Supplementary Material for
“Characterizing the COVID-19 Dynamics with a New Epidemic Model: Susceptible-Exposed-Symptomatic-Asymptomatic-Active-Removed”

Section S1: Additional Information for the SEASAR Model

Figure S.1: Dynamic of the transmission among the six subpopulations at time \( t \)
Table S.1: Initial state size for the SEASAR model

| Variable | Meaning | Value |
|----------|---------|-------|
| $S(\tau_0)$ | The initial number of susceptible cases | $N - E(\tau_0) - I_a(\tau_0) - I_s(\tau_0) - A(\tau_0) - R(\tau_0)$ |
| $E(\tau_0)$ | The initial number of exposed cases | $r_3 \times \sum_{t=\tau_0}^{\lfloor \tau_0 + Z \rfloor} Q(t)$ |
| $I_a(\tau_0)$ | The initial number of asymptomatic infections | $r_2 \{ \sum_{t=\tau_0}^{t \leq \tau_0} Q(t) - C_0 \}$ |
| $I_s(\tau_0)$ | The initial number of symptomatic infections | $\sum_{t=\tau_0}^{t \leq \tau_0} Q(t) - C_0$ |
| $A(\tau_0)$ | The initial number of active cases | $C_0 - R_c - D_0$ |
| $R(\tau_0)$ | The initial number of removed cases | $D_0 + (1 + r_1)R_c$ |

Section S2: Appendices of Technical Derivations

Appendix A: Expression of the Basic Reproduction Number

Following the arguments outlined in Diekmann et al. (2010), we derive the basic reproduction number $R_0$. The basic idea is to extract the information on the occurrence of new infections from the model, and then calculate the expected number of new infections generated from a case in the population under the constraint $S(t) = N$.

Since a confirmed case must be quarantined, any individuals in the compartment for active cases cannot be infectious. In addition, any individuals in the compartments for the susceptible individuals or the removed cases cannot be infectious. Therefore, only Equations (2), (3) and
(4) in the SEASAR model provide information concerning new infections generated by infected cases who are not yet confirmed, and thus are not being quarantined. With the assumption that all individuals in the population are susceptible (i.e., setting $S(t) = N$), Equations (2)-(4) become

$$\frac{dE(t)}{dt} = \theta I_s(t) + \mu \theta I_a(t) - \frac{E(t)}{Z};$$  \hspace{1cm} (S.1)

$$\frac{dI_a(t)}{dt} = \left(1 - \frac{\alpha}{Z}\right) \frac{E(t)}{Z} - \beta I_a(t) - \gamma I_a(t);$$  \hspace{1cm} (S.2)

$$\frac{dI_s(t)}{dt} = \frac{\alpha E(t)}{Z} - \frac{I_s(t)}{F} + \beta I_a(t);$$  \hspace{1cm} (S.3)

which can be written in the equivalent compact form

$$\frac{d\phi(t)}{dt} = G\phi(t),$$  \hspace{1cm} (S.4)

where $\phi(t) = (E(t), I_a(t), I_s(t))^T$ and $G = \begin{bmatrix} -\frac{1}{Z} & \mu \theta & \theta \\ \frac{1}{Z} \beta & -(\beta + \gamma) & 0 \\ \frac{1}{Z} & \beta & -\frac{1}{F} \end{bmatrix}$.

Note that only Equation (S.1) includes the terms, $\theta I_s(t) + \mu \theta I_a(t)$, which represents the occurrence of new infections; now rewrite the matrix $G$ as $G = G_1 + G_2$, where $G_1 = \begin{bmatrix} 0 & \mu \theta & \theta \\ \frac{1}{Z} \beta & -(\beta + \gamma) & 0 \\ \frac{1}{Z} & \beta & -\frac{1}{F} \end{bmatrix}$, $G_2 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$.

