The Paradox of Neglecting Changes in Behavior: How Standard Epidemic Models Misestimate Both Transmissibility and Final Epidemic Size

Binod Pant[1,2,*], Marko Lalovic[3,*], István Z. Kiss[3,4], and Mauricio Santillana[1,2,5†]

[1] *Machine Intelligence Group for the Betterment of Health and the Environment, Northeastern University, Boston, MA, 02115, USA*

[2] *Network Science Institute, Northeastern University, Boston, MA, 02115, USA*

[3] *Network Science Institute, Northeastern University London, London, E1W 1LP, United Kingdom*

[4] *Department of Mathematics, Northeastern University, Boston, MA 02115, USA*

[5] *Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, 02115, USA*

Abstract

During epidemic outbreaks, populations adapt their behavior in response to disease burden, fundamentally altering transmission dynamics. Despite this, most compartmental models assume constant contact rates throughout outbreaks. To quantify biases from this assumption, we fitted a baseline SEIRD model with constant transmission and three behavioral variants—incorporating mortality-driven transmission reduction via exponential, rational, and mixed functional forms—to COVID-19 mortality data from 30 US locations during the first pandemic wave (March–July 2020). All three behavioral models achieved a lower median normalized sum of squared error in at least 28 of 30 locations, and Bayesian model selection favored them in at least 28 of 30 locations. More importantly, we identified systematic biases when behavioral responses are ignored: the baseline model consistently underestimated the basic reproduction number ($\mathcal{R}_0$) while paradoxically overestimating the final epidemic size. Median $\mathcal{R}_0$ estimates from the behavioral models exceeded the baseline estimates across all 30 locations, yet baseline models predicted larger cumulative infection burdens. Controlled synthetic experiments—where mortality trajectories were generated from behavioral models with known parameters—confirmed these biases result from model misspecification rather than data quality or stochastic variation. We prove analytically that for any fixed $\mathcal{R}_0$, the baseline model overestimates cumulative infections compared to behavioral models where mortality reduces transmission, regardless of functional form. This dual bias poses serious risks for pandemic response: standard models may simultaneously underestimate pathogen contagiousness (delaying critical early action) while overestimating infection burden (causing excessive late-phase resource allocation). Our findings across 30 geographically diverse locations demonstrate that incorporating behavioral change substantially improves both model fit and estimation of epidemiological parameters essential for public health policy.

Author Summary. When diseases spread, people change their behavior—avoiding crowds, wearing masks, washing hands more frequently. Yet most mathematical models used to predict epidemics assume people behave the same way throughout an outbreak. We asked: What happens when models ignore these behavioral changes? Using COVID-19 data from 30 US locations, we compared a traditional model (assuming constant behavior) against models where people reduce contact as deaths increase. We discovered a troubling paradox: models ignoring behavior consistently suggest diseases are less contagious than they really are, yet simultaneously predict that more people will get infected. This creates a dangerous mismatch for decision-makers: underestimating how easily a disease spreads may delay urgent early actions like school

*These authors contributed equally to this study

†Corresponding author: Email: *msantill@g.harvard.edu*

closures or travel restrictions, while overestimating total infections wastes resources preparing for scenarios that behavioral adaptation prevents. We confirmed this paradox through computer simulations where we knew the true answer, and proved analytically that behavioral adaptation reduces final epidemic size even when the basic reproduction number is held constant. Our work shows that incorporating human behavioral responses is not just a modeling refinement—it is essential for accurate epidemic predictions that inform life-or-death policy decisions.

1 Introduction

Behavioral responses to epidemic threats fundamentally shape disease transmission. During the 1918 influenza pandemic, communities closed schools, banned public gatherings, and adopted mask-wearing to reduce transmission [1, 2]. The HIV/AIDS epidemic saw widespread adoption of condom use, averting an estimated 100 million infections between 1990 and 2019 [3]. More recently, the COVID-19 pandemic triggered rapid behavioral shifts, including masking and social distancing, with mask mandates alone associated with significant reductions in cases and deaths [4, 5]. Recent survey data from the United States suggests that human behavior, which varied spatially and temporally during COVID-19, was strongly associated with changes in COVID-19 mortality and hospitalization [6], indicating that observed disease severity indicators drive behavioral adaptation. These examples demonstrate that human populations do not passively experience epidemics—they adapt their behavior in response to perceived risk, fundamentally altering transmission dynamics.

Compartmental models based on Susceptible-Exposed-Infected-Recovered (SEIR) frameworks have been widely used to understand epidemic trajectories and guide public health interventions [7]. Despite historical and contemporary evidence showing that people actively adapt their behavior in response to disease threats, most models assume populations passively experience outbreaks by maintaining constant contact patterns. Although the assumption of constant transmission rates is mathematically convenient, it ignores the dynamic feedback between disease burden and behavioral change. Recent modeling efforts have begun incorporating behavioral responses through various mechanisms, including prevalence-dependent [8–11], hospitalization-driven [12], and mortality-driven feedback [8, 13, 14]. Lejeune et al. [15] classify these approaches as either endogenous (behavior changes as a function of disease state variables within the model, creating a feedback loop) or exogenous (behavior changes through fixed or time-varying parameters informed by external data sources such as government policies and mobility patterns, or through modelers' assumptions, rather than a state variable within a model itself).

The critical question is whether systematic biases arise when models fail to account for behavioral change. If populations adapt their behavior in response to disease burden—as historical and contemporary evidence suggests—then models assuming constant transmission may systematically mischaracterize epidemic dynamics. Such mischaracterization could have serious consequences for public health decision-making: underestimating pathogen transmissibility might lead to insufficient interventions, while overestimating infection burden could result in misallocated resources. Despite these stakes, few studies have systematically investigated—either numerically or analytically—whether and to what extent omitting behavioral feedback introduces biases in epidemic models. Recent work has demonstrated that incorporating behavioral responses improves model fit to observed data [8, 12, 14, 16, 17], with behavioral models capturing epidemic trajectories better than their behavior-free counterparts. Behavioral models have also shown superior performance in out-of-sample validation [8, 12]. Theoretical studies have examined how behavioral mechanisms—including vaccination behavior [18], awareness-driven protection [19], and incidence-dependent contact reduction [9]—alter epidemic outcomes such as cumulative infection burden; notably, models ignoring behavior change have systematically overshoot final size projections [20]. However, systematic comparison of key epidemiological quantities—such as basic reproduction number and final size—estimated through models with versus without behavioral feedback, using real data across many locations, remains lacking.

We address this gap by systematically comparing compartmental models with and without behavioral feedback, fitted to COVID-19 mortality data from 30 US locations (29 states and the District of Columbia) during the first pandemic wave (March–July 2020). We implement three behavioral models where the transmission rate decreases as mortality increases, alongside a baseline model with constant transmission. Using

Bayesian methods, we fit each model to observed mortality trajectories, estimate posterior distributions for the basic reproduction number and final epidemic size, and compare model performance both qualitatively and quantitatively. To test whether biases arise from model misspecification rather than data limitations, we conduct controlled synthetic experiments where true parameter values are known by design. We also analytically derive the relationship between final epidemic size in baseline versus behavioral models when the basic reproduction number is identical.

Our analysis reveals three key findings with important public health implications. First, behavioral models consistently outperform baseline models across nearly all locations examined, as evidenced by both qualitative and quantitative comparisons. Second, baseline models systematically underestimate the basic reproduction number while counterintuitively overestimating the total proportion of the population infected compared to all three behavioral model formulations. We confirm this dual bias through synthetic experiments, demonstrating that these biases emerge when behavioral feedback is present in the data but models do not explicitly account for behavioral change. Third, we prove analytically that this paradox—lower estimated transmissibility yet higher predicted final epidemic size—is an inevitable consequence of ignoring behavioral responses. These findings demonstrate that incorporating behavioral dynamics is not merely a modeling refinement but essential for accurate epidemic parameter estimation and evidence-based pandemic response.

This paper is organized as follows. In Section 2, we formulate the baseline model that does not explicitly account for behavioral change and three behavioral model variants, describe the COVID-19 mortality data used for calibration, and outline our Bayesian parameter inference approach. Section 3 presents our three key findings: behavioral models' superior fit to observed data, systematic underestimation of the basic reproduction number by baseline models, and overestimation of infection burden when behavioral responses are ignored. We support these findings through both real data analysis across 30 US states and controlled synthetic experiments. Section 4 discusses the public health implications and limitations of our work.

2 Methods

2.1 Model Formulation

We developed a compartmental SEIRD model incorporating behavioral responses to disease-induced mortality. The population is stratified into susceptible (S), exposed (E), infectious (I), recovered (R), and deceased (D) compartments, with total living population $N(t) = S(t) + E(t) + I(t) + R(t)$ at time t. The model dynamics are governed by the following system of differential equations:

$$\begin{aligned}
\frac{dS}{dt} &= -\beta(t)\frac{SI}{N}, \\
\frac{dE}{dt} &= \beta(t)\frac{SI}{N} - \sigma E, \\
\frac{dI}{dt} &= \sigma E - (\gamma + \delta)I, \\
\frac{dR}{dt} &= \gamma I, \\
\frac{dD}{dt} &= \delta I,
\end{aligned} \tag{2.1}$$

where σ is the incubation rate, γ is the recovery rate, and δ is the disease-induced mortality rate. The transmission rate $\beta(t)$ varies between model variants to capture different behavioral assumptions, whereby the severity of the outbreak is assumed to lead to behavioral adaptation aimed at reducing infection risk.

Baseline model. The baseline model assumes constant transmission $\beta(t) = \beta_0$, representing traditional SEIRD dynamics without explicit behavioral feedback.

Behavioral models. Behavioral models incorporate dynamic feedback through mortality-dependent transmission, with $\beta(t)$ declining as deaths increase to reflect behavioral adaptation to perceived risk. The

transmission rate $\beta(t)$ depends on three quantities: β_0, the baseline transmission rate in the absence of behavioral response; $\delta I(t)$, the instantaneous death rate due to disease; and ζ, which quantifies the population's sensitivity to mortality. Larger values of ζ correspond to stronger behavioral responses. We consider three functional forms of $\beta(t)$:

(a) *Exponential form:*

$$\beta(t) = \beta_0\, e^{-\zeta\, \delta I(t)}, \tag{2.2}$$

where transmission rate decays exponentially with increasing mortality.

