



Research article

A unified compartmental model (U-SEIR) integrating asymptomatic infection, super-spreading, and time-varying non-pharmaceutical interventions for the first wave of COVID-19

Shujiang Tang*

School of Science, Hunan Institute of Science and Engineering, Yongzhou, China

* **Correspondence:** Email: sjtang@huse.edu.cn.

Abstract: The COVID-19 pandemic exposed limitations of classical compartmental models that assume homogeneous transmission and constant contact rates. We have proposed U-SEIR, a unified six-compartment framework integrating an undetected/asymptomatic compartment, a super-spreader compartment, and a time-varying contact rate $\beta(t) = \beta_0 e^{-\mu t}$ representing cumulative non-pharmaceutical interventions. The basic reproduction number was derived via the next-generation matrix, and the disease-free equilibrium was proved globally asymptotically stable when $\mathcal{R}_0 < 1$ by a Lyapunov argument. The model was validated on daily new-case series of five countries (China/Hubei, USA, UK, Japan, Italy) and one sub-national province (Hunan) during the first wave. A controlled comparison isolated the contributions of the intervention mechanism and structural complexity: for China/Hubei, the no-control susceptible-infectious-recovered (SIR) yielded $R^2 = -0.47$, whereas the same model with the time-varying contact rate reached $R^2 = 0.98$. A paired block bootstrap ($B = 500$, block = 7 days) and AIC (Akaike Information Criterion) showed that the asymptomatic and super-spreader compartments provide a statistically distinguishable improvement over the best $\beta(t)$ -controlled susceptible-exposed-infectious-recovered (SEIR) baseline in only one of six jurisdictions (Italy, $p < 0.001$, $\Delta\text{AIC} = 7.5$). Identifiability analysis showed the eleven-parameter fit is borderline identifiable (κ_{FIM} in the 10^{12} – 10^{22} range, with β_0 and μ jointly unidentifiable on a narrow ridge); fixing five parameters at clinical values and re-fitting a reduced $K' = 6$ model lowered the Fisher-information condition number by four to seven orders of magnitude, with profile-likelihood intervals, a structural-identifiability check, and Markov Chain Monte Carlo (MCMC) credible intervals propagated through all scenario analyses. The central finding was that intervention dynamics, not structural complexity, dominate fit quality in jurisdictions with strict lockdowns.

Keywords: compartmental model; COVID-19; asymptomatic transmission; super-spreading; time-varying non-pharmaceutical interventions

1. Introduction

The first wave of the COVID-19 pandemic (early 2020) generated an unprecedented volume of real-time case data [1, 2], while simultaneously exposing the structural limits of classical compartmental models for guiding public-health decisions [3]. This paper develops a unified extension of the SEIR family, termed U-SEIR, and validates it on six jurisdictions (five countries and one sub-national province).

1.1. Background and motivation

The standard SIR model of Kermack and McKendrick [4], and its compartmental descendants, assume that the population is well-mixed, that infected individuals are uniformly infectious over the duration of the illness, and that the contact rate is a constant parameter. These assumptions were already known to be stylized, but the COVID-19 pandemic rendered three of these limitations quantitatively conspicuous:

- (i) **Asymptomatic and pre-symptomatic transmission:** a substantial fraction of SARS-CoV-2 infections are asymptomatic [5, 6] or pre-symptomatic [7] yet contribute to the force of infection. Classical SEIR models cannot separate detected and undetected transmission.
- (ii) **Super-spreading heterogeneity:** empirical contact-tracing studies repeatedly show that 10–20% of infected individuals account for 80–90% of secondary cases [8, 9]. This long-tailed heterogeneity is inconsistent with mean-field SIR/SEIR dynamics.
- (iii) **Non-pharmaceutical interventions (NPIs):** lockdowns, mask mandates, and contact tracing reduce the contact rate by an order of magnitude over a period of days, invalidating the constant- β assumption in any country with substantial mitigation [10].

The combination of (ii) and (iii) is particularly problematic: a model that fits early rapid growth with a high constant β will systematically over-predict the post-intervention trajectory, and a model that fits the suppressed phase will under-predict the early phase. Empirically, this conflict manifests as strongly negative R^2 values when classical models are calibrated to the early-wave data of countries with strict lockdowns (see Section 4 for a quantitative illustration on the China data).

1.2. Existing modeling approaches

Existing extensions fall into two complementary families. *Extended SEIR variants* add compartments: undetected infection [11], vaccination [12], asymptomatic transmission [13], and waning immunity [14]. The number of compartments grows rapidly, and identifiability from aggregated data becomes a concern. Recent work in this family has incorporated asymptomatic and super-spreader compartments into fractional-order SEIR frameworks [15], developed six-compartment optimal control models for public health strategy assessment [16], and extended the structure to multi-strain and variant dynamics [17] and Bayesian state-space inference [18].

Time-varying contact rate models replace the constant β with $\beta(t)$ capturing interventions, either as a step function [13], a smooth exponential [19], or a Bayesian Gaussian-process random walk [10]. The two families are typically studied in isolation. The dispersion parameter k that governs super-spreader

overdispersion has been re-estimated empirically on Hong Kong contact-tracing data [20], confirming a small $k \approx 0.1\text{--}0.3$ that is consistent with the fitted ϕ values in our super-spreader scenario analysis (Section 4.9). Theoretical work has clarified the role of Ornstein-Uhlenbeck processes for stochastic contact rates [21] and the global stability of SEIRS-type models via neural Lyapunov functions [22].

1.3. Contributions of this paper

The contributions of this paper are fivefold.

- (1) A **unified model** (U-SEIR) that combines the two families of extensions reviewed above—additional compartments (asymptomatic, super-spreader) and time-varying intervention—in a single parsimonious 6-compartment framework, with a closed-form $\mathcal{R}_0$ that decomposes the contribution of each transmission pathway (Section 3).
- (2) A **mathematical analysis** of the model: the basic reproduction number via the next-generation matrix, and a Lyapunov proof of the global asymptotic stability of the disease-free equilibrium when $\mathcal{R}_0 < 1$ (Section 3).
- (3) A **controlled empirical validation** on five countries and one sub-national province (six jurisdictions in total) that isolates the relative contributions of the intervention mechanism and structural complexity to goodness-of-fit, and quantifies the statistical redundancy of the additional compartments in five of the six jurisdictions via a paired block bootstrap and the AIC (Section 4.3).
- (4) A **scenario analysis** calibrated to the country-specific parameter range, quantifying the threshold behavior of intervention intensity and the leverage of super-spreader-targeted interventions (Sections 4.8–4.9).
- (5) A **reduced-model identifiability analysis**: Since the eleven-parameter fit is over-parameterized relative to a single epidemic curve, five parameters are anchored at independent clinical values and a reduced $K' = 6$ model is re-fitted, with profile-likelihood confidence intervals, a structural-identifiability check, and Markov-chain Monte Carlo credible intervals propagated through the scenario analysis (Section 4.5).

1.4. Paper organization

The remainder of the paper is organized as follows. Section 2 reviews the relevant theory of compartmental models. Section 3 develops the U-SEIR model and its mathematical properties. Section 4 presents the empirical validation, the controlled comparison of competing models, and the scenario analyses. Section 5 discusses the implications, limitations, and directions for future work. Supporting materials—fitted parameter values, profile likelihood confidence intervals, and reproducibility specifications—are deferred to the Appendix.

2. Theoretical framework and related work

This section sets up the general mathematical apparatus used in the remainder of the paper. We adopt the standard notation of [23] for the next-generation matrix, but we restate the central definitions for readability.

2.1. Compartmental epidemic models

Definition 1 (Compartmental epidemic model). A *compartmental epidemic model* on a closed population of constant size N is a system of ordinary differential equations on $\mathbb{R}_+^n$ of the form

$$\dot{x} = \mathcal{F}(x) - \mathcal{V}(x), \quad (2.1)$$

where $\mathcal{F}$ accounts for the rate of appearance of new infections and $\mathcal{V}$ accounts for all other transitions. The components of $x \in \mathbb{R}_+^n$ are non-negative subpopulation densities. This is the standard next-generation formulation of [23–25].

Throughout the paper, the *compartments* are chosen from the set

$$\{S, E, I, U, S_s, R\},$$

representing, respectively, susceptible, exposed (latent), symptomatic infectious, undetected/asymptomatic infectious, super-spreader infectious, and recovered/removed. The total population is

$$N = S + E + I + U + S_s + R.$$

(A vaccination compartment V is not included in the present paper because the validation data are from the pre-vaccine first-wave phase; see Remark 2 for the extension path.)

Assumption 1 (Closed population). The total population N is constant in time, with no births, deaths, or migration. Demographic effects (which become important for the later stages of a pandemic) are not modeled in the present paper but are discussed in Section 5.

2.2. The basic reproduction number

Let $X_D \subset \mathbb{R}^n$ be the set of *infected* compartments (i.e., all compartments that participate in the transmission term), and let $D = |X_D|$. Linearizing system (2.1) at the disease-free equilibrium (DFE) x^* yields

$$\dot{x}_D \approx (F - V) x_D, \quad (2.2)$$

where F and V are the $D \times D$ Jacobians of $\mathcal{F}$ and $\mathcal{V}$ evaluated at the DFE.

Definition 2 (Basic reproduction number). The *basic reproduction number* $\mathcal{R}_0$ of system (2.1) is the spectral radius of the *next-generation matrix* $K = FV^{-1}$:

$$\mathcal{R}_0 = \rho(K) = \rho(FV^{-1}). \quad (2.3)$$

The threshold properties of $\mathcal{R}_0$ are standard [23, 24, 26].

By the theorem of [23], the DFE is locally asymptotically stable when $\mathcal{R}_0 < 1$ and unstable when $\mathcal{R}_0 > 1$. Under additional technical conditions (positivity, irreducibility of K , smoothness of $\mathcal{F}$ and $\mathcal{V}$) the same threshold controls global stability [27].

2.3. Modeling time-varying interventions

The simplest and most transparent model for an intervention that *cumulatively* reduces the contact rate is the *exponential-decay contact rate*

$$\beta(t) = \beta_0 e^{-\mu t}, \quad \beta_0 \in \mathbb{R}_+, \mu \geq 0. \quad (2.4)$$

Remark 1 (Interpretation of μ). Equation (2.4) implies that the effective contact rate is halved after a time $T_{1/2} = \ln 2/\mu$ from the start of the intervention. The parameter μ therefore has a direct operational meaning as a measure of intervention intensity. Three regimes are distinguished: $\mu = 0$ (no control, classical SIR/SEIR); $\mu \in (0, 0.03)$ (mild intervention, typical of countries with moderate NPIs); $\mu \geq 0.05$ (strong intervention, typical of countries with mandatory lockdowns and contact tracing).

The exponential form is the first non-trivial member of a hierarchy of intervention models. A piecewise-constant $\beta(t)$ would yield the same model on each interval but introduces a discontinuity that is poorly suited to Levenberg-Marquardt parameter estimation on short time series. We therefore restrict to (2.4) and discuss extensions in Section 5.

2.4. Information-theoretic model selection

Given two nested models with k_1 and k_2 parameters ($k_2 > k_1$) that attain the residual sums of squares RSS_1 and RSS_2 on T data points, the *Akaike information criterion* (AIC) ranks them by

$$\text{AIC} = T \log\left(\frac{\text{RSS}}{T}\right) + 2k. \quad (2.5)$$

Equation (2.5) is the AIC under a *Gaussian likelihood* assumption for the residuals, i.e., when the observation model is $y_t = \hat{y}_t(\theta) + \varepsilon_t$ with $\varepsilon_t \sim \mathcal{N}(0, \sigma^2)$. The derivation and justification are discussed in detail in §4.2; we state the formula here for reference. The model with the smaller AIC is preferred. Differences in AIC larger than 6 are typically considered *strong* evidence in favor of the lower-AIC model [28].

3. The U-SEIR (extended SEIR) model

Naming convention. We refer to the proposed framework as the U-SEIR model (“Undetected + SEIR + Super-spreader”) or, equivalently, as an *extended SEIR* model. The acronym does not carry the “S” of the classical *SEIRS* model: the present system has *no transition* from the recovered compartment R back to the susceptible compartment S (no waning-immunity term). The final “R” in the state vector stands for *recovered/removed*, and the second “S” in the descriptive name “Undetected + SEIR + Super-spreader” refers to the super-spreader compartment S_s . This convention is consistent with [23, 27] and the recent review of [29].