Although the definition of $R_0$ is conceptually intuitive, its determination is mathematically complicated (Delamater et al., 2019). Diekmann et al. (1990) derived $R_0$ as the dominant eigenvalue of the next-generation matrix which, for the Equations (S.4), we can show equals $-G_1G_2^{-1}$ by adapting the arguments of Diekmann et al. (2010). Since

$$G_2^{-1} = \begin{bmatrix} -\frac{Z}{\beta + \gamma} & 0 & 0 \\ \frac{1}{\beta + \gamma} & -\frac{1}{\beta + \gamma} & 0 \\ \frac{-\alpha}{\beta + \gamma} & \frac{-\alpha}{\beta + \gamma} F & -F \end{bmatrix},$$

it follows that $-G_1(G_2)^{-1} = \begin{bmatrix} -\frac{\theta (\alpha - 1)}{\beta + \gamma} & \frac{\theta (\alpha + \beta + \gamma) F}{\beta + \gamma} & \frac{\mu \theta}{\beta + \gamma} + \frac{\theta \beta F}{\beta + \gamma} & \theta F \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, which has the dom-
\[ \text{inant } \frac{\theta(F_\alpha + \beta F - \mu \alpha + \mu)}{\beta + \gamma}. \] Therefore, \( R_0 = \frac{\theta(F_\alpha + \beta F - \mu \alpha + \mu)}{\beta + \gamma}. \]

**Appendix B: Establishing Equations (1)-(6)**

In this appendix, we outline the logic that justifies Equations (1)-(6) found in Section 2.2 of the associated article.

For \( \Delta t > 0 \), define \( N_{SE}(t, \Delta t) \) as the average number of individual transitions from \( S \) to \( E \) between time \( t \) and time \( t + \Delta t \). Similar results apply to \( N_{EI_a}(t, \Delta t) \), \( N_{EI_s}(t, \Delta t) \), \( N_{I_aI_s}(t, \Delta t) \), \( N_{I_aR}(t, \Delta t) \), \( N_{I_sA}(t, \Delta t) \), and \( N_{AR}(t, \Delta t) \). According to the flowchart found in Figure 1 in the primary article, it is clear that the change in the average number among the six subpopulations between \( t \) and \( t + \Delta t \) can be expressed as

\[
S(t + \Delta t) - S(t) = -N_{SE}(t, \Delta t); \tag{S.5}
\]
\[
E(t + \Delta t) - E(t) = N_{SE}(t, \Delta t) - N_{EI_a}(t, \Delta t) - N_{EI_s}(t, \Delta t); \tag{S.6}
\]
\[
I_a(t + \Delta t) - I_a(t) = N_{EI_a}(t, \Delta t) - N_{I_aI_s}(t, \Delta t) - N_{I_aR}(t, \Delta t); \tag{S.7}
\]
\[
I_s(t + \Delta t) - I_s(t) = N_{EI_s}(t, \Delta t) - N_{I_sA}(t, \Delta t) + N_{I_aI_s}(t, \Delta t); \tag{S.8}
\]
\[
A(t + \Delta t) - A(t) = N_{I_sA}(t, \Delta t) - N_{AR}(t, \Delta t); \tag{S.9}
\]
\[
R(t + \Delta t) - R(t) = N_{I_aR}(t, \Delta t) + N_{AR}(t, \Delta t). \tag{S.10}
\]

Now we examine the quantities on the right hand side of Equations (S.5)-(S.10) in terms of the parameters defined in Section 2.1 of the primary article. According to the definition of \( \theta \), during the interval \([t, t + \Delta t]\), the product \( \theta I_s(t) \Delta t \) represents the average number of people infected by symptomatic cases and \( S(t)/N \) is the proportion of susceptible cases, so \( \theta S(t) I_s(t) \Delta t / N \) equals the average number of susceptible cases who are infected by symptomatic cases. Similarly, \( \mu \theta S(t) I_a(t) \Delta t / N \) equals the average number of susceptible cases who are infected by asymptomatic individuals. It follows that
\[ N_{SE}(t, \Delta t) = \frac{\theta S(t) I_a(t) \Delta t}{N} + \frac{\mu \theta S(t) I_a(t) \Delta t}{N}. \]  \hfill (S.11)

