when $Y_\theta(t) = 1$. We take $Y_\theta(t) = 1$ for times $t < \theta_1$, i.e., for the initial period of the COVID-19 pandemic in the US that we take to have started on January 21, 2020. We take $Y_\theta(t)$ to have the form of a step function with distinct values greater than 1 for periods starting at $t = \theta_1$ and $\theta_{k+1} > \theta_k$ for $k = 1, \dots, m$. Thus, the model allows for m distinct periods of variant strain dominance delimited by a set of start times $\theta = \{\theta_1, \dots, \theta_m\}$. We considered $m = 2$. We assume that variants differ only in transmissibility.



Appendix Figure 2. Comparison of an Erlang cumulative distribution function with shape parameter $n_V = 6$ and rate parameter $k_V = 0.3\ \mathrm{d}^{-1}$ and the empirical cumulative distribution of waiting times ($\tilde{t}_V$ values) observed in the longitudinal study of Korodi et al. [4]. The waiting time $\tilde{t}_V$ is the time between vaccination of a previously uninfected person and detection of vaccine-induced SARS-CoV-2-specific IgG antibodies.



Appendix Figure 3. Example of a parsimony check with criteria that are violated for the MSA surrounding Houston. (A) Alpha and Delta surges (identified by sequencing data) are not explained, at least in part, by increased transmissibility, while all other criteria are satisfied. Each broken vertical black line from left to right indicates the date of onset of a social-distancing stage (i.e., with n = 3 for this illustration). Each broken vertical red line from left to right indicates the takeoff date of a variant (namely, Alpha or Delta). (B) Not all social-distancing setpoint parameter values are feasible, while all other criteria are satisfied. (C) Social-distancing changes proximal to an Alpha or Delta surge (if any) precede an increase in transmissibility, while all other criteria are satisfied. In Panel A, we see that the orange segment between the two broken vertical red lines, which denotes the relative transmissibility of Alpha, is no larger than the orange segment to its left, which denotes the relative transmissibility of ancestral strains of SARS-CoV-2. In Panel B, the rightmost blue segment, which denotes the social-distancing setpoint parameter value for the final social-distancing stage, has a value of roughly 0.16, which is deemed infeasible by our criterion. Finally, in Panel C, the rightmost broken vertical black line

is proximal to the leftmost broken vertical red line, which corresponds to the takeoff time of the Alpha surge, and the social-distancing setpoint parameter value changes prior to the Alpha surge. Note that all three panels show MLE curves, each of which constitutes an acceptable fit to the data.