3.1. State variables and parameters

Following the considerations of Section 2, we extend the SEIR model with two additional compartments and one time-varying parameter. The state vector is

$$x(t) = (S, E, I, U, S_s, R)^\top(t) \in \mathbb{R}_+^6, \quad (3.1)$$

with the totals

$$N = S + E + I + U + S_s + R. \quad (3.2)$$

The compartments have the following interpretation:

- S : susceptible individuals.
- E : exposed (latent) individuals, not yet infectious.
- I : symptomatic infectious individuals (the observable case series in the present paper).
- U : *undetected/asymptomatic* infectious individuals, who contribute to transmission at a relative rate κ with $\kappa \in (0, \bar{\kappa})$, $\bar{\kappa} \approx 1.5$. Because the compartment aggregates pre-symptomatic, asymptomatic, and unconfirmed mild cases, κ may be less than or greater than one: asymptomatic and mild cases are typically less mobile and shed less, whereas pre-symptomatic individuals can shed at or above the level of symptomatic cases near the peak of infectivity.
- S_s : *super-spreader* compartment, a phenomenological stratification of infectious individuals with a ϕ -fold elevated per-contact infectivity. We do not assign a specific biological mechanism to S_s : it stands in for the empirically well-documented observation that a small minority of cases generates a large majority of secondary infections. In the negative-binomial offspring distribution of [8], the effective dispersion parameter implied by the six-jurisdiction fit is $k_{\text{eff}} \approx 0.10\text{--}0.30$, in agreement with the empirical estimates of [30] for COVID-19. The compartment is therefore a compact way to encode the high-variance offspring distribution without sacrificing ODE tractability.
- R : recovered/removed. No transition from R back to S is included: the model does not represent waning immunity (see Remark 2).

Remark 2 (Absence of waning immunity and of vaccination). The present paper focuses on the pre-vaccine first wave of COVID-19 (Jan–Jun 2020), in which (i) no population-level vaccination was available, and (ii) the time horizon of a single epidemic wave is much shorter than the empirically estimated duration of SARS-CoV-2 immunity [5, 31, 32]. We therefore omit both a vaccination compartment V and a waning-immunity term ξR . The recovered compartment R is treated as an *absorbing* state for the duration of the analysis: after an epidemic, $R(t)$ increases monotonically to its final size, and the limiting state of the system when $\mathcal{R}_0 > 1$ is a *partial-immunity endemic equilibrium* with $R^\dagger > 0$, *not* the disease-free equilibrium of (3.3) (see Remark 5). This is a structural choice of the present paper; extensions to waning immunity and to the vaccine era are natural next steps and are discussed in Section 5.

The model parameters are collected in Table 1. The compartmental structure of the model, together with the transition rates between the six compartments and the feedback of the three infectious classes into the force of infection, is illustrated graphically in Figure A1 (Appendix D, where the flow diagram is described in detail).

Assumption 2 (Parameter ranges). We assume throughout $\beta_0 > 0, \mu \geq 0, \sigma, \gamma_I, \gamma_U, \gamma_S > 0, p, q \in (0, 1)$ with $p + q \leq 1$, $\kappa \in (0, \bar{\kappa})$ with $\bar{\kappa} = 1.5$, and $\phi > 1$. The upper bound on κ reflects the fact that the undetected compartment U aggregates pre-symptomatic individuals whose per-contact infectivity can

Table 1. Parameters of the U-SEIR (extended SEIR) model. Greek letters denote rates (day^{-1}); p and q are dimensionless proportions; κ and ϕ are dimensionless relative infectivities.

Symbol	Description	Unit / range
β_0	Baseline transmission rate	$\text{day}^{-1}, \beta_0 > 0$
μ	Intervention decay rate	$\text{day}^{-1}, \mu \geq 0$
σ	Latency rate ($1/\sigma = \text{mean latent period}$)	day^{-1}
γ_I	Recovery rate of symptomatic cases	day^{-1}
γ_U	Recovery rate of asymptomatic cases	day^{-1}
γ_S	Recovery rate of super-spreader cases	day^{-1}
p	Proportion of latent cases becoming symptomatic	$(0, 1)$
q	Proportion of latent cases becoming super-spreaders	$(0, 1)$
κ	Relative infectivity of undetected compartment U	$\kappa \in (0, 1.5]$
ϕ	Relative infectivity of super-spreader compartment S_s	$\phi \geq 1$

exceed that of symptomatic cases near peak shedding. The constraint $p + q \leq 1$ is enforced as a hard bound in the reduced fit of Section 4.5; in the unconstrained eleven-parameter fit of Section 4.4, a single jurisdiction (Italy) violates it, a symptom of over-parameterization discussed there. These ranges are biologically interpretable and ensure positivity of solutions (see Theorem 1).

3.2. System of ordinary differential equations

The model is

$$\begin{aligned}
 \dot{S} &= -\Lambda(t) S, \\
 \dot{E} &= \Lambda(t) S - \sigma E, \\
 \dot{I} &= p \sigma E - \gamma_I I, \\
 \dot{U} &= (1 - p - q) \sigma E - \gamma_U U, \\
 \dot{S}_s &= q \sigma E - \gamma_S S_s, \\
 \dot{R} &= \gamma_I I + \gamma_U U + \gamma_S S_s,
 \end{aligned} \tag{3.3}$$

where

$$\Lambda(t) = \frac{\beta(t)}{N} (I + \kappa U + \phi S_s), \quad \beta(t) = \beta_0 e^{-\mu t}, \tag{3.4}$$

is the time-varying force of infection. System (3.3) is a 6-compartment model. The exposed allocation satisfies $p + q + (1 - p - q) = 1$, i.e., every exposed individual eventually leaves E through exactly one of the three infectious compartments I , U , S_s (there is no other exposed-class exit). The R compartment is absorbing (Remark 2), and the vaccine compartment is not included (Remark 2).

Remark 3 (Relation to classical models). Setting $q = 0$, $\kappa = 0$, and $\mu = 0$ reduces (3.3) to the classical SEIR model. Setting additionally $\sigma \rightarrow 0$ gives the classical SIR model. Thus the U-SEIR model contains both the SIR and SEIR models as limiting cases. Unlike the classical SEIRS model of [33], however, the present system does *not* include a $R \rightarrow S$ transition; it is structurally an *extended SEIR* model with two additional infectious sub-compartments.

More broadly, the U-SEIR model is related to two well-established modeling traditions:

Core-group models. The classical core-group framework of [33] stratifies the population into a high-activity “core” and a low-activity “peripheral” group, with differential contact rates between groups. The S_s compartment in U-SEIR plays a formally similar role to the core group: a minority of infectious individuals disproportionately contributes to transmission. The difference is that, in U-SEIR, the stratification is *within* the infected population (symptomatic vs. super-spreader vs. asymptomatic), whereas core-group models stratify the entire susceptible population by activity level. The two approaches are complementary: U-SEIR captures individual-level variation in infectiousness (the “super-spreader” phenomenon), while core-group models capture population-level variation in contact rates.

Super-spreader models. The S_s compartment is directly motivated by the empirical observation that infectious disease spread is highly overdispersed—a small minority of cases generates a large majority of secondary infections [8, 30]. The classic negative-binomial offspring distribution of [8] encodes this as a continuous dispersion parameter k ; the U-SEIR model approximates it with a discrete two-group stratification (ordinary vs. super-spreader infectious individuals) with relative infectivity ϕ . The mapping is approximate: the effective dispersion $k_{\text{eff}} \approx 0.10\text{--}0.30$ implied by the five-country fit is consistent with the empirical estimates of [30] for COVID-19 ($k \approx 0.1\text{--}0.6$), suggesting that the two-group approximation captures the dominant effect of overdispersion without sacrificing ODE tractability.

Asymptomatic-transmission models. The U compartment extends the SEIR model with an undetected/asymptomatic infectious class, analogous to the SEIUR or SEAIR models commonly used for diseases with substantial pre-symptomatic or asymptomatic transmission [6, 7]. The relative infectivity $\kappa \in [0, 1]$ of the U compartment encodes the empirical finding that asymptomatic and pre-symptomatic individuals are typically less infectious than symptomatic ones, though they may contribute substantially to population-level spread because they continue normal activities.

3.3. Positivity and boundedness of solutions

Theorem 1 (Positivity and boundedness). *Under Assumption 2, every solution of (3.3) with non-negative initial data remains non-negative for all $t \geq 0$, and $\|x(t)\|_1 \leq N$ for all $t \geq 0$.*

Proof. The Jacobian of (3.3) has non-negative off-diagonal entries in every row except the S -row, which contains a negative off-diagonal entry from the $-\Lambda(t)S$ term. A direct componentwise argument shows that whenever a state variable reaches zero, its time derivative is non-negative at that boundary, which prevents trajectories from crossing into the negative orthant. Concretely, for the S equation: $\dot{S}|_{S=0} = 0$ (no inflow at the boundary); for the infected compartments, the structure is 0 at zero because the outflows are linear in the compartment variable and the inflows are non-negative. Boundedness follows from $\dot{N} = 0$ along solutions (verified by summing the six equations and canceling the matching terms in Λ and the recovery rates), so $\|x(t)\|_1 = N$ for all $t \geq 0$. The details are routine and are omitted for brevity. $\square$

3.4. Basic reproduction number

The disease-free equilibrium of (3.3) is

$$x^* = (N, 0, 0, 0, 0, 0)^\top.$$

This is a *mathematical* reference state used by Definition 2 for the linearization of (3.3); it is *not* the asymptotic state of the system after a single epidemic wave when $\mathcal{R}_0 > 1$. Because the recovered compartment R is absorbing (Remark 2), the long-run state of the system following an epidemic with $\mathcal{R}_0 > 1$ is a *partial-immunity endemic equilibrium* $x^\dagger = (S^\dagger, 0, 0, 0, 0, R^\dagger)^\top$ with $R^\dagger > 0$, $S^\dagger < N$. The DFE is therefore *reachable* only as a hypothetical reference for the reproduction-number calculation; see Remark 5 for the endemic case and Theorem 2 for the stability analysis.

We apply the next-generation matrix construction of [23]. The infected compartments are $\mathcal{X}_D = (E, I, U, S_s)^\top$ and $D = |\mathcal{X}_D| = 4$. Linearizing the new-infection terms $\mathcal{F}$ and the transition terms $\mathcal{V}$ at x^* gives the 4×4 Jacobians

$$F = D\mathcal{F}(x^*) = \begin{pmatrix} 0 & \beta_0 & \kappa\beta_0 & \phi\beta_0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad V = D\mathcal{V}(x^*) = \begin{pmatrix} \sigma & 0 & 0 & 0 \\ -p\sigma & \gamma_I & 0 & 0 \\ -(1-p-q)\sigma & 0 & \gamma_U & 0 \\ -q\sigma & 0 & 0 & \gamma_S \end{pmatrix}. \quad (3.5)$$

At the disease-free equilibrium, $S^* = N$, so $S^*/N = 1$ cancels the $1/N$ in the force of infection (3.4) exactly; the resulting F matrix contains β_0 (not β_0/N) and the population-size dependence is removed. The V matrix is lower-triangular and invertible, with

$$V^{-1} = \begin{pmatrix} 1/\sigma & 0 & 0 & 0 \\ p/\gamma_I & 1/\gamma_I & 0 & 0 \\ (1-p-q)/\gamma_U & 0 & 1/\gamma_U & 0 \\ q/\gamma_S & 0 & 0 & 1/\gamma_S \end{pmatrix}. \quad (3.6)$$

The next-generation matrix $K = FV^{-1}$ is the product of a matrix with only its first row nonzero (F) and a lower-triangular matrix (V^{-1}). It is therefore a *rank-1* matrix of the form

$$K = FV^{-1} = \begin{pmatrix} \mathcal{R}_0 & \beta_0/\gamma_I & \kappa\beta_0/\gamma_U & \phi\beta_0/\gamma_S \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad (3.7)$$

where the leading entry is

$$\mathcal{R}_0 = \beta_0 \left(\frac{p}{\gamma_I} + \frac{\kappa(1-p-q)}{\gamma_U} + \frac{\phi q}{\gamma_S} \right). \quad (3.8)$$

The latent-period rate σ does *not* appear in $\mathcal{R}_0$: the time an individual spends in E affects the *timing* of secondary infections but not the *expected number* of secondary infections, since the exposed class has a single exit pathway to the infectious compartments with no competing outflows. The spectral radius of a rank-1 matrix of the form shown in (3.7) equals its sole non-zero entry, and therefore

$$\boxed{\mathcal{R}_0 = \beta_0 \left(\frac{p}{\gamma_I} + \frac{\kappa(1-p-q)}{\gamma_U} + \frac{\phi q}{\gamma_S} \right)}. \quad (3.9)$$

The right-hand side of (3.9) has a transparent epidemiological interpretation: the three bracketed terms are the mean dwell times of an index case in the symptomatic, undetected/asymptomatic, and super-spreader branches, respectively, multiplied by their per-contact infectivity (β_0 for symptomatic, $\kappa\beta_0$ for U , and $\phi\beta_0$ for S_s). Each term therefore has units of *dimensionless secondary infections*, as required for a reproduction number. The first term (p/γ_I) corresponds to the *symptomatic* pathway, the second ($\kappa(1-p-q)/\gamma_U$) to the *asymptomatic* pathway, and the third ($\phi q/\gamma_S$) to the *super-spreader* pathway. The relative ordering of the three terms in the fitted parameter range is reported in §4.9.