Since \( \frac{E(t) \Delta t}{Z} \) represents the average number of exposed cases who complete the latent period and become infectious during the time interval \( [t, t + \Delta t] \), it follows from the definition of \( \alpha \) that

\[ N_{EI_a}(t, \Delta t) = (1 - \alpha) \frac{E(t)}{Z} \Delta t; \]  \hfill (S.12)
\[ N_{EI_s}(t, \Delta t) = \alpha \frac{E(t)}{Z} \Delta t. \]  \hfill (S.13)

Likewise, by the definition of \( \beta \), \( \beta I_a(t) \Delta t \) represents the average number of infections who are initially asymptomatic but later exhibit symptoms during the time interval \( [t, t + \Delta t] \), so

\[ N_{I_a I_s}(t, \Delta t) = \beta I_a(t) \Delta t. \]  \hfill (S.14)

Similarly, according to the definition of \( \gamma \), \( \gamma I_a(t) \Delta t \) equals the average number of asymptomatic cases who recover from COVID-19, so

\[ N_{I_a R}(t, \Delta t) = \gamma I_a(t) \Delta t. \]  \hfill (S.15)

According to the definition of \( F \), \( I_s(t) \Delta t / F \) represents the average number of symptomatic cases whose infections are confirmed during the interval \( [t, t + \Delta t] \); therefore

\[ N_{I_s A}(t, \Delta t) = \frac{I_s(t)}{F} \Delta t. \]  \hfill (S.16)

Finally, from the definitions of \( B \) and \( J \), \( A(t) \Delta t / B \) represents the average number of active cases who recover from COVID-19 and \( A(t) \Delta t / J \) equals the average number of active cases who die from COVID-19 during \( [t, t + \Delta t] \); thus

\[ N_{AR}(t, \Delta t) = \left( \frac{A(t)}{B} + \frac{A(t)}{J} \right) \Delta t. \]  \hfill (S.17)
By substituting Equations (S.11)-(S.17) into Equations (S.5)-(S.10), next dividing each resulting equation by $\Delta t$ and then letting $\Delta t$ approach zero, we obtain Equations (1)-(6) in the primary article.

**Appendix C: The Logic Underlying Algorithm 1**

Iterated filtering (IF) is a useful method for estimating the time-invariant parameters of partially-observed nonlinear stochastic dynamic systems, also known as state-space models (Ionides, Bretó & King 2006). In our framework, the state-space model consists of an unobserved process, $\{\phi(t) : t > 0\}$, and a partially observable process $\{Y(t) : t > 0\}$ with observed values $D_c = \{y_t : t \in \mathcal{T}\}$ at discrete time points $\mathcal{T}$. In situations where the likelihood function does not have a tractable form, the IF method basically produces estimates of the model parameters sequentially using partially observed data for a dynamic process. A computationally efficient IF approach uses the ensemble adjustment Kalman filter (EAKF) as the filtering technique. The iterated filtering-ensemble adjustment Kalman filter (IF-EAKF) algorithm is a plug-and-play inference method which has been useful for non-linear stochastic dynamic models. It has been discussed by many authors such as Bretó et al. (2009), Ionides et al. (2006), and He, Ionides & King (2010).