(b) *Rational form:*

$$\beta(t) = \frac{\beta_0}{1 + \zeta\, \delta I(t)}, \tag{2.3}$$

which produces similar behavior to the exponential form but saturates more slowly as deaths rise.

(c) *Mixed form:*

$$\beta(t) = \frac{\beta_0 e^{-\zeta\, \delta I(t)}}{1 + \zeta\, \delta I(t)}, \tag{2.4}$$

which incorporates both exponential and rational forms. This hybrid form saturates faster than either individual form as mortality increases.

All behavioral models reduce to the baseline model when $\zeta = 0$. Hereafter, we refer to these as the *Behavior (Exponential)*, *Behavior (Rational)*, and *Behavior (Mixed)* models.

Daily mortality. While individuals change behavior in response to instantaneous mortality δI, we compare model predictions to observed daily mortality data. Model-simulated daily mortality on day t is defined as:

$$M_{\text{sim}}(t) = D(t) - D(t-1), \tag{2.5}$$

representing the daily increment in cumulative deaths. The time series $\big(M_{\text{sim}}(t)\big)_t$ is compared to observed mortality data to estimate model parameters (Section 2.3).

Basic reproduction number. The *basic reproduction number*, defined as the average number of secondary infections produced by a typical infectious individual in a wholly susceptible population (i.e., $S(0) \approx N(0)$), is given by:

$$\mathcal{R}_0 = \frac{\beta_0}{\gamma + \delta}. \tag{2.6}$$

Effective reproduction number. The *effective reproduction number*, denoted by $\mathcal{R}_e(t)$, is the average number of new infections generated by a typical infected individual at time t [21]. The effective reproduction numbers for the baseline model, $\mathcal{R}_e^f(t)$, and behavioral model, $\mathcal{R}_e^b(t)$, are given respectively by:

$$\mathcal{R}_e^f(t) = \frac{\beta_0}{\gamma + \delta} \times \frac{S(t)}{N(t)} \quad \text{and} \quad \mathcal{R}_e^b(t) = \frac{\beta(t)}{\gamma + \delta} \times \frac{S(t)}{N(t)}. \tag{2.7}$$

Final epidemic size. The final epidemic size, denoted A, represents the total proportion of the population that becomes infected over the course of an outbreak:

$$A = \frac{R(\infty) + D(\infty)}{N}, \tag{2.8}$$

where $R(\infty)$ and $D(\infty)$ denote the cumulative recovered and deceased individuals as $t \to \infty$. Numerically, we compute A from simulated SEIRD trajectories by integrating until the infectious compartment falls below a threshold ($I(t) < 1$) and calculating $A = (N - S(t_{\text{end}}))/N$, where t_{end} is the time at which the epidemic effectively terminates.

2.2 Data and State Selection

We calibrated the models to COVID-19 mortality data from the COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University [22, 23]. Our analysis focused on the first wave of the pandemic in the United States, covering the 123-day period from March 1, 2020 to July 1, 2020.

State selection criteria. From the 51 states and territories in the dataset, we selected 30 locations for analysis based on three criteria designed to ensure data quality and the presence of a well-defined epidemic wave:

1. *Sufficient mortality signal:* States with very low daily death counts produced noisy 7-day moving averages unsuitable for reliable parameter inference (excluded: AK, HI, ID, ME, MT, ND, OR, SD, UT, VT, WV, WY).
2. *Clear single-wave structure:* States must exhibit a distinct epidemic wave with a recognizable peak and subsequent decline within the observation window (excluded: AL, AR, AZ, CA, FL, SC, TN, TX).
3. *Data integrity:* States with apparent reporting anomalies, such as sudden discontinuities in cumulative deaths, were excluded (excluded: CO).

The 30 selected locations were: Connecticut, Delaware, District of Columbia, Georgia, Illinois, Indiana, Iowa, Kansas, Kentucky, Louisiana, Maryland, Massachusetts, Michigan, Minnesota, Mississippi, Missouri, Nebraska, New Hampshire, New Jersey, New Mexico, Nevada, New York, North Carolina, Ohio, Oklahoma, Pennsylvania, Rhode Island, Virginia, Washington, and Wisconsin. Details on the 21 excluded locations and their mortality trajectories are provided in Supplementary Figure S1.

Data preprocessing. We applied a trailing 7-day moving average to daily mortality counts to reduce reporting artifacts.

Population data. Initial state population sizes $N(0)$ were obtained from the JHU CSSE dataset, ranging from 705,749 (District of Columbia) to 19,453,561 (New York).

2.3 Parameter Inference

We infer parameters using Approximate Bayesian Computation with Sequential Monte Carlo (ABC–SMC) [24], as implemented in the `pyABC` library [25], with the deterministic SEIRD model (2.1). The complete codebase for simulation, inference, and figure generation is available at `https://github.com/markolalovic/epi-behavior-models` [26]. For each location and model, we simulate daily deaths over the 123-day window and apply the same 7-day trailing average used for the observed data.

We reparameterize in epidemiological terms, inferring $\mathcal{R}_0$, the initial prevalence $\pi_0 = I_0/N$, mortality rate δ, and behavioral response parameter ζ, with weakly informative priors:

$$\mathcal{R}_0 \sim \mathcal{U}(1.2, 6.0), \quad \log \pi_0 \sim \mathcal{U}(\log 10^{-8}, \log 10^{-3}), \quad \log \delta \sim \mathcal{U}(\log 10^{-6}, \log 10^{-2}), \quad \zeta \sim \mathcal{U}(0, 0.05).$$

ABC–SMC uses 1000 particles per generation, a multivariate normal transition kernel adapted from the previous population, and a quantile-based tolerance schedule over up to 10 generations. Additional details are provided in Supplement Section S2.

2.4 Synthetic Data Experiments

To rigorously test whether behavioral effects can be recovered from mortality data and to quantify systematic biases in parameter estimation, we conducted controlled synthetic data experiments. We generated mortality trajectories from behavioral models using parameters calibrated to Massachusetts COVID-19 data, with varying values of the behavioral parameter ζ. To approximate observational variability, we added small independent Gaussian noise to each simulated series:

$$M_t^{\text{obs}} = M_t + \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{N}(0, \sigma^2), \qquad \sigma = 0.025 \max_t M_t.$$

Negative values were truncated to zero before applying the 7-day moving average. We then fitted both the data-generating behavioral model and the baseline model to these synthetic mortality datasets. This approach allows us to assess: (i) whether the behavioral model correctly recovers the known underlying dynamics; (ii) whether the baseline model produces systematic biases when behavioral feedback is present but ignored; and (iii) how the magnitude of behavioral response (controlled by varying ζ) affects the ability to distinguish between model structures. By knowing the true data-generating mechanism, we can definitively attribute any discrepancies in estimation of $\mathcal{R}_0$ and final epidemic size to model misspecification rather than data limitations or stochastic variation. Additional details are provided in Supplement Section S3.

2.5 Model Comparison

We evaluate the relative performance of baseline and behavioral models using Bayesian model selection via ABC–SMC. We describe below the distance metric and Bayes factor calculation used for model comparison.

Distance metric. To quantify the discrepancy between model-simulated and observed mortality trajectories, we use the *Normalized Sum of Squared Errors* (NSSE):

$$d_{\mathrm{NSSE}}(\boldsymbol{y}_{\mathrm{sim}}, \boldsymbol{y}_{\mathrm{obs}}) = \frac{\sum_{t=1}^{T} (\, y_{\mathrm{sim}}(t) - y_{\mathrm{obs}}(t)\,)^2}{\sum_{t=1}^{T} y_{\mathrm{obs}}(t)^2}\,, \tag{2.9}$$

where $y_{\mathrm{sim}}(t)$ is the model-simulated daily mortality time series and $y_{\mathrm{obs}}(t)$ is the observed 7-day averaged daily mortality series over the $T = 123$ day observation window.

The NSSE is the sum of squared errors normalized by a constant that depends only on the observed data for each location. This normalization makes the distance scale-invariant across locations with different mortality burdens, enabling comparable tolerance thresholds and performance metrics across all states while preserving the same ordering of parameter sets that would arise from the unnormalized sum of squared errors within any given state.

Bayes factor. We performed ABC–SMC model selection [24] comparing the baseline model against each of the three behavioral model formulations (exponential, rational, and mixed) separately, assigning equal prior probabilities to each pair. To avoid extinction artifacts, we report results at the first stabilized generation g^*, defined as the earliest g for which both model probabilities change by less than 0.02 over 3 consecutive generations and $P(\text{Baseline} \mid D, g) > 0$. At the stabilized generation g^*, we summarize the evidence for a behavioral model via the Bayes factor under equal model priors:

$$\mathrm{BF}_{\mathrm{eq}} = \frac{P(M_{\mathrm{Behavior}} \mid D, g^*)}{P(M_{\mathrm{Baseline}} \mid D, g^*)}.$$

We interpret $\mathrm{BF}_{\mathrm{eq}}$ following the scale from [27]. Values of $\mathrm{BF}_{\mathrm{eq}} > 1$ indicate evidence favoring the behavioral model over the baseline model, with strength categorized as: 1–3 (*very weak*), 3–20 (*positive*), 20–150 (*strong*), and > 150 (*very strong*). Values < 1 indicate evidence favoring the baseline model.