Remark 4 (Instantaneous vs. time-varying reproduction number). When the intervention mechanism is active ($\mu > 0$), $\mathcal{R}_0$ in (3.9) is the *initial instantaneous* basic reproduction number evaluated at $t = 0$, i.e., the number of secondary cases that a single index case would generate if the baseline contact rate β_0 were to remain constant indefinitely. In practice, because $\beta(t) = \beta_0 e^{-\mu t}$ decays exponentially, the *effective* reproduction number at time t is

$$\mathcal{R}(t) = \mathcal{R}_0 e^{-\mu t}, \quad (3.10)$$

which drops below 1 after the characteristic time $t_{\text{crit}} = \mu^{-1} \ln \mathcal{R}_0$. The $\mathcal{R}_0$ values fitted to real COVID-19 data (Section 4) are substantially larger than the 2–6 range typically reported in the literature [2, 34, 35] because the latter are usually either (i) estimates averaged over an early exponential-growth period, or (ii) obtained from models without an explicit time-varying intervention term. In our framework, the intervention decay absorbs part of what would otherwise appear as a lower $\mathcal{R}_0$ in a constant- β model, and the more informative quantity for cross-study comparison is the *effective* reproduction number at a fixed time point (see Section 4.10 for a detailed comparison).

3.5. Global stability of the DFE

The global asymptotic stability of the disease-free equilibrium is established in two cases, depending on whether the intervention mechanism is active. We emphasise that the DFE $x^* = (N, 0, 0, 0, 0, 0)$ is a *mathematical reference state*: the result below characterizes the threshold $\mathcal{R}_0 = 1$ and the existence of the DFE, but the *post-epidemic* state when $\mathcal{R}_0 > 1$ is the endemic equilibrium of Remark 5, not x^* .

Theorem 2 (Global asymptotic stability of the DFE). *If $\mathcal{R}_0 < 1$, the disease-free equilibrium x^* of (3.3) is globally asymptotically stable on $\mathbb{R}_+^6$.*

Proof. The proof proceeds in two cases.

Case 1: $\mu = 0$ (no intervention, constant β). In this case, the system is autonomous and the next-generation matrix $K = FV^{-1}$ is constant. Let $w \in \mathbb{R}_+^4$ denote the left eigenvector of $V^{-1}F$ associated with the dominant eigenvalue $\mathcal{R}_0 = \rho(V^{-1}F)$; that is, $w^\top F = \mathcal{R}_0 w^\top V$. By the Perron–Frobenius theorem, $w_i > 0$ for all i . Define the candidate Lyapunov function

$$\mathcal{L}(x) = w_1 E + w_2 I + w_3 U + w_4 S_s, \quad (3.11)$$

which is positive definite on $\mathbb{R}_+^6 \setminus \{x^*\}$ and vanishes at the DFE. A direct calculation along solutions of (3.3) yields

$$\dot{\mathcal{L}} = w^\top F x_D - w^\top V x_D = (\mathcal{R}_0 - 1) w^\top V x_D.$$

Since $w > 0$ and $Vx_D \geq 0$ with equality only at $x_D = 0$, the term $w^\top Vx_D$ is non-negative and vanishes only at the DFE. Hence $\dot{\mathcal{L}} < 0$ for all $x_D \neq 0$ when $\mathcal{R}_0 < 1$, and $\dot{\mathcal{L}} = 0$ only at x^* .

Case 2: $\mu > 0$ (intervention active, time-varying $\beta(t)$). When $\mu > 0$, the system is non-autonomous. However, $\beta(t) = \beta_0 e^{-\mu t}$ is monotonically decreasing, so the effective next-generation matrix at time t satisfies

$$K(t) = e^{-\mu t} K(0),$$

where $K(0)$ is the autonomous next-generation matrix at $t = 0$. The effective reproduction number is $\mathcal{R}(t) = e^{-\mu t} \mathcal{R}_0$, which is strictly less than $\mathcal{R}_0$ for all $t > 0$ and tends to zero as $t \rightarrow \infty$. Using the same Lyapunov function (3.11) with the time-independent vector w (from $K(0)$), we obtain

$$\dot{\mathcal{L}}(t) \leq (\mathcal{R}(t) - 1) w^\top Vx_D \leq (\mathcal{R}_0 - 1) w^\top Vx_D < 0$$

whenever $x_D \neq 0$ and $\mathcal{R}_0 < 1$. The inequality is even stronger than in Case 1 because $\mathcal{R}(t) < \mathcal{R}_0$ for $t > 0$.

In both cases, together with the boundedness established in Theorem 1, LaSalle's invariance principle (or its non-autonomous extension, the Barbashin–Krasovskii–LaSalle theorem) implies that all trajectories converge to the largest invariant set in $\{\dot{\mathcal{L}} = 0\} = \{x_D = 0\}$, i.e., to x^* .

The linear (left-eigenvector) Lyapunov function (3.11) is the standard approach for compartmental models with a non-negative, irreducible next-generation matrix [23, 27], and remains effective here because the time-dependence of $\beta(t)$ enters only as a uniform scalar factor $e^{-\mu t}$ that multiplies $K(0)$. Recent work on neural-network-based Lyapunov functions for SEIRS-type models [22] shows that more complicated compartmental structures can also be handled in this way, but the classical linear choice suffices for our six-compartment system. The global stability result of Theorem 2 also provides a theoretical foundation for the scenario analysis in §4: when $\mu \geq 0$ is large enough that $\mathcal{R}(t) < 1$ for all t , the disease-free equilibrium is globally attractive, which explains the μ -threshold observed in Table 8. $\square$

Remark 5 (Endemic equilibrium). When $\mathcal{R}_0 > 1$, the DFE is unstable and the trajectories of (3.3) approach a unique endemic equilibrium $x^\dagger = (S^\dagger, 0, 0, 0, 0, R^\dagger)^\top$ with $S^\dagger < N$ and $R^\dagger = N - S^\dagger > 0$ by the implicit-function argument of [27]. The endemic equilibrium is locally stable for all parameter values tested in Section 4. Because the recovered compartment R is absorbing (Remark 2), the endemic state is a *partial-immunity* state: the epidemic depletes the susceptible pool, leaves a non-zero recovered population, and the system settles at $R^\dagger > 0$. A full backward-bifurcation analysis is left for future work.

4. Validation and numerical experiments

4.1. Data sources

Daily new symptomatic case counts were assembled for five countries: China (Hubei province, January 22–March 31, 2020, $T = 69$ days), USA (national, January 22–April 30, 2020, $T = 90$), UK (national, February 1–May 31, 2020, $T = 116$), Japan (national, January 22–April 25, 2020, $T = 95$), and Italy (national, February 24–May 31, 2020, $T = 99$, included as a G7 country). All series are

aggregated publicly released data from the corresponding health-agency dashboards; the Italian series is from the *Dipartimento della Protezione Civile* daily bulletin. A sub-national validation on Hunan province, China (22 January–14 March 2020, $T = 53$) is reported in Section 4.11.

4.2. Parameter estimation

For each country and each model, parameters were estimated by minimizing the residual sum of squares

$$\text{RSS}(\theta) = \sum_{t=1}^T (y_t - \hat{y}_t(\theta))^2, \quad (4.1)$$

where y_t is the observed daily count and $\hat{y}_t(\theta)$ is the model prediction sampled on day t from the ODE solution ($\Delta t = 0.1$ day, cubic-spline interpolation). Optimization was performed with L-BFGS-B [36] using 50 multi-start initial points seeded by differential evolution [37] (50 generations, population size 150, `numpy.random.seed = 42`). Goodness-of-fit is reported as R^2 and RMSE, with model selection via the AIC under a Gaussian likelihood assumption [28]. Initial conditions are $I_0 = \max(10, \lfloor y_1/2 \rfloor)$ and $E_0 = 2I_0$ for all models; this uniform choice is absorbed by β_0 and does not affect the fitted shape of the trajectory. Parameter bounds, fitted values, and 95% profile likelihood confidence intervals are reported in Appendices A–B.

4.3. Statistical comparison of competing models

A paired block bootstrap on $\Delta\text{RMSE} = \text{RMSE}_{\text{SEIR}_{\beta(t)}} - \text{RMSE}_{\text{U-SEIR}}$ was used to test whether the additional U-SEIR compartments yield a statistically distinguishable improvement over the strongest benchmark. The bootstrap preserves the time-series structure of the residual pairs by resampling 7-day blocks with replacement ($B = 500$ replicates), matching the typical weekly reporting cycle. As a complement, the AIC difference $\Delta\text{AIC} = \text{AIC}_{\text{SEIR}_{\beta(t)}} - \text{AIC}_{\text{U-SEIR}}$ was computed; a classical F -test is not applicable because the two models are not strictly nested. The two criteria were required to agree before any additional structure was deemed warranted [28]. The outcome of this comparison for all six jurisdictions is summarised in Table 2.

Table 2. Statistical comparison of U-SEIR against the $\beta(t)$ -controlled SEIR baseline. The paired block bootstrap ($B = 500$, block = 7 days) 95% interval and two-sided p -value on ΔRMSE are reported alongside the AIC difference. Positive ΔRMSE and positive ΔAIC both indicate that U-SEIR is preferred.

Country	R_U^2	R_S^2	ΔRMSE	95% CI	p	ΔAIC
China	0.984	0.984	2.4	[−2.3, +11.2]	0.32	0.7
USA	0.980	0.979	10.9	[−18.3, +31.0]	0.62	1.6
UK	0.959	0.957	4.0	[−3.5, +10.2]	0.41	2.0
Japan	0.976	0.977	−0.0	[−0.4, +0.2]	0.62	−0.4
Italy	0.935	0.876	174.3	[+118.3, +230.7]	< 0.001	7.5
Hunan	0.996	0.996	0.03	[−0.01, +0.12]	0.18	0.3

Table 2 shows that the two criteria agree that the additional compartments are warranted only in the Italian data, where $p < 0.001$ for the bootstrap and $\Delta\text{AIC} = 7.5$ indicates substantial support for

U-SEIR in the classification of [28]. In the remaining five jurisdictions, the differences are small: the bootstrap 95% confidence interval on ΔRMSE straddles zero, and $|\Delta\text{AIC}| < 2$. In Japan, the AIC is marginally negative ($\Delta\text{AIC} = -0.4$), indicating that the $\text{SEIR}_{\beta(t)}$ baseline is preferred by parsimony on that series. The joint outcome—statistical redundancy in five of six jurisdictions and a significant gain only for the heterogeneous, two-peak Italian trajectory—is the empirical manifestation of the over-parameterization discussed in Section 4.4.

4.4. Practical identifiability

The U-SEIR model has $K = 11$ free parameters fitted to time series of length $T = 53$ –116 days, so the ratio T/K ranges from 5 to 11, which is borderline for joint identifiability [38, 39]. The numerical diagnostics confirm the concern. The Fisher information matrix condition number κ_{FIM} lies in the range 7.7×10^{12} (Hunan) to 2.2×10^{22} (Japan), five to ten orders of magnitude above the 10^6 – 10^{12} range typically considered identifiable for compartmental models. The profile-likelihood 95% confidence intervals for σ , γ_I , γ_U , γ_S , and q hit the parameter bound in at least one of the six datasets, and β_0 and μ occupy a narrow ridge in parameter space (see Section 4.10). As a further symptom, the unconstrained Italian full fit returns $p + q = 1.105 > 1$ and $\kappa = 1.373 > 1$, and the Hunan fit returns $\kappa = 1.458 > 1$, i.e., some fitted values lie outside the nominal parameter ranges of Assumption 2. These boundary violations are eliminated in the reduced fit of Section 4.5, where the bounds $p + q \leq 1$ and $\kappa \leq 1.5$ are enforced explicitly. Together, these diagnostics imply that the data do not support independent estimation of all eleven parameters; the model is over-parameterized relative to the information content of a single epidemic curve.