Here we adapt the IF-EAKF algorithm in combination with the fourth-order Runge-Kutta (RK4) method (e.g., Süli & Mayers 2003, p.328) to estimate the SEASAR model parameters $\eta$. The basic idea is to embed the original SEASAR model with the *time-invariant* parameter $\eta$ in an expanded model with a *time-varying* parameter, say $\eta(t)$, so that estimation of the original parameter $\eta$ becomes estimation of $\eta(t)$, where the enhanced model assumes the same form as that specified in Equations (1)-(6) in the primary article with the parameter $\eta$ replaced by $\eta(t) = (\theta(t), \mu(t), \alpha(t), \beta(t), \gamma(t), F(t), B(t), J(t))^T$. 
Using the terminology of the “joint state-observation vector” (Anderson 2001, p.2886) which includes time-dependent variables, let the joint state-observation vector at time $t$ be $O(t) = (\eta(t)^T, \phi(t)^T, \mu_c(t))^T$. Using prior distributions for the elements of $\eta(\tau_0)$ as well as the initial sizes of $\phi(\tau_0)$, we aim to evaluate the posterior distribution of the joint state-observation vector $O(t)$ sequentially for $t \in \mathcal{T}$ based on the observed data $\mathcal{D}_c = \{y_t : t \in \mathcal{T}\}$. To reduce the associated computational costs, we adapt the discussion of (Anderson, 2001, p.2888) and consider a simple implementation procedure by examining $O(t)$ via its paired subvectors separately, where we pair $\mu_c(t)$ with each of the elements in $\eta(t)$ and $\phi(t)$, and then approximate the posterior distributions for those pairs. Such a pairing procedure allows us to update estimates of the model parameters dynamically using the sequentially collected data $\mathcal{D}_c$, with the stochastic variability accommodated via the distribution identified in Equation (7) of the primary article. We start by assuming a normal prior distribution for each pair, and then evaluate its posterior distribution using the detailed specifications identified in Equation (7) and the observed data $\mathcal{D}_c$. A detailed derivation of the posterior distribution of the pair of $\mu_c(t)$ and $\theta(t)$ may be found in Appendix D.

**Appendix D: Deriving the Posterior Distribution of $(\theta(t), \mu_c(t))$**

Rather than attempt to find an analytic expression for the posterior distribution of $(\theta(t), \mu_c(t))^T$ for $t \geq \tau_0$, the EAKF algorithm generates posterior values of $(\theta(t), \mu_c(t))^T$ by using prior values of $(\theta(t), \mu_c(t))^T$ and the observation $y_t$. Let $Z_t = (\theta(t), \mu_c(t))^T$, and let $H = (0, 1)$ so that $HZ_t = \mu_c(t)$. For ease of exposition, in this appendix we may sometimes write $U = Y(t)$ and $Z = Z_t$ for any given $t \geq \tau_0$. Let $H_1 = (1, 0)$. Then since $Z = (\theta(t), \mu_c(t))^T$ and $H = (0, 1)$, we can write $\theta(t) = H_1Z$ and $\mu_c(t) = HZ$.

Now we show that $U$ is conditionally independent of $H_1Z$, given $HZ$. In fact, Equation (7)
\[ U = HZ + \epsilon_t, \]

where \( \epsilon_t \) follows the normal distribution with zero mean and variance as specified in Equation (8); moreover, \( U \) is independent of \( H_1Z = \theta(t) \) conditioned on \( HZ = \mu_c(t) \). Then, for any \( u, z_1, z_2 \in \mathbb{R} \), the conditional cumulative distribution function of \( U \) and \( H_1Z \), given \( HZ = z_2 \), is

\[
F_{U,H_1Z|HZ}(u,z_1|z_2) = P(U \leq u, H_1Z \leq z_1|HZ = z_2) \\
= P(HZ + \epsilon_t \leq u, H_1Z \leq z_1|HZ = z_2) \\
= P(\epsilon_t \leq u - z_2, H_1Z \leq z_1|HZ = z_2) \\
= P(\epsilon_t \leq u - z_2|HZ = z_2)P(H_1Z \leq z_1|HZ = z_2) \\
= P(U \leq u|HZ = z_2)P(H_1Z \leq z_1|HZ = z_2) \\
= F_{U|HZ}(u|z_2)F_{H_1Z|HZ}(z_1|z_2)
\]

where the second equality follows from \( U = HZ + \epsilon_t \), the fourth equality comes from the fact that \( \epsilon_t \) is independent of \( H_1Z \) conditioned on \( HZ = \mu_c(t) \), and the second last step holds because that \( P(\epsilon_t \leq u - z_2|HZ = z_2) = P(\epsilon_t \leq u - HZ|HZ = z_2) = P(HZ + \epsilon_t \leq u|HZ = z_2) = P(U \leq u|HZ = z_2) \). Therefore, we conclude that \( U \) is independent of \( H_1Z \), given \( HZ \).