3 Results

We fitted COVID-19 mortality data from 30 US locations (29 states and the District of Columbia) during the first pandemic wave (March–July 2020) to a baseline SEIRD model with constant transmission rate and three behavioral SEIRD models where transmission rate declines in response to disease-induced mortality. We selected these 30 locations from the 50 states and DC based on data quality criteria including sufficient mortality signal, clear single-wave structure, and data integrity (see Section 2 for detailed selection criteria). Our analysis reveals three key findings: (i) behavioral models provide significantly improved fits to observed mortality data across almost all locations examined, with both median normalized squared error and Bayes factors favoring behavioral models in 28 of 30 locations; (ii) baseline models that ignore behavioral change systematically underestimate the basic reproduction number ($\mathcal{R}_0$), with median $\mathcal{R}_0$ from behavioral models

consistently higher than from baseline models across all states; and (iii) baseline models overestimate total infection burden despite underestimating transmissibility, creating a paradoxical dual bias with serious policy implications.

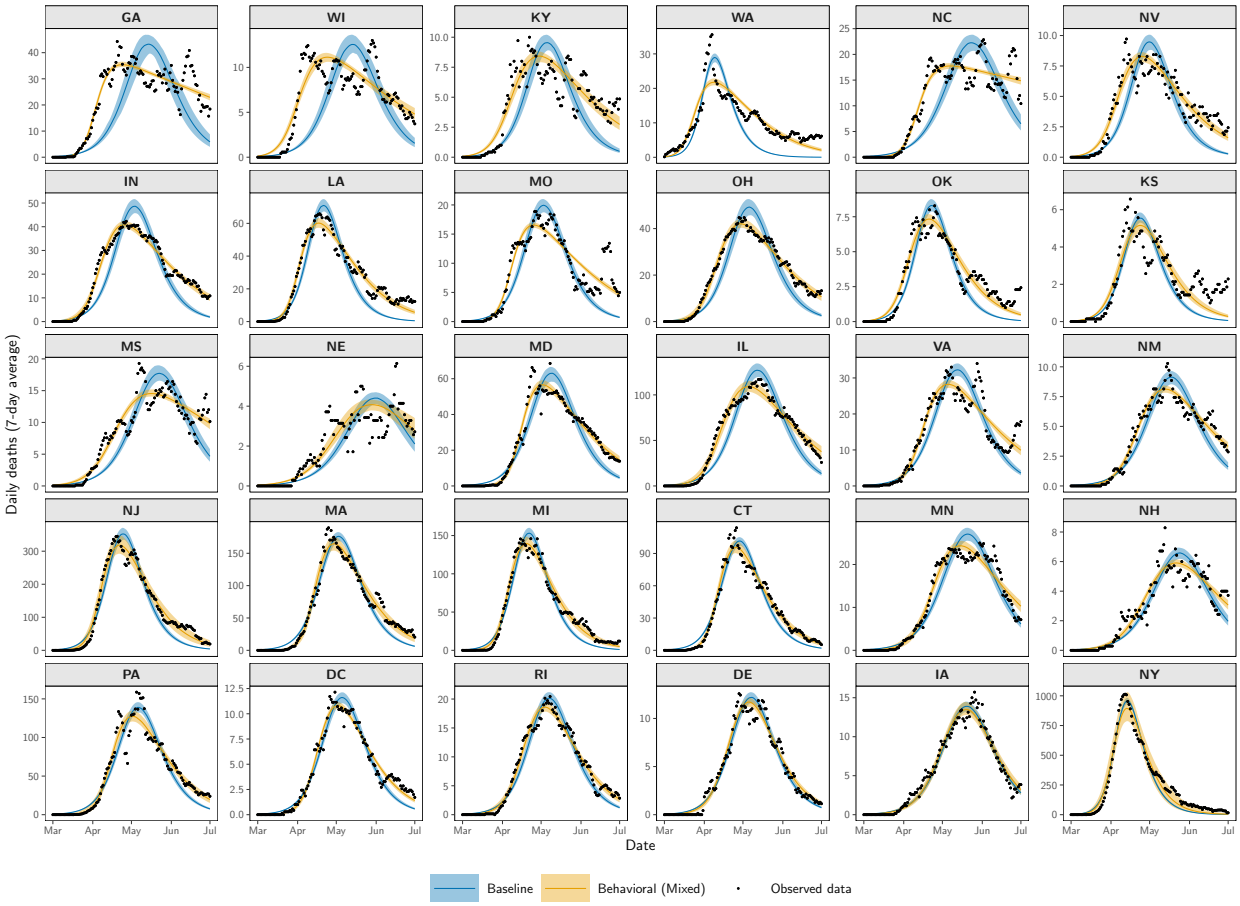

Figure 1: Behavioral models capture observed mortality dynamics better than baseline models across 30 US locations. Posterior predictive fits comparing baseline and behavioral (mixed) models to 7-day averaged daily COVID-19 deaths during the first pandemic wave (March–July 2020). Black points show observed mortality data. Solid lines show posterior predictive medians and shaded regions show 90% credible intervals (5th–95th percentiles), computed from 1,000 weighted posterior samples drawn from the final ABC–SMC generation. The behavioral model accurately captures the characteristic asymmetric mortality curves with sharp initial rises and gradual extended declines observed across most locations. States are ordered from left to right and top to bottom by decreasing improvement in fit: those where the behavioral model outperforms the baseline model most dramatically (largest difference in median NSSE) appear first, progressing to states where the two models perform most similarly. This ordering visually emphasizes that behavioral feedback provides substantial improvements in model fit across the vast majority of locations examined.

3.1 Behavioral Models Show Superior Fit to COVID-19 Mortality Data

All three behavioral models (exponential, rational, and mixed) consistently outperformed the baseline model when fitted to COVID-19 mortality data from 30 locations during the first pandemic wave. Visually, the behavioral models capture the characteristic asymmetric mortality curves observed in many locations—sharp initial rises followed by gradual, extended declines—while the baseline model produces curves that decay too rapidly (Figure 1 shows the mixed behavioral model; Supplementary Figures S4 and S5 show exponential and rational formulations, respectively).

This superior performance is confirmed quantitatively across all three formulations. For the mixed behavioral

model, we observed lower median normalized sum of squared errors (NSSE) in 28 of 30 locations, with Bayesian model selection via ABC–SMC providing evidence for the mixed behavioral model in these same 28 locations (Table 1). Only Iowa and New York favored the baseline model. The exponential and rational formulations of behavioral models demonstrated comparable performance (see Supplementary Tables S1 and S2). Overall, across all three behavioral formulations, at least 28 of 30 locations favored behavioral models over baseline models based on NSSE and Bayes factor.

The superior performance of behavioral models varies systematically across locations, with the largest improvements in states exhibiting extended mortality decline phases. Figure 1 orders states by decreasing difference in median NSSE between baseline and behavioral models, with locations showing the largest improvement appearing first. States like Washington and Indiana (top left) exhibit the most striking contrast—the behavioral models capture gradual, extended mortality declines that baseline models fail to reproduce. Moving toward Rhode Island and Delaware (bottom right), the performance gap narrows, with both model types achieving similar fit quality.

Sensitivity analysis reveals a key feature of behavioral models: increasing the behavioral response parameter (ζ) produces lower peak mortality and slower decline in daily deaths following the peak (Supplementary Figure S8), a mechanism absent in baseline models. This pattern is reminiscent of the "flattening the curve" phenomenon, where public health interventions reduce peak mortality while extending the duration of the outbreak. In all three behavioral formulations, $\beta(t)$ exhibits a characteristic pattern: sharp decline during the mortality surge followed by gradual recovery as deaths decline (see Supplementary Section S6). This time-varying transmission rate better reflects the reality of behavioral adaptation during epidemics, consequently producing superior fits to observed mortality data compared to the constant transmission rate assumption of the baseline model.

We validated this finding through controlled synthetic data experiments where mortality trajectories were generated from behavioral models with varying levels of behavioral sensitivity (ζ), then fitted to both the data-generating behavioral model and the baseline model. Using the mixed behavioral formulation, behavioral models accurately recovered the true dynamics across all ζ values, while the baseline model increasingly failed to capture mortality trajectories as behavioral feedback strengthened (Figure 3A). Results were consistent across all three functional forms examined (Supplementary Figures S2A and S3A show exponential and rational forms, respectively), demonstrating that baseline models systematically mischaracterize epidemic trajectories when behavioral responses are present.

3.2 Baseline Models Systematically Underestimate the Basic Reproduction Number

Baseline models that ignore behavioral change systematically underestimate the basic reproduction number ($\mathcal{R}_0$) compared to all three behavioral models when fitted to the same COVID-19 mortality data across 30 US locations (Figure 2A shows comparison with mixed formulation). Furthermore, $\mathcal{R}_0$ estimates from all three behavioral model variants are similar to each other, and the systematic underestimation by baseline models is consistent across exponential, rational, and mixed formulations (Supplementary Figures S6 and S7 show $\mathcal{R}_0$ comparisons with exponential and rational forms, respectively).

The effective reproduction number for all four models is provided in Supplementary Section S7. In all cases, $\mathcal{R}_e(t)$ converges to values below 1 by late April–May across all states (coinciding with peak infection), though behavioral models exhibit steeper initial declines corresponding to rapid behavioral responses during the early stages of the outbreak.

Synthetic experiments where true $\mathcal{R}_0$ was known by design confirmed this systematic underestimation. Baseline models consistently underestimated true $\mathcal{R}_0$ across all behavioral sensitivity values, with bias increasing as behavioral feedback strengthened (increasing ζ), while behavioral models accurately recovered true $\mathcal{R}_0$ across all ζ values (Figure 3B for mixed; Supplementary Figures S2B and S3B for exponential and rational forms). This confirms that $\mathcal{R}_0$ underestimation is a systematic consequence of model misspecification.

Intuitively, this systematic underestimation of $\mathcal{R}_0$ arises because baseline and behavioral models, although fitted to the same observed mortality data, assume fundamentally different transmission mechanisms. The

Table 1: Comparison of the baseline model and the behavioral (mixed) model across 30 U.S. states. For each state, Median $\mathrm{NSSE_B}$ and Median $\mathrm{NSSE_M}$ denote the median posterior normalized sum of squared errors (NSSE) under the baseline B and behavioral (mixed) M models, respectively. $\mathrm{BF_{eq}}$ is the Bayes factor under equal model priors in favor of the mixed model versus the baseline model at the stabilized ABC–SMC generation, and "Evidence" gives the corresponding verbal strength of evidence. See Section 2 for interpretation guidelines for $\mathrm{BF_{eq}}$.