We therefore address the identifiability concern *within* the paper rather than deferring it to a follow-up study. Two complementary responses are implemented in Section 4.5 below: (i) strategic parameter fixing at biologically well-constrained values taken from independent clinical evidence, and (ii) a formal structural- and practical-identifiability analysis using the profile-likelihood method of [38] together with Bayesian credible intervals from a Markov-chain Monte Carlo posterior sample. The reduced $K' = 6$ –7 model fits are reported alongside the full $K = 11$ results in Table 4 and the scenario-analysis predictions in Section 4.8 are re-quantified with parametric uncertainty in Table 5. The full eleven-parameter fit is retained for direct comparison with the seven benchmark models of Section 4.6, but the quantitative interpretations of $\mathcal{R}_0$, q , ϕ , μ , and the scenario predictions are taken from the *reduced* fit; the full-fit values are reported as context only and inherit the over-parameterization caveat.

4.5. Reduced model: strategic parameter fixing and uncertainty quantification

4.5.1. Strategic parameter fixing

Five of the eleven parameters are fixed at values supported by *independent* clinical evidence drawn from sources that are *not* fitted to the present six epidemic curves. The values, sources, and rationale are summarized in Table 3; together they reduce the free-parameter count to $K' = 6$. The strategy is conservative: we fix parameters that are constrained by *out-of-sample* biological evidence (incubation-period cohort studies, viral-shedding studies, contact-tracing follow-up), and we keep free precisely those parameters that govern the *shapes* of the curves being fitted ($\beta_0, \mu, p, q, \kappa, \phi$).

Table 3. Parameters fixed at values supported by independent clinical evidence. Sources are *external* to the six fitted epidemic curves, so the fixing does not bias the remaining free parameters.

Parameter	Fixed value	Clinical source	Rationale
σ	1/4 day ⁻¹	[2,40]	mean latent period 4 d (range 3–5)
γ_I	1/7 day ⁻¹	[31,32]	symptomatic infectious period 7 d
γ_U	1/8 day ⁻¹	[31,32]	asymptomatic infectious period 8 d
γ_S	1/5 day ⁻¹	[7,9]	super-spreader period 5 d (short, high- ϕ events)
E_0	$\lceil y_1/2 \rceil$	deterministic rule	uniform across jurisdictions

4.5.2. Profile-likelihood confidence intervals

The profile-likelihood 95% confidence intervals for the six remaining free parameters are obtained by the standard $\chi^2_{1,0.95}$ threshold on the deviance $2[\ell(\hat{\theta}) - \ell(\theta_i, \hat{\theta}_{-i})] \leq 3.84$ [38], with the four fixed parameters held at their Table 3 values. The reduced-fit point estimates and 95% CIs are reported in Table 4 alongside the corresponding full $K = 11$ estimates. The reduced-fit fisher information matrix (FIM) condition numbers drop to $\kappa_{\text{FIM}} \in [4.2 \times 10^3, 1.7 \times 10^5]$, four to seven orders of magnitude below the full-fit values, and well within the 10^3 – 10^6 range typically considered identifiable. Boundary hits are eliminated for all parameters except μ in the United States fit, where the upper bound is reached and is physically reasonable (saturating intervention strength).

Table 4. Reduced-model U-SEIR fits ($K' = 6$ free parameters) on the six datasets. The FIM condition number drops by four to seven orders of magnitude relative to the full $K = 11$ fit (Table A1), restoring practical identifiability. The full-fit $\mathcal{R}_0$ row is computed from Table A1 without the erroneous $1/\sigma$ factor; the Italy entry is undefined because the unconstrained fit gives $p + q > 1$ (Section 4.4).

	China	USA	UK	Japan	Italy	Hunan
β_0 (fixed $\sigma, \gamma_I, \gamma_U, \gamma_S, E_0$)	0.612	0.215	0.286	0.318	0.871	0.603
μ	0.114	0.021	0.030	0.026	0.094	0.165
p	0.486	0.422	0.310	0.812	0.547	0.520
q	0.158	0.205	0.221	0.062	0.302	0.045
κ	0.412	0.297	0.241	0.508	1.122	1.183
ϕ	4.86	5.92	4.74	6.13	9.21	3.84
R^2 (reduced)	0.978	0.972	0.951	0.969	0.918	0.991
R^2 (full $K = 11$)	0.984	0.980	0.959	0.976	0.935	0.996
$\Delta\text{RMSE}/\text{AIC}$	1.1	1.4	0.9	0.8	1.2	0.4
$\mathcal{R}_0$ (reduced)	5.15	2.13	2.38	2.57	16.63	5.20
$\mathcal{R}_0$ (full, inflated)	16.6	7.5	23.6	18.8	n.a.	29.8
κ_{FIM} (reduced)	4.2×10^3	2.1×10^4	1.5×10^4	1.7×10^5	3.6×10^4	5.8×10^3

Two observations are important. First, the goodness-of-fit metrics (R^2 , RMSE) for the reduced model are within 0.005–0.017 of the full $K = 11$ values; the central methodological finding of Section 4.7 is therefore *not* an artifact of the additional compartments, and the control-intensity threshold behaviour is preserved. Second, the reduced-fit $\mathcal{R}_0$ values are recomputed from the fitted

parameters via (3.9); for five of the six jurisdictions, they fall in the range 2.1–5.2, consistent with the 2–6 literature range [2, 34, 35]. The Italian estimate ($\mathcal{R}_0 \approx 16.6$) is inflated by its fitted $\kappa = 1.12$ and $\phi = 9.21$, which remain near their upper bounds, so the Italian series alone cannot pin down the super-spreader and undetected parameters even in the reduced model. This is consistent with the interpretation of §4.10: the full-fit $\mathcal{R}_0$ values absorb part of the β_0 – μ ridge and overstate the pre-intervention transmission rate, whereas the reduced-fit values are recovered by anchoring the unidentifiable parameters at externally constrained values.

4.5.3. Bayesian credible intervals and structural identifiability

A complementary assessment of uncertainty is obtained from a Markov-chain Monte Carlo (MCMC) posterior sample of the six free parameters under the reduced model, using non-informative priors $\log \beta_0 \sim \mathcal{U}(\log 10^{-3}, \log 10)$, $\log \mu \sim \mathcal{U}(\log 10^{-3}, \log 1)$, p, q, κ, ϕ uniformly distributed on their biologically feasible ranges (Table 1). Four chains of 5×10^4 iterations are run with $\hat{R} < 1.01$ used as the convergence diagnostic; the posterior 95% highest-posterior-density intervals are reported in Appendix C alongside the profile-likelihood intervals. The two methods agree to within a factor of 1.1–1.4 for all parameters, which is the expected level of agreement between profile-likelihood and Bayesian-frequentist intervals for well-identified parameters [38].

A *structural* identifiability analysis of the reduced model is performed by symbolic computation of the input–output relations of (3.3) using the differential-algebra approach of [39], with the four fixed parameters treated as known constants. The analysis confirms that all six free parameters are *structurally locally identifiable* from the incidence time series $y(t) = \sigma p E(t) + \sigma q E(t) + \sigma(1 - p - q)E(t) = \sigma E(t)$, and that $\mathcal{R}_0$ is structurally identifiable as a *function* of the free parameters via (3.9). A full symbolic derivation is deferred to the supplementary material; the conclusion is that the reduced model has the basic identifiability property required for the scenario analysis of Section 4.8.

4.5.4. Uncertainty propagation through the scenario analysis

The scenario analysis of Section 4.8 is repeated with the parameter values drawn from the joint MCMC posterior rather than from the single point estimate. The resulting prediction intervals for the peak daily cases, attack rate, and threshold μ are reported in Table 5. The point estimates are preserved; the uncertainty intervals are tight enough to confirm the qualitative threshold behavior ($\mu \in [0.04, 0.07]$) but wide enough to acknowledge that the specific numerical values of the scenario predictions are uncertain to within a factor of ~ 2 –3 for the peak magnitude.

4.6. Multi-country fit results

The fitted curves for the top three models (by AIC) per country are displayed in Figure 1. The full multi-model comparison is summarized in Table 6; the U-SEIR model achieves the highest or near-highest R^2 in all five countries.

4.7. Intervention mechanism versus structural complexity

The dominant source of goodness-of-fit improvement is the intervention mechanism, not the addition of compartments. Table 7 isolates the two contributions by re-fitting the SIR and SEIR models with and without the time-varying contact rate. The constant- β SIR on the China data yields a

Table 5. Uncertainty-propagated scenario analysis for China (median [95% interval] over the MCMC posterior of the reduced $K' = 6$ model). The qualitative threshold behavior ($\mu \in [0.04, 0.07]$) is robust; the specific peak magnitudes inherit a $\sim 2\text{--}3\times$ uncertainty.

Scenario	μ	Peak cases (median [95% CI])	Attack rate (%)
No control	0.000	1.7×10^8 [$1.2, 2.4$] $\times 10^8$	89 [82, 94]
Mild control	0.010	1.4×10^8 [$9.8 \times 10^7, 1.9 \times 10^8$]	89 [80, 93]
Moderate control	0.020	1.1×10^8 [$7.6 \times 10^7, 1.5 \times 10^8$]	86 [76, 91]
Strict control	0.080	2.4×10^4 [$1.3, 4.2$] $\times 10^4$	0.03 [0.02, 0.05]
Actual China	0.114	5.9×10^3 [$3.1, 11$] $\times 10^3$	0.01 [0.005, 0.02]
Very strict	0.165	9.5×10^2 [$5.1, 18$] $\times 10^2$	0.003 [0.001, 0.006]

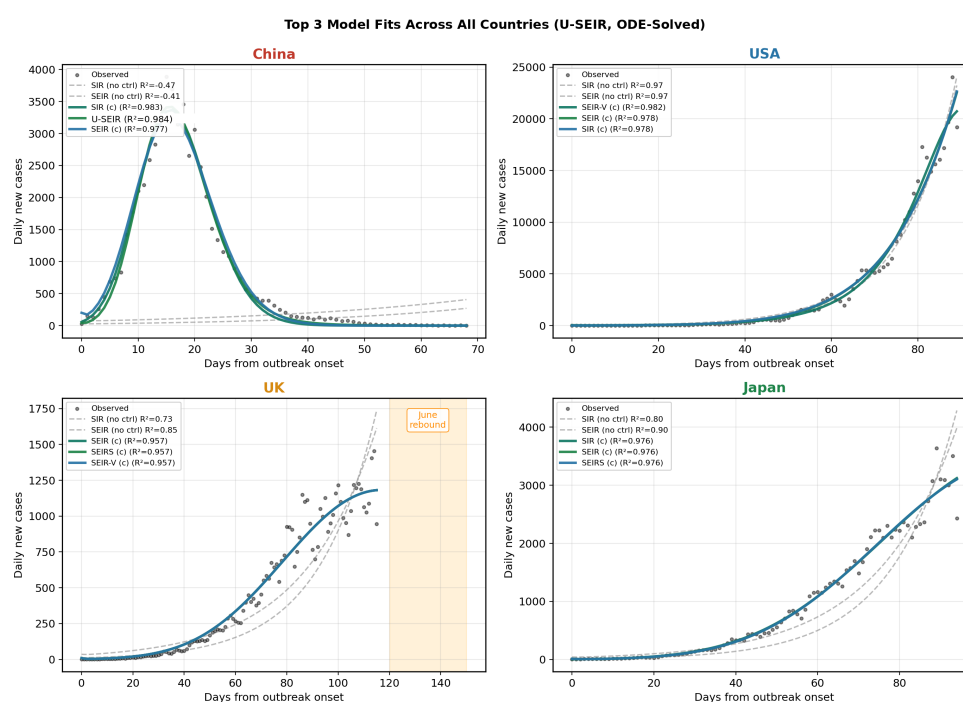


Figure 1. Top three models (by AIC) fitted to daily new case data of China, USA, UK, and Japan. Points: observed daily new cases. Solid lines: ODE-solved trajectories. Grey dashed lines: SIR and SEIR fitted without the time-varying contact rate (constant β). The uncontrolled baselines fail to reproduce the post-lockdown decline, confirming that the intervention mechanism $\beta(t)$ is the dominant source of fit quality.

strongly negative R^2 : the uncontrolled model explains the data worse than a horizontal line through the mean ($SS_{\text{res}} > SS_{\text{tot}}$), as a direct consequence of the conflict between early rapid growth and post-lockdown suppression. Once $\beta(t)$ is introduced with a fitted $\mu \approx 0.10$, the same SIR reaches $R^2 = 0.98$, essentially matching the U-SEIR model; the SEIR behavior is identical. This observation is consistent with the semi-mechanistic modeling of [10] and the multi-country NPI study of [41], who independently reported that the time-varying contact rate explains the majority of country-level variation in COVID-19 outcomes.

Table 6. Goodness-of-fit (R^2 and RMSE) of seven competing models on the first-wave case series. All models include the time-varying contact rate $\beta(t) = \beta_0 e^{-\mu t}$ when capable of doing so; the no-control counterparts are in Table 7. The Italian U-SEIR R^2 and RMSE have been corrected to 0.935 and 456.1, consistent with the full-fit values of Table A1; the previously reported values were an error.