Next, we find the conditional distribution of \( U \) given \( Z \), namely \( f_{U|Z}(u|z) \), where \( z = (z_1, z_2)^T \) with \( z_1, z_2 \in \mathbb{R} \), and \( u \in \mathbb{R} \). To do this, we use the definition,

\[
f_{U|Z}(u|z) = \frac{f_{U,Z}(u,z)}{f_Z(z)} \tag{S.18}
\]

and first determine the joint distribution of \( U \) and \( Z \), \( f_{U,Z}(u,z) \). By the definition of \( Z \), we have

\[
f_{U,Z}(u,z) = f_{U,H_1Z,HZ}(u,z_1,z_2)
\]
which equals $f_{U, H, Z|HZ}(u, z_1|z_2)f_{HZ}(z_2)$. Since $U$ is conditionally independent of $H_1Z$, given $HZ$, it follows that

$$f_{U, Z}(u, z) = f_{U|HZ}(u|z_2)f_{H, Z|HZ}(z_1|z_2)f_{HZ}(z_2)$$

$$= f_{U|HZ}(u|z_2)f_{H, Z, HZ}(z_1, z_2)$$

$$= f_{U|HZ}(u|z_2)f_{Z}(z), \quad (S.19)$$

where the last equality is due to the definition of $Z = (H_1Z, HZ)^T$. Combining Equations (S.19) and (S.18) we obtain

$$f_{U|Z}(u|z) = f_{U|HZ}(u|z_2), \quad (S.20)$$

which is determined by Equation (7) in the primary article. Therefore, using our original notation for $U$ and $Z$, by Equation (S.20) the conditional density function of $Y(t)$ given $Z_t$ equals

$$f_{Y(t)|Z_t}(y_t|z_t) \propto \exp\left\{-\frac{1}{2\sigma_t^2}(y_t - Hz_t)^T(y_t - Hz_t)\right\}. \quad (S.21)$$

Finally, we determine the posterior distribution of $Z_t$, i.e., the conditional distribution of $Z_t$ given $Y(t)$. To this end, we assume that the prior distribution of $Z_t$ is represented or reasonably approximated by a Gaussian distribution with mean value equal to $\overline{z}_t^p$, and covariance matrix $\Sigma_t^p$. That is, the density function of the prior distribution of $Z_t$ equals

$$f_{Z_t}(z_t) \propto \exp\left\{-\frac{1}{2}(z_t - \overline{z}_t^p)^T(\Sigma_t^p)^{-1}(z_t - \overline{z}_t^p)\right\}. \quad (S.22)$$

By combining Equations (S.21) and (S.22) we obtain the density function of the posterior dis-
tribution of $Z_t$

$$
    f_{Z_t|Y(t)}(z_t|y_t) \propto f_{Z_t}(z_t) \times f_{Y(t)}(y_t|z_t) \\
    \propto \exp \left[ -\frac{1}{2} \left\{ z_t^T \left( \Sigma_t^p \right)^{-1} z_t - \bar{z}_t^p \left( \Sigma_t^p \right)^{-1} \bar{z}_t - \bar{z}_t^T \left( \Sigma_t^p \right)^{-1} \bar{z}_t^p \right. \\
    + z_t^T \frac{H_t^T}{\sigma_t^2} z_t - (H z_t)^T \frac{1}{\sigma_t^2} y_t - y_t \frac{1}{\sigma_t^2} H z_t \right] \\
    \propto \exp \left\{ -\frac{1}{2} \left[ z_t - \left\{ \left( \Sigma_t^p \right)^{-1} + \frac{H_t^T}{\sigma_t^2} \right\} \left\{ \left( \Sigma_t^p \right)^{-1} \bar{z}_t^p + \frac{H_t^T y_t}{\sigma_t^2} \right\} \right]^T \\
    \left[ \left( \Sigma_t^p \right)^{-1} + \frac{H_t^T H_t}{\sigma_t^2} \right] \left[ z_t - \left\{ \left( \Sigma_t^p \right)^{-1} + \frac{H_t^T}{\sigma_t^2} \right\} \left\{ \left( \Sigma_t^p \right)^{-1} \bar{z}_t^p + \frac{H_t^T y_t}{\sigma_t^2} \right\} \right], \right. \\
$$