State	Median $\mathrm{NSSE_B}$	Median $\mathrm{NSSE_M}$	$\mathrm{BF_{eq}}$	Evidence
DC	0.0285	**0.0131**	11551.24	very strong
MS	0.0882	**0.0337**	5051.63	very strong
NH	0.0571	**0.0381**	2105.67	very strong
NE	0.1390	**0.0859**	1852.99	very strong
RI	0.0250	**0.0111**	852.43	very strong
NM	0.0488	**0.0151**	688.87	very strong
MO	0.1214	**0.0539**	211.95	very strong
KY	0.2154	**0.0586**	210.55	very strong
NC	0.1440	**0.0310**	164.63	very strong
IN	0.0896	**0.0086**	126.43	strong
CT	0.0330	**0.0121**	121.80	strong
WI	0.2540	**0.0558**	105.04	strong
VA	0.0805	**0.0387**	100.58	strong
MN	0.0392	**0.0188**	79.31	strong
GA	0.2254	**0.0265**	66.31	strong
WA	0.1786	**0.0655**	64.52	strong
OH	0.0690	**0.0060**	37.48	strong
IL	0.0572	**0.0110**	35.61	strong
OK	0.1005	**0.0394**	28.71	strong
MA	0.0374	**0.0145**	22.96	strong
PA	0.0471	**0.0311**	19.86	positive
LA	0.0842	**0.0165**	13.46	positive
MD	0.0621	**0.0151**	11.97	positive
MI	0.0344	**0.0123**	11.80	positive
DE	0.0210	**0.0167**	10.67	positive
NJ	0.0433	**0.0157**	3.67	positive
NV	0.1456	**0.0359**	1.47	very weak
KS	0.1627	**0.1054**	1.34	very weak
IA	**0.0141**	0.0169	0.51	favors Baseline
NY	**0.0277**	0.0316	0.02	favors Baseline

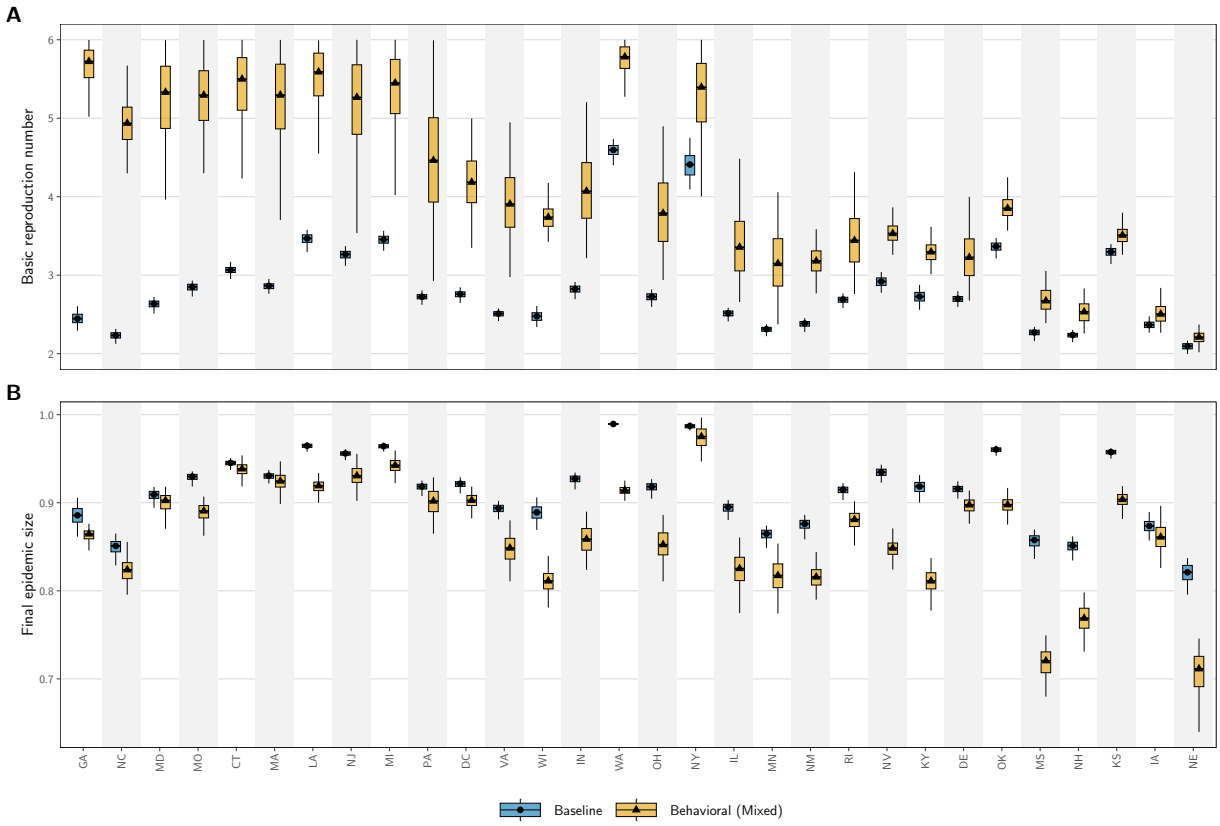

Figure 2: Baseline models systematically underestimate basic reproduction number while overestimating final epidemic size across all 30 locations. Posterior distributions of key epidemiological quantities estimated by fitting the baseline and behavioral (mixed) models to COVID-19 mortality data from 30 US locations during the first pandemic wave (March–July 2020). **(A)** Basic reproduction number $\mathcal{R}_0$. The behavioral model consistently estimates higher $\mathcal{R}_0$ values than the baseline model across all 30 locations, indicating that models ignoring behavioral feedback systematically underestimate intrinsic transmissibility. **(B)** Final epidemic size (proportion of population infected). Despite estimating lower $\mathcal{R}_0$, the baseline model predicts larger cumulative infection burdens than the behavioral model in most locations, demonstrating the paradoxical consequence of ignoring behavioral adaptation. Boxplots show median (center line), interquartile range (box), and whiskers extending to $1.5 \times$ IQR within the sample support computed from weighted posterior samples. States are ordered from left to right by decreasing difference in median basic reproduction number between the behavioral (mixed) and baseline models.

baseline model uses a constant transmission rate β_0 throughout the epidemic, fitting the data with a moderate value that remains fixed. In contrast, behavioral models assume a time-varying transmission rate $\beta(t)$ that declines as people adapt their behavior in response to rising mortality. To reproduce the same initial epidemic growth observed in the data, behavioral models require a substantially higher initial transmission rate β_0—which then declines as behavioral responses take effect. As evident from Equation (2.6), $\mathcal{R}_0$ is an increasing function of β_0; hence models that estimate higher values of β_0 necessarily yield higher estimates of $\mathcal{R}_0$. Thus, behavioral models, which require higher β_0 to account for subsequent behavioral adaptation, naturally produce higher $\mathcal{R}_0$ estimates than baseline models. This higher initial transmission rate captured by behavioral models represents the intrinsic transmissibility of the pathogen before any behavioral adaptation occurs. The baseline model, lacking this behavioral feedback mechanism, systematically underestimates this initial transmissibility because it interprets the entire epidemic trajectory—including periods when behavior has already adapted—through the lens of a single, constant transmission rate.

3.3 Ignoring Behavioral Responses Overestimates Infection Burden

Despite underestimating $\mathcal{R}_0$, baseline models overestimate the total proportion of the population that becomes infected (Figure 2B shows comparison with mixed; Supplementary Figures S6B and S7B show comparisons with exponential and rational forms, respectively). This overestimation is most pronounced in locations like Illinois, where baseline models predict median final epidemic sizes over 10 percentage points higher than behavioral models. The three behavioral variants produce similar final size estimates across all 30 locations, while baseline models consistently overestimate infection burden compared to all three behavioral models.

Synthetic experiments confirmed this systematic overestimation of final epidemic size. Baseline models consistently overestimated true final epidemic sizes while behavioral models accurately recovered them across all ζ values considered (Figure 3C shows comparison with mixed; Supplementary Figures S2C and S3C show comparisons with exponential and rational forms, respectively). The bias in overestimating final epidemic size increases with stronger behavioral responses, demonstrating that ignoring behavioral dynamics leads to inflated infection burden estimates regardless of functional form.

To gain theoretical insight into the relationship between behavioral feedback and epidemic outcomes, we analytically examine final epidemic size under a controlled scenario where $\mathcal{R}_0$ is identical for both the baseline and a behavioral model. For analytical tractability, we assume a constant total population $N = S+E+I+R+D$, which differs from the time-varying living population $N(t) = N(0) - D(t)$ used in our numerical simulations. We prove that for any fixed value of $\mathcal{R}_0$, behavioral models incorporating mortality-driven feedback produce final epidemic sizes less than or equal to those from baseline models. We first demonstrate this result for the exponential behavioral model.

Theorem 3.1 (Final Size Reduction: Exponential Form)**.** *Consider an epidemic model with behavioral response function $\beta(t) = \beta_0 \exp(-\alpha I(t))$ where $\alpha = \zeta\delta$. Furthermore, we assume that the total population size is fixed with $N = S + E + I + R + D$. Let A_0 denote the final size of the baseline model (i.e., $\alpha = 0$). The final size of the behavioral model, A_α, can be expressed as:*

$$A_\alpha = A_0 + \alpha a_1 + O(\alpha^2),$$

where $a_1 < 0$ is the first-order correction to the final epidemic size due to behavioral feedback. Hence, for small α, the final size of the behavioral model is smaller than that of the baseline model.

The proof of Theorem 3.1, based on asymptotic expansion, is given in Appendix A.