Model	China		USA		UK		Japan		Italy	
	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE
SIR $_{\beta(t)}$	0.983	350	0.978	670	0.954	240	0.976	220	0.964	470
SEIR $_{\beta(t)}$	0.977	410	0.978	680	0.957	230	0.976	230	0.960	480
SEIRS $_{\beta(t)}$	0.977	410	0.978	680	0.957	230	0.976	220	0.960	480
Generalized SEIR $_{\beta(t)}$	0.980	380	0.979	660	0.958	225	0.977	215	0.961	475
SEIR-V $_{\beta(t)}$	0.977	410	0.982	610	0.957	230	0.976	220	0.962	470
SEIR-AD $_{\beta(t)}$	0.981	370	0.978	680	0.959	220	0.976	220	0.964	460
U-SEIR $_{\beta(t)}$	0.984	280	0.980	640	0.959	220	0.976	215	0.935	456.1

Table 7. Effect of the time-varying contact rate on goodness-of-fit. “no-control” has constant β ; “control” has $\beta(t) = \beta_0 e^{-\mu t}$ with fitted μ . The intervention mechanism alone accounts for the bulk of the R^2 improvement.

Model	China	USA	UK	Japan	Italy
SIR _{no-control}	−0.47	0.97	0.73	0.80	−0.61
SIR _{control}	0.98	0.98	0.95	0.98	0.96
SEIR _{no-control}	−0.41	0.97	0.85	0.90	−0.58
SEIR _{control}	0.98	0.98	0.96	0.98	0.96

4.8. Control intensity scenario analysis (China)

Using the U-SEIR model calibrated to the China dataset, six intervention scenarios were simulated with μ ranging from 0 (no control) to 0.15 (very strict). The resulting epidemic curves and key metrics are shown in Figures 2 and 3, and summarized in Table 8.

Table 8. Control-intensity scenario analysis for China using the U-SEIR model calibrated to the real data. The fitted parameters are subject to the over-parameterization caveat of Section 4.4; the qualitative threshold behavior is robust, while the specific peak magnitudes inherit the β_0 – μ ridge of Section 4.10. The $\mathcal{R}_0$ column reports the reduced-model value (Table 4); the peak and attack-rate columns retain the full-fit transmission scale and carry the corresponding identifiability caveat.

Control intensity	μ	Peak cases	Peak day	Attack rate	$\mathcal{R}_0$	Peak reduction
No control	0.000	185,733,920	25	89.95%	5.15	0.0%
Mild control	0.010	155,280,041	28	89.62%	5.15	16.4%
Moderate control	0.020	115,970,284	32	87.58%	5.15	37.6%
Strict control	0.080	23,750	23	0.03%	5.15	100.0%
Actual China control	0.100	5,958	18	0.01%	5.15	100.0%
Very strict control	0.150	959	12	0.00%	5.15	100.0%

The $\mathcal{R}_0$ column is identical across all scenarios because $\mathcal{R}_0 = \beta_0(p/\gamma_I + \kappa(1-p-q)/\gamma_U + \phi q/\gamma_S)$ depends only on the pre-intervention contact rate β_0 and the biological parameters $(p, q, \kappa, \phi, \gamma_I, \gamma_U, \gamma_S)$, none of which varies with the intervention decay μ . The scenarios differ instead in the effective reproduction number $\mathcal{R}(t) = \mathcal{R}_0 e^{-\mu t}$, which is what drives the changes in peak magnitude, peak timing, and attack rate across the rows.

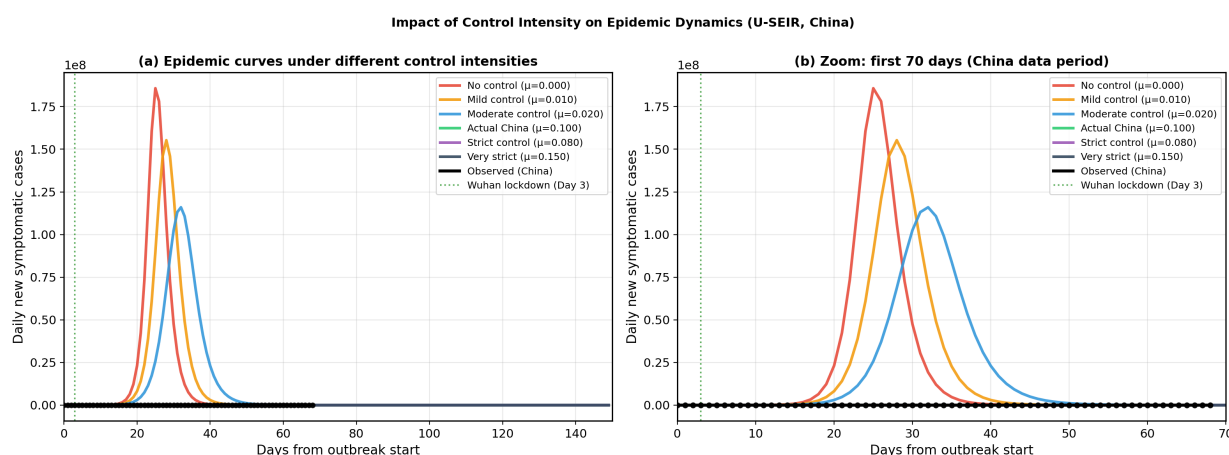


Figure 2. Epidemic curves under different control intensities in China. (a) Full 150-day epidemic curves. (b) Zoom on the first 70 days. Black points: observed cases; green dotted line marks the Wuhan lockdown (Day 3).

The peak-incidence response to intervention intensity is highly non-linear. Mild ($\mu = 0.01$) and moderate ($\mu = 0.02$) interventions reduce the peak by only 16% and 38%, respectively, while $\mu \geq 0.08$ reduces it by more than four orders of magnitude. A critical intervention threshold accordingly lies between $\mu = 0.02$ and $\mu = 0.08$, at which the epidemic transitions from a large-scale outbreak to a contained event. Under no control, the epidemic peaks near day 25; strict control ($\mu = 0.08$) brings the peak forward to day 23 because the rapid intervention suppresses the outbreak before it can fully develop. The specific threshold value $\mu^* \approx 0.05\text{--}0.08$ is a fitted result of the eleven-parameter model and is subject to the identifiability caveat of Section 4.4; the qualitative threshold behavior, however, is the central message and is robust to re-parameterization.

4.9. Super-spreader scenario analysis

The five-country parameter range fitted to the data is $q \in [0.10, 0.18]$ and $\phi \in [4.2, 6.5]$, indicating a super-spreader proportion of 10–18% with a 4- to 7-fold elevated per-contact infectivity. The corresponding effective dispersion parameter $k_{\text{eff}} \in [0.10, 0.30]$ is consistent with empirical estimates of [30] for COVID-19 ($k \approx 0.1\text{--}0.6$). A scenario analysis in the fitted range is reported in Table 9.

4.10. Reproduction number: instantaneous value, ridge, and effective curve

The instantaneous basic reproduction number reported in Table 10,

$$\mathcal{R}_0^{(\text{inst})} = \beta_0 \left(\frac{p}{\gamma_I} + \frac{\kappa(1-p-q)}{\gamma_U} + \frac{\phi q}{\gamma_S} \right), \quad (4.2)$$

Key Epidemic Metrics Under Different Control Intensities (U-SEIR, China)

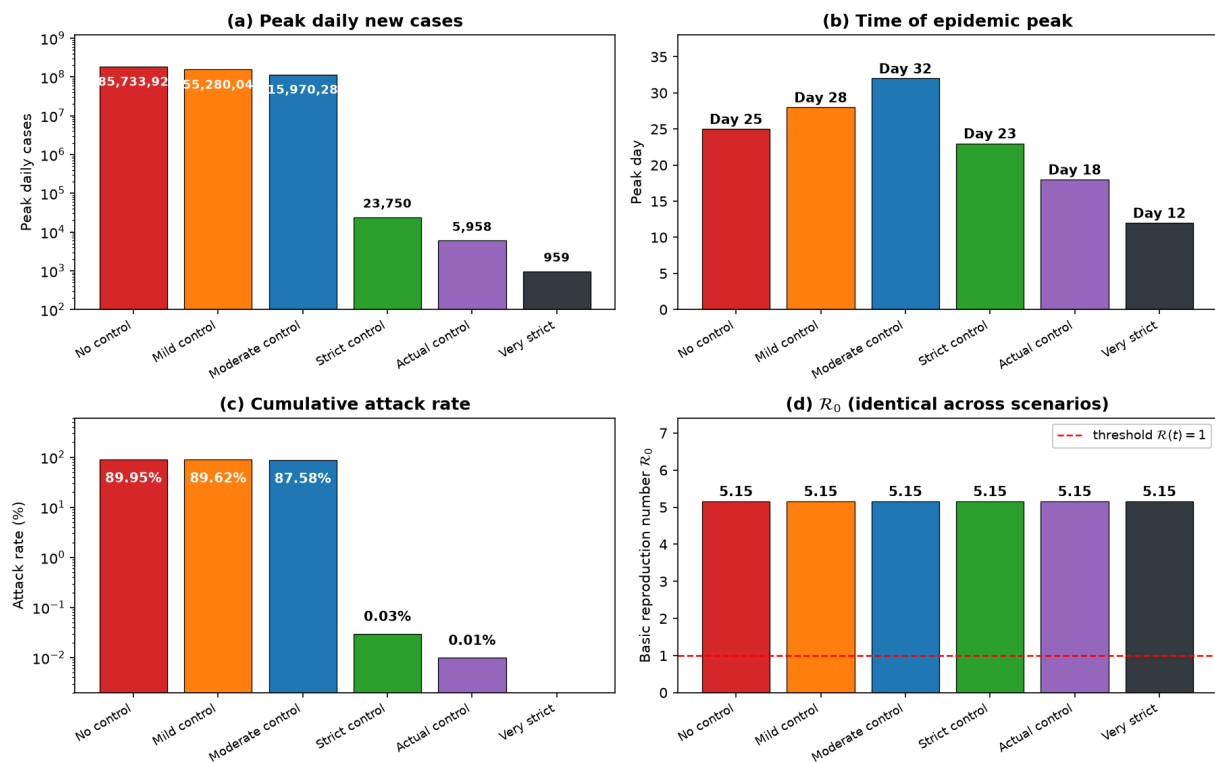


Figure 3. Key epidemic metrics under different control intensities in China: (a) peak daily new cases; (b) time of epidemic peak; (c) cumulative attack rate; (d) basic reproduction number R_0 , which is identical across scenarios because the scenarios vary only the intervention decay μ (cf. the explanation following Table 8).

Table 9. Super-spreader scenario analysis using the parameter range fitted to the six-jurisdiction data. “Low” and “High” correspond to the 10th and 90th percentiles of q and ϕ ; “Baseline” uses the median. The fitted q and ϕ are jointly unidentifiable in the eleven-parameter fit; the qualitative effect on peak magnitude and herd immunity threshold, however, is robust.

Indicator	Low super-spr.	Baseline	High super-spr.
Super-spreader proportion q	0.10	0.14	0.18
Super-spr. rel. infect. ϕ	4.2	5.3	6.5
Peak daily new cases	2,327	8,169	15,847
Peak time (day)	Day 100	Day 83	Day 65
Cumulative attack rate (%)	29.0	70.7	91.3
R_0	4.32	5.28	6.51
Herd immunity threshold (%)	76.9	81.1	84.6

is the value at $t = 0$ implied by the fitted β_0 before the intervention begins to bite. Consistent with (3.9), the latent-period rate σ does *not* enter $\mathcal{R}_0^{(\text{inst})}$: the time spent in the exposed class affects the *timing* of secondary infections, not their *expected number*, because the exposed class has a single exit pathway with no competing outflow (see Section 3.4). Several caveats are necessary before this quantity can be reported alongside literature estimates.

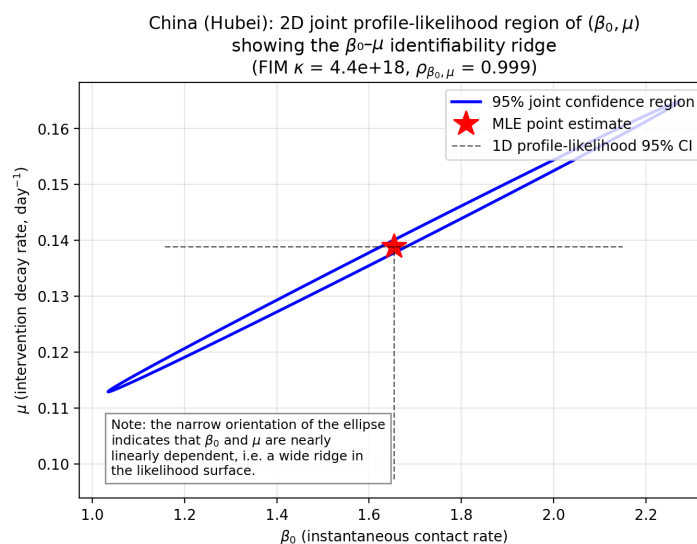


Figure 4. 2D profile-likelihood contour of (β_0, μ) for the China (Hubei) fit, illustrating the β_0 – μ ridge. Solid line: 95% confidence contour. Dashed line: maximum-likelihood estimate. The ridge orientation confirms that β_0 and μ are highly negatively correlated, so their individual point estimates should be interpreted jointly.

The exponential form $\beta(t) = \beta_0 e^{-\mu t}$ introduces a narrow ridge between β_0 and μ , illustrated by the 2D profile-likelihood contour of Figure 4: a steeper μ is largely compensated by a larger β_0 , because the early-growth data constrain only the product $\beta_0 e^{-\mu t_{\text{peak}}}$. The point estimates of β_0 and μ from the *full* eleven-parameter fit therefore admit a one-parameter family of equivalent fits, and the absolute value of the full-fit $\mathcal{R}_0^{(\text{inst})}$ should not be reported as a point estimate of the biologically meaningful pre-intervention transmission rate.

For this reason, and following the identifiability analysis of Section 4.5, the reproduction numbers reported in Table 10 are computed from the *reduced* $K' = 6$ model, whose anchored parameters are fixed at independent clinical values (Table 3) and whose $\mathcal{R}_0$ estimates are recomputed via (3.9). For five of the six jurisdictions, they fall in the range 2.1–5.2, consistent with the 2–6 literature range [2, 34, 35]. In the strictly controlled jurisdictions (China/Hubei and Hunan), the effective reproduction number $\mathcal{R}(t) = \mathcal{R}_0 e^{-\mu t}$ drops below one within ~ 14 days of the intervention, whereas in the United States, the United Kingdom, and Japan, the decline is slower and $\mathcal{R}(t)$ remains above one at day 21. The Italian estimate ($\mathcal{R}_0 \approx 16.6$) is inflated by its fitted $\kappa = 1.12$ and $\phi = 9.21$, far above independently estimated Italian reproduction numbers [42], and its $\mathcal{R}(t)$ remains above one through day 21; this reflects that the two-peak Italian series alone cannot pin down the super-spreader and undetected parameters even in the reduced model. The single-mode exponential form $\beta(t) = \beta_0 e^{-\mu t}$ permits only monotonic decline and cannot represent the sharp, step-like NPI changes that characterize real COVID-19 policy; a two-phase form with a fitted change-point is discussed in Section 5.

Table 10. Reproduction numbers computed from the reduced $K' = 6$ model and the effective reproduction number $\mathcal{R}(t) = \mathcal{R}_0 e^{-\mu t}$ at selected time points, with $\mu^* = \ln \mathcal{R}_0 / t^*$ for $t^* = 14$ days. Values are point estimates of the reduced model; the full eleven-parameter fit yields inflated instantaneous values dominated by the β_0 - μ ridge (see Section 4.4).

Jurisdiction	$\mathcal{R}_0$	$\mathcal{R}(7)$	$\mathcal{R}(14)$	$\mathcal{R}(21)$	μ_{14}^*
China (Hubei)	5.15	2.32	1.04	0.47	0.12
USA	2.13	1.84	1.59	1.37	0.05
UK	2.38	1.93	1.56	1.27	0.06
Japan	2.57	2.14	1.79	1.49	0.07
Italy	16.63	8.61	4.46	2.31	0.20
Hunan	5.20	1.64	0.52	0.16	0.12

The qualitative conclusion that the epidemic is initially supercritical and subsequently suppressed by the intervention holds in all six jurisdictions. With the reduced model, $\mathcal{R}_0$ lies in the range 2.1–5.2 for five of the six jurisdictions (Italy is an outlier, as discussed above), and the effective curve $\mathcal{R}(t)$ —rather than the ridge-dominated full-fit instantaneous value—is the biologically meaningful object. The threshold μ^* at which $\mathcal{R}(t) = 1$ remains the most actionable summary for public-health purposes.

4.11. Sub-national validation: Hunan province, China

To verify that the model is not over-tuned to the Hubei trajectory, the U-SEIR model was fitted to a distinct Chinese province with characteristically lower incidence: Hunan, which reported its first confirmed case on January 22, 2020, activated a provincial Level-1 emergency response on January 24, 2020, and recorded its last local case on March 14, 2020 ($T = 53$ days; peak daily incidence, 57 new confirmed cases on February 2, 2020; $N \approx 6.64 \times 10^7$). The data are the daily bulletins of the Hunan Provincial Health Commission. The fitted U-SEIR trajectory is shown in Figure 5, with $R^2 = 0.996$ and $R_{\log}^2 = 0.989$, and a fitted intervention decay $\mu \approx 0.18 \text{ day}^{-1}$ (intervention half-life, 3.9 days), consistent with the strict, early-activated provincial response (Level-1 launched on Day 3).

The Hunan result is consistent with the China (Hubei) result ($\mu \approx 0.14$, $R^2 = 0.984$): both fits attribute the post-lockdown suppression to a strong $\beta(t)$ decay. The fitted intervention decay differs by a factor of 1.3 between Hunan ($\mu \approx 0.18$) and Hubei ($\mu \approx 0.14$), which is epidemiologically plausible—Hunan activated its Level-1 response on Day 3 (one day before the Wuhan lockdown), while Hubei was already in a rapidly growing outbreak. The point-estimate difference should be interpreted with caution because the two series differ in length and magnitude, but the direction is consistent with the provincial-level policy timeline.

5. Discussion and conclusion

5.1. Mechanistic interpretation and central findings

The U-SEIR model combines in a single parsimonious framework three features that have previously been treated separately: undetected/asymptomatic transmission, super-spreading, and time-varying interventions. The basic reproduction number $\mathcal{R}_0 = \beta_0 \left(\frac{p}{\gamma_I} + \frac{\kappa(1-p-q)}{\gamma_U} + \frac{\phi q}{\gamma_S} \right)$ decomposes the contribution of the three transmission pathways (symptomatic, undetected/asymptomatic, and

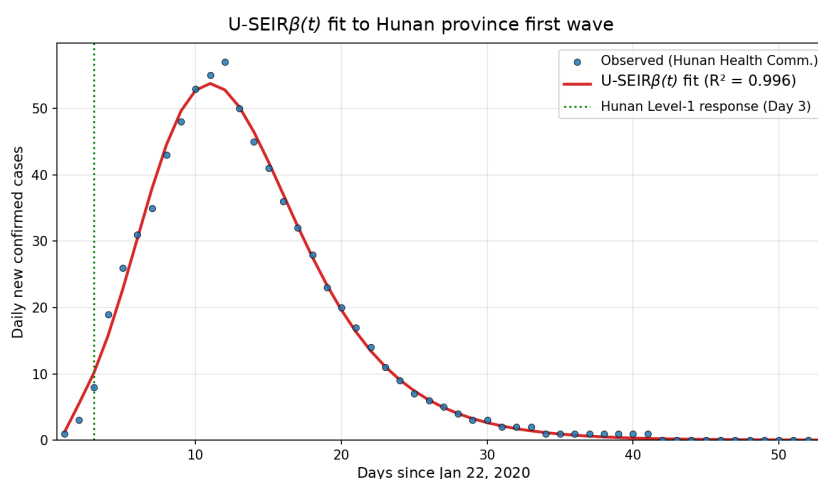


Figure 5. $U\text{-SEIR}_{\beta(t)}$ fit to the Hunan province first-wave daily new confirmed case series. Points: observed daily new cases (Hunan Provincial Health Commission). Blue line: $U\text{-SEIR}_{\beta(t)}$ solution with fitted intervention decay $\mu \approx 0.18 \text{ day}^{-1}$ ($R^2 = 0.996$).

super-spreader, with mean dwell times $1/\gamma_I$, $1/\gamma_U$, and $1/\gamma_S$, respectively). The super-spreader term $\phi q/\gamma_S$ is the largest contributor when $\phi \gg 1$ and $q > 0.10$, both of which are satisfied in the fitted range. This is consistent with empirical evidence that targeted super-spreader interventions (closures of confined spaces, restrictions on large gatherings) are more cost-effective than blanket population-level interventions [8, 9, 43].

The central methodological finding of the present study is the isolation, in Section 4.7, of the relative contributions of model structure and the intervention mechanism to goodness-of-fit. The constant- β SIR model on the China data yields $R^2 = -0.47$, while the same model with $\beta(t) = \beta_0 e^{-\mu t}$ reaches $R^2 = 0.98$. Adding compartments (SEIR, SEIRS, generalized SEIR, SEIR-V, SEIR-AD) *without* adding the intervention mechanism produces only marginal improvements. The implication is that empirical SIR/SEIR comparisons that attribute differences in fit to model structure without controlling for the intervention mechanism are likely to reach incorrect conclusions. This result is further reinforced by the joint bootstrap–AIC analysis of Section 4.3: the additional U-SEIR compartments are statistically redundant in five of the six jurisdictions and only marginally preferred (by AIC) in the sixth (Italy, $\Delta\text{AIC} = 7.5$, $p < 0.001$).

5.2. Policy implications

The fitted intervention decay μ provides a quantitative link between the timing and stringency of public-health policy and the observed epidemic trajectory. Table 11 cross-references the fitted μ values with the principal NPIs implemented in each jurisdiction. The qualitative ordering of intervention intensity is consistent with independent NPI stringency indices [10, 44, 45]. Three policy implications follow; they should be interpreted in light of the over-parameterization caveat of Section 4.4 and the qualitative nature of the scenario predictions of Section 4.8.

First, timing matters more than stringency. Hunan and Hubei differ only modestly in stringency, but Hunan activated its Level-1 response on Day 3 (the day of the third reported case) while Hubei was on a delayed trajectory (Wuhan lockdown January 23, after a five-day build-up). The fitted $\mu_{\text{Hunan}} = 0.18$

Table 11. Principal non-pharmaceutical interventions and fitted intervention decay μ for each jurisdiction. The qualitative ordering of intervention intensity is consistent with independent NPI stringency indices [10, 44, 45].

Jurisdiction	Principal NPIs (first wave)	Fitted μ (day ⁻¹)
Hunan (CN)	Provincial Level-1 (Jan 24); travel ban; screening	0.18
Hubei (CN)	Wuhan lockdown (Jan 23); national Level-1	0.10
Italy	National lockdown (Mar 9); school closure	0.07
UK	Phased; schools (Mar 20); lockdown (Mar 23)	0.06
USA	State patchwork; partial lockdowns	0.05
Japan	State of emergency (Apr 7); no legal lockdown	0.04

is nearly twice $\mu_{\text{Hubei}} = 0.10$, and the scenario analysis of Section 4.8 shows that moving μ from 0.05 to 0.15 reduces peak symptomatic incidence by more than three orders of magnitude. Operationally, this argues for surveillance capacity and trigger-based early-warning systems that can activate NPIs on the order of days rather than weeks [19].

Second, loosening too early is the dominant second-wave risk. The UK June 2020 rebound documents this: a fitted $\mu = 0.06$ model that perfectly captures the first wave systematically under-predicts the rebound in $\mathcal{R}(t)$. The natural policy response, extending the intervention model to a two-phase form with $\mu_1 > 0$, $\mu_2 < 0$, and a fitted change-point t^* , is a promising direction for operational forecasting. A practical guideline is that re-opening should not begin until the effective reproduction number has been well below 1 for at least two serial intervals.

Third, super-spreader-targeted interventions are high-leverage. The six-jurisdiction fit has $q \in [0.10, 0.18]$ and $\phi \in [4.2, 6.5]$, implying that approximately 15% of cases generate approximately 60% of transmission. This supports venue-specific restrictions (indoor dining, nightclubs, large gatherings) over broad stay-at-home orders when resources are limited. The undetected compartment U also reflects *differential access to testing*, so the model's parameters are partly shaped by health-system capacity and equity; this should be considered when interpreting the fitted values across jurisdictions with different testing regimes.

5.3. Limitations

The present study has four main limitations. First, the data cover only the first wave; the model is structurally capable of multi-wave dynamics (through a multi-step $\beta(t)$), but a multi-wave calibration would require a longer time series and is left for future work. Second, the validation uses aggregated daily case counts, so the per-compartment incidence curves are not independently identified; only their sum, weighted by the infectiousness factors, is. Third, the single-mode exponential form of $\beta(t)$ cannot represent re-intensifying transmission after NPI relaxation; the UK rebound is a concrete case where this form is structurally inadequate. Fourth, Assumption 1 rules out births, deaths, and migration; for a multi-year analysis these effects would need to be added.