This result extends beyond the exponential form. For any behavioral response function satisfying $0 < \beta(t) \leq \beta_0$—including the exponential, rational, and mixed forms considered in this study—we show that behavioral models always produce smaller or equal final epidemic sizes compared to baseline models when $\mathcal{R}_0$ is the same for both models.

Theorem 3.2 (Final Size Reduction: General Form)**.** *Consider the SEIRD model* (2.1) *where $N = S + E + I + R + D$ is the fixed total population size and the transmission rate $\beta_\alpha(t) = \beta_0 \Psi(\alpha, I_\alpha(t))$ decreases with increasing mortality through a behavioral response function satisfying $0 < \Psi(\alpha, I_\alpha(t)) \leq 1$, with β_0 being the maximum transmission rate and $\alpha = \zeta\delta$. If a behavioral model (with $\alpha > 0$) and the baseline model (with $\alpha = 0$ and constant transmission β_0) have equal basic reproduction numbers, given by $\mathcal{R}_0 = \beta_0/(\gamma + \delta)$, then the final epidemic size of the baseline model (A_0) is always larger than or equal to that of the behavioral model (A_α).*

The proof of Theorem 3.2 is given in Appendix B. We obtain final size relationships for the baseline model with constant transmission rate β_0 and for behavioral models with $\beta(t) < \beta_0$. By defining an auxiliary function based on these final size equations and exploiting its properties, including its unique maximum and concavity, we show that the final size of the behavioral model is always less than or equal to that of the baseline model when both share the same basic reproduction number $\mathcal{R}_0$.

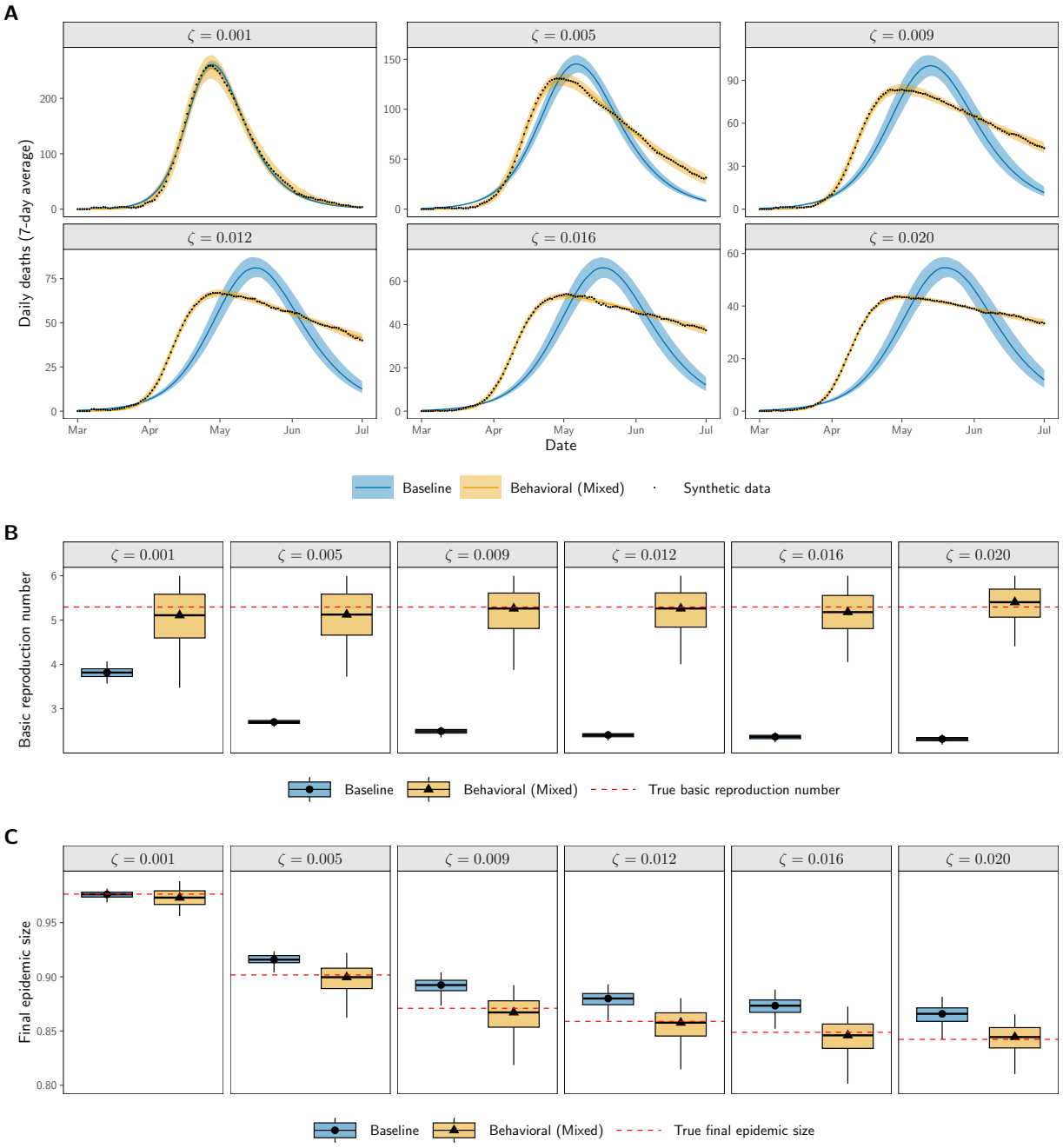

Figure 3: Baseline models systematically misestimate epidemic quantities when behavioral feedback is present but ignored. Synthetic data experiments where mortality trajectories were generated using the behavioral (mixed) model with known parameters and varying behavioral sensitivity ζ, then fitted with both baseline and behavioral models. **(A)** Posterior predictive fits to synthetic mortality data (black points). Solid lines show posterior medians and shaded regions show 90% credible intervals. The behavioral (mixed) model accurately recovers the true dynamics across all ζ values, while the baseline model increasingly fails to capture the extended post-peak decline as behavioral feedback strengthens with increasing ζ values. **(B)** Posterior distributions of the basic reproduction number $\mathcal{R}_0$. The behavioral model correctly recovers the true $\mathcal{R}_0$ (red dashed line) across all scenarios, while the baseline model systematically underestimates $\mathcal{R}_0$, with bias increasing as ζ increases. **(C)** Posterior distributions of final epidemic size. The behavioral model accurately estimates the true final size (red dashed line), while the baseline model systematically overestimates infection burden despite underestimating transmissibility.

4 Discussion

Our analysis of COVID-19 mortality data from 30 US locations during the first pandemic wave reveals systematic biases that arise when epidemic models fail to account for behavioral responses to disease burden. Across all three behavioral models, at least 28 of 30 locations favored behavioral models—where transmission rates decline as mortality increases—over the traditional model with a constant transmission rate, based on both NSSE and Bayes factor. More critically, we identified a problematic paradox: models that ignore behavioral change systematically underestimate pathogen transmissibility while simultaneously overestimating total infection burden. We confirmed these biases through controlled synthetic experiments where the true dynamics were known by design, demonstrating that the discrepancies arise from model misspecification rather than data quality or stochastic effects. These findings held across three distinct mathematical formulations of behavioral models, strengthening confidence in our conclusions.

The dual bias we identify creates dangerous implications for public health decision-making. Models that ignore behavioral responses underestimate transmissibility as measured by the basic reproduction number, potentially delaying critical early interventions when aggressive action matters most. Yet these same models overestimate total infections, leading to excessive resource allocation for later epidemic phases that behavioral adaptation may prevent. Decision-makers relying on such models face a systematic mismatch: insufficiently aggressive early response combined with overpreparation for infection burdens that may not materialize. This paradox can result in both under-intervention during critical early phases and misallocation of resources for scenarios that behavioral change prevents.

The mechanism underlying these biases is intuitive once recognized. Baseline models with constant transmission rates fit observed mortality data using relatively modest transmission parameters because the transmission rate remains fixed throughout the entire epidemic wave. In contrast, behavioral models require substantially higher initial transmission rates to fit the same mortality data precisely because transmission declines as mortality increases and people adapt their behavior. This difference directly manifests in the basic reproduction number: the baseline model settles on a lower value that suffices under constant transmission, while the behavioral model correctly identifies the higher intrinsic transmissibility that characterized the pathogen before behavioral responses took effect. The bias in final epidemic size follows naturally from this same mechanism. When populations reduce contact rates in response to rising deaths—through actions like avoiding crowds, wearing masks, or increasing hygiene—fewer people ultimately become infected compared to scenarios where behavior remains unchanged.

This study has several important limitations. First, our models assume that behavioral changes are driven exclusively by location-specific reported mortality, whereas behavior in reality may also respond to mortality signals from nearby regions, national trends, policy announcements, or media coverage [6]. Second, using mortality as a proxy for public risk perception does not capture the complex array of influences on behavior, including policy actions, misinformation, informal information channels, and individual risk tolerance. This simplification may also introduce delays between changes in actual risk and behavioral adaptation, which our instantaneous feedback formulation does not capture [28, 29].

Third, we assume homogeneous populations within each location, ignoring substantial variation in risk tolerance, contact patterns, and protective behavior adoption across age groups, socioeconomic strata, and communities. Real populations exhibit heterogeneous responses to epidemic threats, with some groups adopting protective measures rapidly while others respond more slowly or not at all [6, 28–30]. Fourth, our analysis focuses exclusively on the first pandemic wave (March–July 2020), when behavioral responses were likely strongest due to novelty and uncertainty. Behavioral patterns may differ substantially in later waves as pandemic fatigue sets in or as populations develop more nuanced risk assessments [6].