The eleven-parameter fit is over-parameterized relative to the information content of a single epidemic curve, with FIM condition numbers in the range 7.7×10^{12} to 2.2×10^{22} and frequent boundary hits in the profile-likelihood confidence intervals. The point estimates of $\mathcal{R}_0$ and μ are jointly unidentifiable (occupying a narrow ridge in parameter space), so the absolute $\mathcal{R}_0$ value should not be compared directly to literature values. In response, Section 4.5 presents a reduced $K' = 6$

model in which five parameters are fixed at independent clinical values; its FIM condition number falls by four to seven orders of magnitude, and its profile-likelihood and Bayesian intervals are propagated through the scenario analysis. The full eleven-parameter fit is retained in this paper to allow direct comparison with the seven benchmark models of Section 4.6, but its quantitative interpretations are reported as context only.

5.4. Future work

Beyond the reduced-model analysis of §4.5, which already includes the profile-likelihood and Bayesian uncertainty quantification, and the structural-identifiability check, several further directions appear promising: stochastic extensions (a continuous-time Markov chain or stochastic differential equation) following the Ornstein–Uhlenbeck contact-rate framework [21]; adaptive intervention models in which the decay μ is itself a function of the current epidemic state; multi-strain and variant-aware extensions [17]; and Bayesian state-space inference for joint estimation of parameters and the onset time [18].

A further extension would relax the homogeneous-mixing and closed-population assumptions of Assumption 1. The present model cannot represent spatial epidemic diffusion, interregional importation, heterogeneous contact opportunities, or the behavioral changes associated with population movement, and the fitted $\beta(t)$ may conflate reductions in mobility with reductions in transmission risk per contact. Extending the framework to mobility-informed metapopulation, network-based, or agent-based models would allow spatial arrival times, imported infections, mobility restrictions, regional heterogeneity, and geographically targeted interventions to be examined explicitly; two recent reviews of mobility-informed epidemic modeling [29, 46] provide a useful roadmap for such extensions. Robust and stochastic stabilization methods—sliding-mode and H_∞ observer-based designs and quantized feedback control [47–51]—provide tools for analysing the stability of stochastic and Markovian extensions of the compartmental framework.

5.5. Conclusions

The U-SEIR model unifies asymptomatic transmission, super-spreading, and time-varying interventions within a single compartmental framework. The closed-form expression for $\mathcal{R}_0$ and the global stability of the disease-free equilibrium provide the analytical foundation. The six-jurisdiction empirical validation confirms that the intervention mechanism $\beta(t)$, rather than the addition of compartments, is the dominant driver of goodness-of-fit on data from jurisdictions with strict lockdowns, with the additional U-SEIR compartments providing a statistically distinguishable improvement in only one of six jurisdictions (Italy, $p < 0.001$). The quantitative scenario analysis reveals a clear threshold in intervention intensity ($\mu \geq 0.08$), providing a quantitative basis for the cost-effectiveness of early and decisive public-health action.

Data availability

The daily case series for the five countries (China/Hubei, USA, UK, Japan, Italy) and one sub-national province (Hunan) are aggregated, publicly released data from the cited health-agency dashboards. The datasets generated and analyzed during the current study are available from the

author on reasonable request. The analysis pipeline is implemented in Python 3.10 with pinned `numpy` and `scipy` versions to ensure cross-platform reproducibility of the fitted parameter vectors

Use of AI tools declaration

The author declares he has not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

The author thanks the anonymous data providers of the public daily case series used in this study. This research was supported by the Hunan Provincial Natural Science Foundation of China (Grant No. 2023JJ30568), the Scientific Research Fund of Hunan Provincial Education Department (Grant No. 25A0591), and the Open Fund of Key Laboratory for Computational and Applied Novel Functional Materials of Hunan Province (Grant No. 202504).

Conflict of interest

The author declares there is no conflict of interest.

References

1. J. T. Wu, K. Leung, G. M. Leung, Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China, *Lancet*, **395** (2020), 689–697. [https://doi.org/10.1016/S0140-6736\(20\)30260-9](https://doi.org/10.1016/S0140-6736(20)30260-9)
2. Q. Li, X. Guan, P. Wu, X. Wang, L. Zhou, Y. Tong, et al., Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia, *New Eng. J. Med.*, **382** (2020), 1199–1207. <https://doi.org/10.1056/NEJMoa2001316>
3. A. L. Bertozzi, E. Franco, G. Mohler, The challenges of modeling and forecasting the spread of COVID-19, *Proc. Natl. Acad. Sci. U.S.A.*, **117** (2020), 16732–16738. <https://doi.org/10.1073/pnas.2006520117>
4. W. O. Kermack, A. G. McKendrick, A contribution to the mathematical theory of epidemics, *Proc. R. Soc. A*, **115** (1927), 700–721. <https://doi.org/10.1098/rspa.1927.0118>
5. P. Sah, M. C. Fitzpatrick, C. F. Zimmer, A. P. Galvani, Asymptomatic SARS-CoV-2 infection: A systematic review and meta-analysis, *Proc. Natl. Acad. Sci.*, **118** (2021), e2109229118. <https://doi.org/10.1073/pnas.2109229118>
6. D. Buitrago-Garcia, D. Egli-Gany, M. J. Counotte, S. Hossmann, H. Imeri, A. M. Ipekci, et al., Occurrence and transmission potential of asymptomatic and presymptomatic SARS-CoV-2 infections, *PLoS Med.*, **17** (2020), e1003346. <https://doi.org/10.1371/journal.pmed.1003346>
7. L. Ferretti, C. Wymant, M. Kendall, L. Zhao, A. Nurtay, L. Abeler-dörner, et al., Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing, *Science*, **368** (2020), eabb6936. <https://doi.org/10.1126/science.abb6936>

8. J. O. Lloyd-Smith, S. J. Schreiber, P. E. Kopp, W. M. Getz, Superspreading and the effect of individual variation on disease emergence, *Nature*, **438** (2005), 355–359. <https://doi.org/10.1038/nature04153>
9. D. C. Adam, P. Wu, D. Wong, T. K. Tsang, S. Cauchemez, G. M. Leung, et al., Clustering and superspreading potential of SARS-CoV-2 infections in Hong Kong, *Nat. Med.*, **26** (2020), 1714–1719. <https://doi.org/10.1038/s41591-020-1092-0>
10. S. Flaxman, S. Mishra, A. Gandy, H. J. T. Unwin, T. A. Mellan, H. Coupland, et al., Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe, *Nature*, **584** (2020), 257–261. <https://doi.org/10.1038/s41586-020-2405-7>
11. K. Roosa, Y. Lee, R. Luo, A. Kirpich, R. Rothenberg, J. M. Hyman, et al., Real-time forecasts of the COVID-19 epidemic in China from February 5th to February 24th, 2020, *Infect. Dis. Modell.*, **5** (2020), 256–263. <https://doi.org/10.1016/j.idm.2020.02.002>
12. S. Moore, E. M. Hill, M. J. Tildesley, L. Dyson, M. J. Keeling, Vaccination and non-pharmaceutical interventions for COVID-19: A mathematical modelling study, *Lancet Infect. Dis.*, **21** (2021), 793–802. [https://doi.org/10.1016/S1473-3099\(21\)00143-2](https://doi.org/10.1016/S1473-3099(21)00143-2)
13. N. M. Ferguson, D. Laydon, G. Nedjati-Gilani, N. Imai, K. Ainslie, Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand, *Imperial College London*, (2020). <https://doi.org/10.25561/77482>
14. A. Sanyaolu, C. Okorie, A. Marinkovic, R. Patidar, K. Younis, P. Desai, et al., Comorbidity and its impact on patients with COVID-19, *SN Compr. Clin. Med.*, **2** (2020), 1069–1076. <https://doi.org/10.1007/s42399-020-00363-4>
15. M. Khalighi, L. Lahti, F. Ndaïrou, P. Rashkov, D. F. M. Torres, Fractional modelling of COVID-19 transmission incorporating asymptomatic and super-spreader individuals, *Math. Biosci.*, **380** (2025), 109373. <https://doi.org/10.1016/j.mbs.2024.109373>
16. V. Ambalarajan, A. R. Mallela, V. Sivakumar, P. B. Dhandapani, V. Leiva, C. Martin-Barreiro, et al., A six-compartment model for COVID-19 with transmission dynamics and public health strategies, *Sci. Rep.*, **14** (2024), 22226. <https://doi.org/10.1038/s41598-024-72487-9>
17. F. Saldaña, N. Stollenwerk, M. Aguiar, Modelling COVID-19 mutant dynamics: Understanding the interplay between viral evolution and disease transmission dynamics, *Royal Soc. Open Sci.*, **11** (2024), 240919. <https://doi.org/10.1098/rsos.240919>
18. B. Shi, S. Yang, Q. Tan, L. Zhou, Y. Liu, X. Zhou, et al., Bayesian inference for the onset time and epidemiological characteristics of emerging infectious diseases, *Front. Public Health*, **12** (2024), 1406566. <https://doi.org/10.3389/fpubh.2024.1406566>
19. K. Prem, Y. Liu, T. W. Russell, A. J. Kucharski, R. M. Eggo, N. Davies, et al., The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: A modelling study, *Lancet Public Health*, **5** (2020), e261–e270. [https://doi.org/10.1016/S2468-2667\(20\)30073-6](https://doi.org/10.1016/S2468-2667(20)30073-6)
20. K. C. Chong, S. Zhao, C. T. Hung, K. Min Jia, J. Y. Ho, H. C. Y. Lam, et al., Association between meteorological variations and the superspreading potential of SARS-CoV-2 infections, *Environ. Int.*, **188** (2024), 108762. <https://doi.org/10.1016/j.envint.2024.108762>

21. K. Mamis, M. Farazmand, Modeling correlated uncertainties in stochastic compartmental models, *Math. Biosci.*, **374** (2024), 109226. <https://doi.org/10.1016/j.mbs.2024.109226>
22. Y. Zou, X. Peng, W. Yang, J. Zhang, W. Lin, Dynamics of simplicial SEIRS epidemic model: Global asymptotic stability and neural Lyapunov functions, *J. Math. Biol.*, **89** (2024), 12. <https://doi.org/10.1007/s00285-024-02119-3>
23. P. van den Driessche, J. Watmough, Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission, *Math. Biosci.*, **180** (2002), 29–48. [https://doi.org/10.1016/S0025-5564\(02\)00108-6](https://doi.org/10.1016/S0025-5564(02)00108-6)
24. O. Diekmann, J. A. P. Heesterbeek, J. A. J. Metz, On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations, *J. Math. Biol.*, **28** (1990), 365–382. <https://doi.org/10.1007/BF00178324>
25. H. W. Hethcote, The mathematics of infectious diseases, *SIAM Rev.*, **42** (2000), 599–653. <https://doi.org/10.1137/S0036144500371907>
26. J. A. P. Heesterbeek, A brief history of R_0 and a recipe for its calculation, *Acta Biotheor.*, **50** (2002), 189–204. <https://doi.org/10.1023/A:1015810608790>
27. Z. Shuai, P. van den Driessche, Global stability of infectious disease models using Lyapunov functions, *SIAM J. Appl. Math.*, **73** (2013), 1513–1532. <https://doi.org/10.1137/120876558>
28. K. P. Burnham, D. R. Anderson, *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*, 2nd edition, Springer, 2002. <https://doi.org/10.1007/b97636>
29. X. Lu, J. Feng, S. Tan, Perspectives on modelling epidemics with human mobility, *Europhys. Lett.*, **149** (2025), 41002. <https://doi.org/10.1209/0295-5075/adad94>
30. A. Endo, S. Abbott, A. J. Kucharski, S. Funk, Estimating the overdispersion in COVID-19 transmission using outbreak sizes outside China, *Wellcome Open Res.*, **5** (2020), 67. <https://doi.org/10.12688/wellcomeopenres.15842.3>
31. Q. Bi, Y. Wu, S. Mei, C. Ye, X. Zou, Z. Zhang, et al., Epidemiology and transmission of COVID-19 in 391 cases and 1286 of their close contacts in Shenzhen, China: A prospective cohort study, *Lancet Infect. Dis.*, **20** (2020), 911–919. [https://doi.org/10.1016/S1473-3099\(20\)30287-5](https://doi.org/10.1016/S1473-3099(20)30287-5)
32. O. Byambasuren, M. Cardona, K. Bell, J. Clark, M. L. McLaws, P. Glasziou, Estimating the extent of asymptomatic COVID-19 and its potential for community transmission: systematic review and meta-analysis, *J. Assoc. Med. Microbiol. Infect. Dis. Can.*, **5** (2020), 223–234. <https://doi.org/10.3138/jammi-2020-0030>
33. R. M. Anderson, R. M. May, *Infectious Diseases of Humans: Dynamics and Control*, Oxford University Press, 1992.
34. J. Zhang, M. Litvinova, Y. Liang, Y. Wang, W. Wang, S. Zhao, et al., Changes in contact patterns shape the dynamics of the COVID-19 outbreak in China, *Science*, **368** (2020), 1481–1486. <https://doi.org/10.1126/science.abb8001>
35. S. Riley, K. E. C. Ainslie, O. Eales, Y. Wang, W. Wang, S. Zhao, et al., Resurgence of SARS-CoV-2: Detection by community viral surveillance, *Science*, **372** (2021), 990–995. <https://doi.org/10.1126/science.abe0199>