Fifth, our models imply that extended post-peak mortality declines reflect stronger behavioral responses, but this relationship does not hold universally. New York experienced one of the earliest and most severe outbreaks [31–33] with stringent public health measures and well-documented behavioral changes [6, 34, 35], yet the mortality curve rose and fell relatively symmetrically without the prolonged tail our models associate with strong behavioral feedback. This disconnect suggests that the temporal signature of behavioral adaptation depends on factors we do not explicitly capture, such as the timing and stringency of policy interventions,

underlying contact network structures, population density, or the rapidity with which protective behaviors were adopted. Locations with sharp mortality declines might reflect either weak behavioral responses or, conversely, such rapid behavioral adaptation that transmission was quickly suppressed. Our models cannot distinguish between these possibilities, which limits our ability to infer behavioral mechanisms from mortality data alone. Finally, our behavioral models assume a specific functional relationship between mortality and transmission reduction. While we tested three different functional forms, real behavioral responses may follow more complex patterns involving thresholds, saturation effects, or time-varying sensitivity to disease burden.

Future work should address these limitations through several avenues. Incorporating spatial coupling between locations could capture how mortality in nearby regions influences local behavior [36, 37]. Integrating data on policy interventions, mobility patterns, and direct behavioral surveys could provide more nuanced representations of behavioral change beyond mortality-driven responses. Age-structured models could account for heterogeneous contact patterns and risk perceptions across demographic groups [21]. Multi-wave analyses could reveal how behavioral responses evolve over the course of a pandemic as novelty effects diminish and risk perceptions stabilize. Comparing fits to multiple data streams—cases, hospitalizations, and deaths—could help disentangle behavioral responses to different epidemic indicators and their respective time lags.

More fundamentally, developing frameworks that explicitly model the cognitive and social processes underlying behavioral change—such as risk perception dynamics, information diffusion, and social norm formation—could provide deeper mechanistic understanding beyond phenomenological feedback functions [38]. Such frameworks could better predict behavioral responses to novel pathogens and inform more effective risk communication strategies.

Severing the coupling between pathogen transmission and human behavior produces models that systematically misrepresent reality. Infectious disease outbreaks unfold through feedback: transmission shapes behavior, and behavior alters transmission. Models that assume constant contact rates sacrifice this coupling for mathematical convenience, producing contradictory estimates of key epidemiological quantities—underestimating pathogen transmissibility while overestimating infection burden. These structural failures cascade through every decision that relies on model projections, from intervention timing to resource allocation.

Our findings demonstrate that incorporating behavioral dynamics is not merely a modeling refinement but essential for accurate epidemic parameter estimation and evidence-based pandemic response. Across 30 geographically and demographically diverse US locations, behavioral models consistently provided superior fits to observed data and more plausible estimates of both transmissibility and final epidemic size. The analytical results further establish that this dual bias is an inevitable mathematical consequence of ignoring behavioral feedback when it is present in the data. Future pandemic preparedness must prioritize the development and deployment of models that explicitly account for the dynamic interplay between disease transmission and human behavioral responses.

Software and Data Availability. The complete codebase used for simulation, inference, and figure generation is available at GitHub repository: `https://github.com/markolalovic/epi-behavior-models` [26]. The COVID-19 mortality data used in this study are publicly available from the COVID-19 Data Repository maintained by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University: `https://github.com/CSSEGISandData/COVID-19` [22, 23].

Acknowledgments. BP and MS were supported by cooperative agreement CDC-RFA-FT-23-0069 from the CDC's Center for Forecasting and Outbreak Analytics. The views expressed in written in this publication do not necessarily reflect the official policies of the Department of Health and Human Services/Centers for Disease Control and Prevention. ML acknowledges the PhD studentship support from Northeastern University London.

A Proof of Theorem 3.1

Proof. Let $\beta_\alpha(t) = \beta_0 \exp(-\alpha I_\alpha(t))$, where $\alpha = \zeta\delta$ combines the behavioral sensitivity parameter ζ with the mortality rate δ, and $I_\alpha(t)$ denotes the infectious population trajectory corresponding to α. Furthermore, we assume that the total population size is fixed with $N = S + E + I + R + D$.

The *final epidemic size* corresponding to α is defined as

$$A_\alpha := R(\infty) + D(\infty) = (\gamma + \delta) \int_0^\infty I_\alpha(t)\, dt. \tag{A.1}$$

From the S–equation of model (2.1), we have

$$\frac{dS}{dt} = -\frac{\beta_\alpha(t)}{N} S(t) I_\alpha(t).$$

Dividing both sides by $S(t)$ and integrating in time from 0 to ∞ (noting $S(t) > 0$ for all finite time until extinction):

$$\ln\left(\frac{S(\infty)}{S(0)}\right) = -\frac{1}{N} \int_0^\infty \beta_\alpha(t) I_\alpha(t)\, dt. \tag{A.2}$$

Since the removed individuals at infinity equal the initial population minus susceptibles at infinity, we have

$$S(\infty) = N - A_\alpha, \qquad S(0) = S_0,$$

and hence, using $\beta_\alpha(t) = \beta_0 \exp(-\alpha I_\alpha(t))$, (A.2) can be rewrriten as

$$\ln\left(\frac{N - A_\alpha}{S_0}\right) = -\frac{\beta_0}{N} \int_0^\infty e^{-\alpha I_\alpha(t)} I_\alpha(t)\, dt. \tag{A.3}$$

This holds for all $\alpha \geq 0$. When $\alpha = 0$ (i.e. $\beta(t) \equiv \beta_0$), we have

$$\ln\left(\frac{N - A_0}{S_0}\right) = -\frac{\beta_0}{N} \int_0^\infty I_0(t)\, dt = -\frac{\beta_0}{N(\gamma + \delta)} A_0,$$

with A_0 and $I_0(t)$ denoting the final size and infectious population corresponding to the baseline model. Hence, the classical implicit final-size equation:

$$\ln\left(\frac{N - A_0}{S_0}\right) = -\frac{\beta_0}{N(\gamma + \delta)} A_0. \tag{A.4}$$

Small-α expansion. We now expand for small $\alpha > 0$. Let

$$A_\alpha = A_0 + \alpha a_1 + O(\alpha^2), \qquad I_\alpha(t) = I_0(t) + \alpha J(t) + O(\alpha^2), \tag{A.5}$$

where, a_1 and $J(t)$ are the first-order corrections to the final epidemic size and infectious trajectory, respectively.

We expand both sides of (A.3) using the Taylor series.

Left-hand side expansion. We expand $\ln\left(\frac{N - A_\alpha}{S_0}\right)$ using Taylor series around $A = A_0$:

$$\begin{aligned}
\ln\left(\frac{N - A_\alpha}{S_0}\right) &= \ln\left(\frac{N - A_0}{S_0}\right) + \frac{d}{dA} \ln \frac{N - A}{S_0}\bigg|_{A_0} (A_\alpha - A_0) + O((A_\alpha - A_0)^2) \\
&= \ln\left(\frac{N - A_0}{S_0}\right) + \frac{d}{dA} \ln \frac{N - A}{S_0}\bigg|_{A_0} (\alpha a_1) + O(\alpha^2) \\
&= \ln\left(\frac{N - A_0}{S_0}\right) - \frac{\alpha a_1}{N - A_0} + O(\alpha^2).
\end{aligned} \tag{A.6}$$

Right-hand side expansion. Using $e^{-\alpha I_0} = 1 - \alpha I_0 + O(\alpha^2)$ and I_α from (A.5), we expand:

$$e^{-\alpha I_\alpha} I_\alpha = (1 - \alpha I_0 + O(\alpha^2))(I_0 + \alpha J + O(\alpha^2)) = I_0 + \alpha(J - I_0^2) + O(\alpha^2). \tag{A.7}$$

We expand the internal term in (A.3) using (A.7) as:

$$\int_0^\infty e^{-\alpha I_\alpha} I_\alpha \, dt = \int_0^\infty I_0 \, dt + \alpha \left(\int_0^\infty J \, dt - \int_0^\infty I_0^2 \, dt \right) + O(\alpha^2).$$

Plugging in the expansion of A_α and $I_\alpha(t)$ from (A.5) into (A.1) and comparing α terms gives:

$$\int_0^\infty J(t) \, dt = \frac{a_1}{\gamma + \delta}.$$

Hence, the right-hand side of (A.3) can be finally expanded as:

$$-\frac{\beta_0}{N} \int_0^\infty e^{-\alpha I_\alpha} I_\alpha \, dt = -\frac{\beta_0}{N} \frac{A_0}{\gamma + \delta} - \frac{\beta_0 \alpha}{N} \left(\frac{a_1}{\gamma + \delta} - \int_0^\infty I_0^2 \, dt \right) + O(\alpha^2). \tag{A.8}$$

Matching terms. Now we compare the matching terms of left-hand side expansion, given in (A.6), with right-hand side expansion, given in (A.8). At order $O(1)$, we recover the classical relation (A.4).

At order $O(\alpha)$ we have:

$$-\frac{a_1}{N - A_0} = -\frac{\beta_0}{N} \left(\frac{a_1}{\gamma + \delta} - \int_0^\infty I_0^2 \, dt \right),$$

which can be rewritten as:

$$\left(\frac{1}{N - A_0} - \frac{\beta_0}{N(\gamma + \delta)} \right) a_1 = -\frac{\beta_0}{N} \int_0^\infty I_0(t)^2 \, dt.$$

Therefore, solving for a_1 gives:

$$a_1 = -\frac{\frac{\beta_0}{N} \int_0^\infty I_0(t)^2 \, dt}{\frac{1}{N - A_0} - \frac{\beta_0}{N(\gamma + \delta)}}. \tag{A.9}$$

Since $A_\alpha = A_0 + \alpha a_1 + O(\alpha^2)$, and A_0 corresponds to the final size of the baseline model, to show that the final size of the behavioral model is smaller than that of the baseline model, it is sufficient to show that $a_1 < 0$. Since we consider small α, the $O(\alpha^2)$ terms are negligible compared to the first-order term αa_1, so the sign of a_1 determines whether $A_\alpha < A_0$ or $A_\alpha > A_0$.