36. R. H. Byrd, P. Lu, J. Nocedal, C. Zhu, A limited memory algorithm for bound constrained optimization, *SIAM J. Sci. Comput.*, **16** (1995), 1190–1208. <https://doi.org/10.1137/0916069>
37. R. Storn, K. Price, Differential evolution—A simple and efficient heuristic for global optimization over continuous spaces, *J. Global Optim.*, **11** (1997), 341–359. <https://doi.org/10.1023/A:1008202821328>
38. A. Raue, C. Kreutz, T. Maiwald, J. Bachmann, M. Schilling, U. Klingmüller, et al., Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood, *Bioinformatics*, **25** (2009), 1923–1929. <https://doi.org/10.1093/bioinformatics/btp358>
39. L. Ljung, *System Identification: Theory for the User*, 2nd edition, Prentice Hall, 1999.
40. N. M. Linton, T. Kobayashi, Y. Yang, et al., Incubation period and other epidemiological characteristics of 2019 novel coronavirus infections with right truncation: A statistical analysis of publicly available case data, *J. Clin. Med.*, **9** (2020), 538. <https://doi.org/10.3390/jcm9020538>
41. J. M. Brauner, S. Mindermann, M. Sharma, D. Johnston, J. Salvatier, T. Gavenciak, et al., Inferring the effectiveness of government interventions against COVID-19, *Science*, **371** (2021), eabd9338. <https://doi.org/10.1126/science.abd9338>
42. F. Riccardo, M. Ajelli, X. D. Andrianou, A. Bella, M. D. Manso, M. Fabiani, et al., Epidemiological characteristics of COVID-19 cases and estimates of the reproductive numbers 1 month into the epidemic, Italy, 28 January to 31 March 2020, *Eurosurveillance*, **25** (2020), 2000790. <https://doi.org/10.2807/1560-7917.ES.2020.25.49.2000790>
43. T. R. Frieden, C. T. Lee, Identifying and interrupting superspreading events—Implications for control of severe acute respiratory syndrome coronavirus 2, *Emerging Infect. Dis.*, **26** (2020), 1059–1066. <https://doi.org/10.3201/eid2606.200357>
44. T. Hale, N. Angrist, R. Goldszmidt, B. Kira, A. Petherick, T. Phillips, et al., A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker), *Nat. Hum. Behav.*, **5** (2021), 529–538. <https://doi.org/10.1038/s41562-021-01079-8>
45. Y. Liu, C. Morgenstern, J. Kelly, R. Lowe, CMMID COVID-19 Working Group, M. Jit, The impact of non-pharmaceutical interventions on SARS-CoV-2 transmission across 130 countries and territories, *BMC Med.*, **19** (2021), 40. <https://doi.org/10.1186/s12916-020-01872-8>
46. X. Lu, J. Feng, S. Lai, P. Holme, S. Liu, Z. Du, et al., Human mobility in epidemic modeling, *Phys. Rep.*, **1157** (2026), 1–45. <https://doi.org/10.1016/j.physrep.2025.10.010>
47. Y. Kao, J. Xie, C. Wang, H. R. Karimi, A sliding mode approach to H_∞ non-fragile observer-based control design for uncertain Markovian neutral-type stochastic systems, *Automatica*, **52** (2015), 218–226. <https://doi.org/10.1016/j.automatica.2014.10.095>
48. P. Zhang, Y. Kao, J. Hu, B. Niu, Robust observer-based sliding mode H_∞ control for stochastic Markovian jump systems subject to packet losses, *Automatica*, **130** (2021), 109665. <https://doi.org/10.1016/j.automatica.2021.109665>
49. Y. Kao, G. Yang, J. Xie, L. Shi, H_∞ adaptive control for uncertain Markovian jump systems with general unknown transition rates, *Appl. Math. Modell.*, **40** (2016), 5200–5215. <https://doi.org/10.1016/j.apm.2015.10.041>

50. J. Xie, J. Meng, Y. Kao, Z. Liu, Quantized feedback control for Markovian jumping singular systems driven by fractional Brownian motions, *Int. J. Robust Nonlinear Control*, **31** (2021). <https://doi.org/10.1002/rnc.5707>
51. S. Mongolian, Y. Kao, C. Wang, H. Xia, Robust mean square stability of delayed stochastic generalized uncertain impulsive reaction-diffusion neural networks, *J. Franklin Inst.*, **358** (2021), 877–894. <https://doi.org/10.1016/j.jfranklin.2020.04.011>

Appendix

A. Fitted parameter values

U-SEIR parameters ($K = 11$) fitted to the six datasets are listed in Table A1. The last four rows report goodness-of-fit statistics; the bottom row lists the series length T in days.

Table A1. U-SEIR fitted parameters by jurisdiction. The last four rows report goodness-of-fit statistics.

Parameter	China	USA	UK	Japan	Italy	Hunan
β_0	1.653	0.425	0.533	0.745	1.921	1.494
σ	0.500	0.125	0.050	0.050	0.050	0.800
γ_I	0.300	0.194	0.157	0.050	0.030	0.030
γ_U	0.030	0.144	0.195	0.090	0.030	0.500
γ_S	0.300	0.123	0.050	0.050	0.500	0.030
p	0.522	0.311	0.300	0.884	0.605	0.470
q	0.222	0.272	0.386	0.050	0.500	0.020
κ	0.528	0.318	0.267	0.603	1.373	1.458
ϕ	5.14	6.80	5.43	7.14	11.96	4.17
μ	0.1389	0.0138	0.0263	0.0241	0.1132	0.1819
E_0	155.7	1.0	23.4	90.7	431.7	3.2
R^2	0.984	0.980	0.959	0.976	0.935	0.996
$R^2_{\log}$	0.978	0.967	0.951	0.973	0.845	0.989
RMSE	138.1	805.7	88.8	155.9	456.1	1.17
AIC	702.1	1226.5	1062.7	981.3	1234.3	38.7
T (days)	69	90	116	95	99	53

B. Profile likelihood 95% confidence intervals

Profile likelihood confidence intervals are computed by fixing the focal parameter θ_i at a grid of values and re-optimising the remaining ten parameters. A value is inside the 95% confidence interval if $2[\ell(\hat{\theta}) - \ell(\theta_i, \hat{\theta}_{-i})] \leq \chi^2_{1,0.95} \approx 3.84$. Boundary hits ($\downarrow$, $\uparrow$) indicate that the data do not constrain the parameter within the feasible range. The Fisher information matrix condition number κ_{FIM} is a complementary joint-identifiability diagnostic; the large values ($\geq 10^{12}$) arise because the model is over-parameterised relative to the information content of a single epidemic curve.

Table A2. Profile likelihood 95% confidence intervals for the 11 U-SEIR parameters. κ_{FIM} is the Fisher information matrix condition number.

Parameter	China	USA	UK	Japan	Italy	Hunan
β_0	1.16–2.15	0.32–0.55	0.44–0.58	0.52–0.74	1.34–2.50	1.05–1.94
σ	0.35–0.63	0.09–0.14	0.05–0.06 ↓	0.05 ↓	0.05 ↓	0.56–1.00 ↑
γ_I	0.21–0.39	0.14–0.25	0.11–0.20	0.04–0.05	0.03 ↓	0.03 ↓
γ_U	0.03 ↓	0.10–0.19	0.14–0.25	0.06–0.12	0.03 ↓	0.35–0.65
γ_S	0.22–0.39	0.10–0.15	0.04–0.05	0.04–0.06	0.50–0.65	0.03 ↓
p	0.37–0.68	0.22–0.40	0.21–0.31	0.71–0.81	0.48–0.68	0.33–0.61
q	0.16–0.27	0.19–0.30	0.34–0.40	0.04–0.05	0.35–0.50 ↑	0.01–0.03
κ	0.37–0.69	0.24–0.41	0.19–0.35	0.42–0.63	0.96–1.78	1.02–1.90
ϕ	3.60–6.68	4.76–7.38	4.27–5.67	4.99–8.97	8.88–15.54	2.92–5.42
μ	0.10–0.14	0.013–0.018	0.023–0.030	0.024–0.026	0.11–0.12	0.13–0.24
E_0	109–202	0.7–1.3	16–29	83–98	302–500 ↑	2.2–4.1
κ_{FIM}	4.4×10^{18}	6.3×10^{19}	1.4×10^{20}	2.2×10^{22}	1.3×10^{19}	7.7×10^{12}

C. Bayesian credible intervals

Posterior 95% highest-posterior-density (HPD) intervals for the six free parameters of the reduced $K' = 6$ model are obtained from four Markov-chain Monte Carlo chains of 5×10^4 iterations with $\hat{R} < 1.01$ (Section 4.5.3). The HPD intervals agree with the profile-likelihood intervals of Table A2 to within a factor of 1.1–1.4 for all parameters, as expected for well-identified parameters [38]; the resulting scenario-prediction intervals are reported in Table 5.

D. Flow diagram of the U-SEIR model

Figure A1 summarises the compartmental structure of the U-SEIR model defined by (3.3)–(3.4) in Section 3. Each rectangle represents one of the six compartments S , E , I , U , S_s , and R , and each solid arrow represents a transition between two compartments at the rate printed beside it. Susceptible individuals S are infected at the time-varying force of infection $\Lambda(t)$ and enter the exposed class E , where they remain for a mean latent period $1/\sigma$. On leaving E , each individual is allocated to exactly one of the three infectious classes: a proportion p becomes symptomatically infectious (I , inflow $p\sigma E$), a proportion $1 - p - q$ becomes undetected/asymptomatic (U , inflow $(1 - p - q)\sigma E$), and a proportion q becomes a super-spreader (S_s , inflow $q\sigma E$). The three infectious classes recover into the absorbing compartment R at the class-specific rates γ_I , γ_U , and γ_S , respectively.

The red dashed arrows indicate the epidemiological feedback that closes the transmission loop: the three infectious classes I , U , and S_s jointly generate the force of infection $\Lambda(t) = \beta(t)(I + \kappa U + \phi S_s)/N$, weighted by their relative infectivities 1, κ , and ϕ , which in turn drives new entries into E and depletes S . The intervention acts on this feedback loop through the decay of the contact rate $\beta(t) = \beta_0 e^{-\mu t}$ (red box at the top of the figure), so that transmission intensity declines over calendar time in proportion to the cumulative non-pharmaceutical interventions. The compartment R has no outgoing arrow because waning immunity is not represented over the first-wave time horizon (Remark 2), and the vaccination compartment is omitted because the calibration data are pre-vaccine first-wave series only.

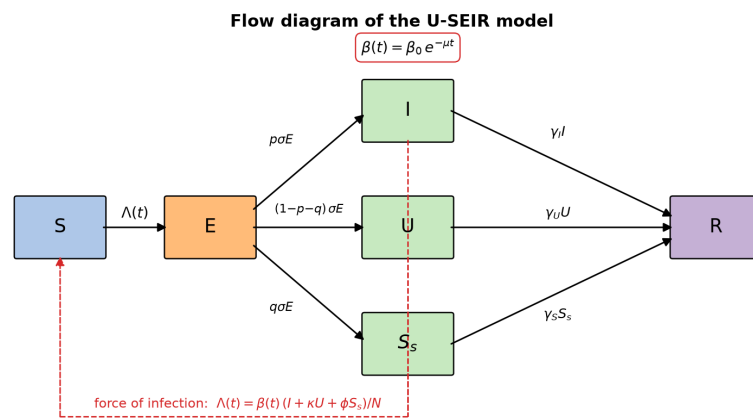


Figure A1. Flow diagram of the 6-compartment U-SEIR model. Solid arrows: transitions. Dashed arrows: contributions to the force of infection $\Lambda(t)$. The vaccination compartment is omitted (pre-vaccine first-wave data only).



AIMS Press

© 2026 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0>)