The numerator of (A.9) is always positive. The denominator of (A.9) is positive provided:

$$N(\gamma + \delta) > \beta_0 (N - A_0),$$

which can be rewritten as:

$$\mathcal{R}_0 \left(1 - \frac{A_0}{N} \right) < 1,$$

which is true since $s_\infty = (1 - \frac{A_0}{N}) < \frac{1}{\mathcal{R}_0}$ (see (B.9) for details). This concludes the proof. □

B Proof of Theorem 3.2

Proof. Let $\beta_\alpha(t) = \beta_0\Psi(\alpha, I_\alpha(t))$, where $\alpha = \zeta\delta$ combines the behavioral sensitivity parameter ζ with the mortality rate δ, and $I_\alpha(t)$ denotes the infectious population trajectory corresponding to α. For example, $\Psi(\alpha, I_\alpha(t)) = \exp(-\zeta\delta I_\alpha(t)) = \exp(-\alpha I_\alpha(t))$. We require that $0 < \Psi(\alpha, I(t)) \leq 1$, and that $\Psi(0, I(t)) = \beta_0$. We assume that the total population size is fixed with $N = S+E+I+R+D$. Furthermore, initial conditions are: $S(0) = S_0 > 0$, $E(0) = E_0 \geq 0$, $I(0) = I_0 \geq 0$, $R(0) = D(0) = 0$ and $S_0 + E_0 + I_0 = N$. The *final epidemic size* is defined as

$$A_\alpha := R(\infty) + D(\infty) = (\gamma + \delta) \int_0^\infty I_\alpha(t)\, dt. \tag{B.1}$$

From the S–equation of model (2.1), we have

$$\frac{dS}{dt} = -\frac{\beta_\alpha(t)}{N} S(t) I_\alpha(t).$$

Dividing both sides by $S(t)$ and integrating in time from 0 to ∞ (noting $S(t) > 0$ for all finite time until extinction):

$$\ln\left(\frac{S(\infty)}{S(0)}\right) = -\frac{1}{N}\int_0^\infty \beta_\alpha(t) I_\alpha(t)\, dt. \tag{B.2}$$

Since removed individuals at infinity equal the initial population minus susceptibles at infinity, we have

$$S(\infty) = N - A_\alpha, \qquad S(0) = S_0,$$

and hence (B.2) can be rewrriten as

$$\ln\left(\frac{N - A_\alpha}{S_0}\right) = -\frac{1}{N}\int_0^\infty \beta_\alpha(t) I_\alpha(t)\, dt. \tag{B.3}$$

Equation (B.3) is exact for any time-dependent Ψ and any solution of the SEIRD model (2.1) satisfying the stated assumptions. Let A_0 denote the final size when $\Psi(\alpha = 0, I(t)) = \beta_0$.

Recall the pointwise inequality assumption

$$0 < \Psi(\alpha, I_\alpha(t)) \leq 1, \qquad \forall t \geq 0,\ \alpha \geq 0,$$

The above implies

$$\beta_\alpha(t) = \beta_0\Psi(\alpha, I_\alpha(t)) I_\alpha(t) \leq \beta_0 \qquad \text{for all } t.$$

Substituting inequality of $\beta_\alpha(t)$ in (B.3) gives:

$$\ln\left(\frac{N - A_\alpha}{S_0}\right) \geq -\frac{1}{N}\int_0^\infty \beta_0 I_\alpha(t)\, dt. \tag{B.4}$$

For the baseline case with $\alpha = 0$ (no behavioral response), we have $\beta_\alpha(t) = \beta_0$ and the infectious trajectory $I_\alpha(t)$ reduces to $I_0(t)$. Substituting $\beta_\alpha(t) = \beta_0$ in (B.3) yields:

$$\ln\left(\frac{N - A_0}{S_0}\right) = -\frac{\beta_0}{N}\int_0^\infty I_0(t)\, dt. \tag{B.5}$$

Equations (B.4) and (B.5) cannot be directly compared because they involve both different final epidemic sizes on the left hand side (A_α versus A_0) and different infectious population trajectories on the right hand side ($I_\alpha(t)$ versus $I_0(t)$). To enable comparison, we apply the relation $\int_0^\infty I_\alpha(t)\, dt = A_\alpha/(\gamma+\delta)$ from equation (B.1), which transforms inequality (B.4) into:

$$\ln\left(\frac{N - A_\alpha}{S_0}\right) \geq -\frac{\beta_0}{N} \cdot \frac{A_\alpha}{\gamma + \delta}, \tag{B.6}$$

and for baseline case, equation (B.5) can rewritten as:

$$\ln\left(\frac{N-A_0}{S_0}\right) = -\frac{\beta_0}{N}\cdot\frac{A_0}{\gamma+\delta}. \tag{B.7}$$

We define the function

$$f(A) := \ln\left(\frac{N-A}{S_0}\right) + \frac{\beta_0}{N(\gamma+\delta)}A. \tag{B.8}$$

Notice that for the baseline model $f(A_0) = 0$ while for behavioral models $f(A_\alpha) \geq 0$. $f(A)$ has two zeros — the trivial solution $A = 0$ (no epidemic) and the nontrivial solution $A = A_0$ (baseline epidemic). The proof proceeds in three steps. First, we show $f(A)$ is strictly concave and compute its unique maximum at $A^* = N(1-1/\mathcal{R}_0)$. Second, since A_0 is defined only implicitly through a transcendental equation (B.7), we cannot directly compare A_0 with A^*. We therefore normalize the susceptible population as $s = S/N$ and express the baseline final size through the transcendental equation $\ln s_\infty = -\mathcal{R}_0(1-s_\infty)$, where $s_\infty = (N-A_0)/N$. By analyzing the auxiliary function $\Phi(s) := \ln s + \mathcal{R}_0(1-s)$, we show that it attains a unique maximum at $s = 1/\mathcal{R}_0$. Furthermore, we prove that $s_\infty < 1/\mathcal{R}_0$, which translates to $A_0 > A^*$. Third, combining $A_0 > A^*$ with the concavity of $f(A)$ and the inequality $f(A_\alpha) \geq 0$, we conclude $A_\alpha \leq A_0$. Details follow.

Differentiating $f(A)$ with respect to A gives (where $c = \frac{\beta_0}{N(\gamma+\delta)}$ is a constant):

$$f'(A) = -\frac{1}{N-A} + c, \qquad f''(A) = -\frac{1}{(N-A)^2} < 0 \quad \text{for } 0 \leq A < N.$$

Thus f is strictly concave on $[0, N)$ and can have at most one critical point (which will be a global maximum).

Solving $f'(A) = 0$ yields $-\frac{1}{N-A} + c = 0$, which gives $N - A = \frac{1}{c}$. Thus the unique critical point is

$$A^* = N - \frac{1}{c},$$

where $c = \mathcal{R}_0/N$. Because $f''(A) < 0$, A^* is the global maximiser of f on $[0, N)$.

Relation to A_0. Introduce the normalized susceptible fraction $s := S/N$, and assume that $S_0 = N$. It follows from (B.7) that the classical final-size equation for $\alpha = 0$ can be written in normalized form as

$$\ln s_\infty = -\mathcal{R}_0(1-s_\infty), \qquad \text{with} \qquad s_\infty := \frac{S(\infty)}{N} = \frac{N-A_0}{N}.$$

Define the auxiliary function

$$\Phi(s) := \ln s + \mathcal{R}_0(1-s), \qquad s \in (0, 1].$$

Then s_∞ is a zero of Φ, i.e. $\Phi(s_\infty) = 0$.

Compute the derivative:

$$\Phi'(s) = \frac{1}{s} - \mathcal{R}_0.$$

Hence $\Phi'(s) > 0$ for $0 < s < 1/\mathcal{R}_0$, $\Phi'(1/\mathcal{R}_0) = 0$, and $\Phi'(s) < 0$ for $s > 1/\mathcal{R}_0$. Thus Φ attains its unique maximum at $s = 1/\mathcal{R}_0$.

Evaluate Φ at the endpoints and at the maximum:

$$\Phi(1) = \ln 1 + \mathcal{R}_0(1-1) = 0,$$

and by the standard inequality $\ln x \leq x - 1$ (with equality only at $x = 1$),

$$\Phi\big(1/\mathcal{R}_0\big) = \ln\big(1/\mathcal{R}_0\big) + \mathcal{R}_0\Big(1 - \frac{1}{\mathcal{R}_0}\Big) = -\ln \mathcal{R}_0 + \mathcal{R}_0 - 1 \geq 0,$$

with strict inequality when $\mathcal{R}_0 \neq 1$. Thus the maximum value $\Phi(1/\mathcal{R}_0)$ is nonnegative (positive for $\mathcal{R}_0 > 1$) while $\Phi(1) = 0$.

Because Φ is continuous on $(0, 1]$, has a unique maximum at $1/\mathcal{R}_0$ with $\Phi(1/\mathcal{R}_0) \geq 0$, and $\Phi(1) = 0$, the equation $\Phi(s) = 0$ has exactly two solutions in $(0, 1]$ when $\Phi(1/\mathcal{R}_0) > 0$: one at $s = 1$ and the other strictly to the left of $1/\mathcal{R}_0$. The epidemiologically relevant solution is the nontrivial root $s_\infty \in (0, 1)$. Because $\Phi(1/\mathcal{R}_0) > 0$ and $\Phi(1) = 0$, this nontrivial root must satisfy

$$s_\infty < \frac{1}{\mathcal{R}_0}. \tag{B.9}$$

Conclude $A^* < A_0$. Translate back to A. Since $s_\infty = (N - A_0)/N$, the inequality $s_\infty < 1/\mathcal{R}_0$ is equivalent to

$$\frac{N - A_0}{N} < \frac{1}{\mathcal{R}_0} \quad \Longrightarrow \quad N - A_0 < \frac{N}{\mathcal{R}_0} = \frac{1}{c},$$

hence

$$A_0 > N - \frac{1}{c} = A^*.$$

Therefore the unique maximiser A^* of $f(A)$, given by (B.8), indeed lies strictly to the left of the classical final size A_0 (for $\mathcal{R}_0 > 1$, i.e. in the nontrivial epidemic regime). Since A_α satisfies inequality (B.6), we conclude that A_α must lie to the left of A_0. Hence, $A_\alpha \leq A_0$.

□

References

[1] B. Luckingham, "To mask or not to mask: a note on the 1918 Spanish influenza epidemic in Tucson," *The Journal of Arizona History*, pp. 191–204, 1984.

[2] B. Dolan, "Unmasking history: Who was behind the anti-mask league protests during the 1918 influenza epidemic in San Francisco?," *Perspectives in Medical Humanities*, vol. 5, no. 19, 2020.

[3] J. Stover and Y. Teng, "The impact of condom use on the HIV epidemic," *Gates Open Research*, vol. 5, p. 91, 2022.

[4] D. Adjodah, K. Dinakar, M. Chinazzi, S. P. Fraiberger, A. Pentland, S. Bates, K. Staller, A. Vespignani, and D. L. Bhatt, "Association between COVID-19 outcomes and mask mandates, adherence, and attitudes," *PloS one*, vol. 16, no. 6, p. e0252315, 2021.

[5] S. Kwon, A. D. Joshi, C.-H. Lo, D. A. Drew, L. H. Nguyen, C.-G. Guo, W. Ma, R. S. Mehta, F. M. Shebl, E. T. Warner, *et al.*, "Association of social distancing and face mask use with risk of COVID-19," *Nature Communications*, vol. 12, no. 1, p. 3737, 2021.

[6] T. Urmi, B. Pant, G. Dewey, A. Quintana-Mathe, I. Lang, J. Druckman, K. Ognyanova, M. Baum, R. Perlis, C. Riedl, *et al.*, "Characterizing population-level changes in human behavior during the COVID-19 pandemic in the United States," *Proceedings of the National Academy of Sciences*, vol. 122, no. 37, p. e2500655122, 2025.

[7] N. M. Ferguson, D. Laydon, G. Nedjati-Gilani, N. Imai, K. Ainslie, M. Baguelin, S. Bhatia, A. Boonyasiri, Z. Cucunubá, G. Cuomo-Dannenburg, *et al.*, *Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand*, vol. 16. Imperial College London London, 2020.

[8] B. Pant, S. Safdar, M. Santillana, and A. B. Gumel, "Mathematical assessment of the role of human behavior changes on SARS-CoV-2 transmission dynamics in the United States," *Bulletin of Mathematical Biology*, vol. 86, no. 8, p. 92, 2024.

[9] S. Del Valle, H. Hethcote, J. M. Hyman, and C. Castillo-Chavez, "Effects of behavioral changes in a smallpox attack model," *Mathematical Biosciences*, vol. 195, no. 2, pp. 228–251, 2005.

[10] N. Perra, D. Balcan, B. Gonçalves, and A. Vespignani, "Towards a characterization of behavior-disease models," *PLOS One*, vol. 6, no. 8, p. e23084, 2011.

[11] M. S. Mahmud, S. Eshun, B. Espinoza, and C. Kadelka, "Adaptive human behavior and delays in information availability autonomously modulate epidemic waves," *PNAS nexus*, vol. 4, no. 5, p. pgaf145, 2025.

[12] A. Oveson, M. Girvan, and A. B. Gumel, "Modeling the impact of hospitalization-induced behavioral changes on the spread of COVID-19 in New York City," *Infectious Disease Modelling*, 2025.

[13] J. S. Weitz, S. W. Park, C. Eksin, and J. Dushoff, "Awareness-driven behavior changes can shift the shape of epidemics away from peaks and toward plateaus, shoulders, and oscillations," *Proceedings of the National Academy of Sciences*, vol. 117, no. 51, pp. 32764–32771, 2020.

[14] N. Gozzi, N. Perra, and A. Vespignani, "Comparative evaluation of behavioral epidemic models using COVID-19 data," *Proceedings of the National Academy of Sciences*, vol. 122, no. 24, p. e2421993122, 2025.

[15] L. LeJeune, N. Ghaffarzadegan, L. M. Childs, and O. Saucedo, "Formulating human risk response in epidemic models: exogenous vs endogenous approaches," *European Journal of Operational Research*, 2025.

[16] D. He, J. Dushoff, T. Day, J. Ma, and D. J. Earn, "Inferring the causes of the three waves of the 1918 influenza pandemic in England and Wales," *Proceedings of the Royal Society B: Biological Sciences*, vol. 280, no. 1766, p. 20131345, 2013.

[17] L. LeJeune, N. Ghaffarzadegan, L. M. Childs, and O. Saucedo, "Mathematical analysis of simple behavioral epidemic models," *Mathematical Biosciences*, vol. 375, p. 109250, 2024.

[18] C. T. Bauch and D. J. Earn, "Vaccination and the theory of games," *Proceedings of the National Academy of Sciences*, vol. 101, no. 36, pp. 13391–13394, 2004.

[19] S. Funk, E. Gilad, C. Watkins, and V. A. Jansen, "The spread of awareness and its impact on epidemic outbreaks," *Proceedings of the National Academy of Sciences*, vol. 106, no. 16, pp. 6872–6877, 2009.

[20] C. Eksin, K. Paarporn, and J. S. Weitz, "Systematic biases in disease forecasting–the role of behavior change," *Epidemics*, vol. 27, pp. 96–105, 2019.

[21] B. Pant and A. B. Gumel, "Mathematical assessment of the roles of age heterogeneity and vaccination on the dynamics and control of SARS-CoV-2," *Infectious Disease Modelling*, vol. 9, no. 3, pp. 828–874, 2024.

[22] E. Dong, H. Du, and L. Gardner, "An interactive web-based dashboard to track COVID-19 in real time," *The Lancet Infectious Diseases*, vol. 20, no. 5, pp. 533–534, 2020.

[23] Johns Hopkins University Center for Systems Science and Engineering, "COVID-19 Data Repository." `https://github.com/CSSEGISandData/COVID-19`, 2020. Accessed: 19 October 2025.

[24] T. Toni, D. Welch, N. Strelkowa, A. Ipsen, and M. P. Stumpf, "Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems," *Journal of The Royal Society Interface*, vol. 6, p. 187–202, July 2008.

[25] Y. Schälte, E. Klinger, E. Alamoudi, and J. Hasenauer, "pyabc: Efficient and robust easy-to-use approximate Bayesian computation," *Journal of Open Source Software*, vol. 7, no. 74, p. 4304, 2022.

[26] B. Pant, M. Lalovic, I. Z. Kiss, and M. Santillana, "epi-behavior-models: Repository." `https://github.com/markolalovic/epi-behavior-models`, 2025. Deposited 24 November 2025.

[27] R. E. Kass and A. E. Raftery, "Bayes factors," *Journal of the american statistical association*, vol. 90, no. 430, pp. 773–795, 1995.

[28] S. K. Gadarian, S. W. Goodman, and T. B. Pepinsky, "Partisanship, health behavior, and policy attitudes in the early stages of the COVID-19 pandemic," *PLOS One*, vol. 16, no. 4, p. e0249596, 2021.

[29] J. Kaashoek, C. Testa, J. T. Chen, L. M. Stolerman, N. Krieger, W. P. Hanage, and M. Santillana, "The evolving roles of US political partisanship and social vulnerability in the COVID-19 pandemic from February 2020–February 2021," *PLOS global public health*, vol. 2, no. 12, p. e0000557, 2022.

[30] R. H. Perlis, K. L. Trujillo, A. Safarpour, A. Quintana, M. D. Simonson, J. Perlis, M. Santillana, K. Ognyanova, M. A. Baum, J. N. Druckman, *et al.*, "Community mobility and depressive symptoms during the COVID-19 pandemic in the United States," *JAMA network open*, vol. 6, no. 9, pp. e2334945–e2334945, 2023.

[31] D. Stadlbauer, J. Tan, K. Jiang, M. M. Hernandez, S. Fabre, F. Amanat, C. Teo, G. A. Arunkumar, M. McMahon, C. Capuano, *et al.*, "Repeated cross-sectional sero-monitoring of sars-cov-2 in new york city," *Nature*, vol. 590, no. 7844, pp. 146–150, 2021.

[32] J. T. Davis, M. Chinazzi, N. Perra, K. Mu, A. Pastore y Piontti, M. Ajelli, N. E. Dean, C. Gioannini, M. Litvinova, S. Merler, *et al.*, "Cryptic transmission of SARS-CoV-2 and the first COVID-19 wave," *Nature*, vol. 600, no. 7887, pp. 127–132, 2021.

[33] K. A. Alroy, A. Crossa, C. Dominianni, J. Sell, K. Bartley, M. Sanderson, S. Fernandez, A. Levanon Seligson, S. Lim, S. M. Wang, *et al.*, "Population-based estimates of coronavirus disease 2019 (COVID-19)–like illness, COVID-19 illness, and rates of case ascertainment, hospitalizations, and deaths—noninstitutionalized New York City residents, March–April 2020," *Clinical Infectious Diseases*, vol. 73, no. 9, pp. 1707–1710, 2021.

[34] T. Kariya, "Rapid spread of COVID-19 in New York and the response of the community," *Global health & medicine*, vol. 2, no. 2, pp. 123–126, 2020.

[35] A. Rădulescu, S. Ballard, K. Gonzalez, and J. Linton, "Dynamic coupling between the covid-19 epidemic timeline and the behavioral response to pause in new york state counties," *Plos one*, vol. 16, no. 8, p. e0255236, 2021.

[36] V. Colizza and A. Vespignani, "Epidemic modeling in metapopulation systems with heterogeneous coupling pattern: Theory and simulations," *Journal of theoretical biology*, vol. 251, no. 3, pp. 450–467, 2008.

[37] D. Balcan, V. Colizza, B. Gonçalves, H. Hu, J. J. Ramasco, and A. Vespignani, "Multiscale mobility networks and the spatial spreading of infectious diseases," *Proceedings of the national academy of sciences*, vol. 106, no. 51, pp. 21484–21489, 2009.

[38] Z. Wang, M. A. Andrews, Z.-X. Wu, L. Wang, and C. T. Bauch, "Coupled disease–behavior dynamics on complex networks: A review," *Physics of life reviews*, vol. 15, pp. 1–29, 2015.