showing that the posterior distribution of $Z_t$ is Gaussian with mean

$$
    \bar{z}_t^u = \left\{ \left( \Sigma_t^p \right)^{-1} + \frac{H_t^T H_t}{\sigma_t^2} \right\} \left\{ \left( \Sigma_t^p \right)^{-1} \bar{z}_t^p + \frac{H_t^T y_t}{\sigma_t^2} \right\} 
    \quad \text{(S.23)}
$$

and covariance matrix

$$
    \Sigma_t^u = \left\{ \left( \Sigma_t^p \right)^{-1} + \frac{H_t^T H_t}{\sigma_t^2} \right\}^{-1}. 
    \quad \text{(S.24)}
$$

**Appendix E: Derivations of the Expressions Found in Section 4.2**

In Appendix D, we demonstrated that the posterior distribution of $Z_t$ is Gaussian with mean value and covariance matrix as specified in Equations (S.23) and (S.24), respectively. In what follows, we show that these quantities are equal to the expressions specified in Section 4.2 of the primary article.

For $t \in \mathcal{T}$, let $\{\theta_{pri,t}^i : i = 1, \cdots, n\}$ and $\{\mu_{c,pri}^i(t) : i = 1, \cdots, n\}$, respectively, denote sequences of prior values of $\theta(t)$ and $\mu_c(t)$ at time $t$ that are described as follows. First, at iteration $l = 1$, the $\theta_{pri,t}^i$ and $\mu_{c,pri}^i(t)$ are generated either from Stage 1 of Part 1 or Stage 1 of Part 2 in Section 4.2. Second, at iteration $l > 1$ prior values $\theta_{pri,t}^i$ and $\mu_{c,pri}^i(t)$ are generated differently for $t = \tau_0$ or $t > \tau_0$. When $t = \tau_0$, they are determined by the corresponding components of
those values generated from the prior distribution \( \mathcal{N}(\hat{\eta}^{(l-1)}, a^{l-1}\Sigma) \), where \( \hat{\eta}^{(l-1)} \), \( a \), and \( \Sigma \) are identified in Section 4.2. When \( t > \tau_0 \), they are generated in a way similar to that outlined in Stage 1 of Part 2 in Section 4.2 in the main text.

Next, within each iteration and for \( t \in \mathcal{T} \), we calculate the sample means, sample covariance, and sample variances, respectively, using the expressions

\[
\bar{\theta}_{\text{pri},t} = \frac{\sum_{i=1}^{n} \theta^i_{\text{pri},t}}{n} ;
\]

\[
\bar{\sigma}_{\text{pri},t} = \frac{\sum_{i=1}^{n} \mu^i_{c,\text{pri}}(t)}{n} ;
\]

\[
\sigma_{\theta,\mu_c(t),\text{pri}}^{\text{cov}} = \frac{1}{n-1} \sum_{i=1}^{n} \left\{ \mu^i_{c,\text{pri}}(t) - \bar{\sigma}_{\text{pri},t} \right\} \left\{ \theta^i_{\text{pri},t} - \bar{\theta}_{\text{pri},t} \right\} ;
\]

\[
\sigma_{\theta,\text{pri},t}^{2} = \frac{\sum_{i=1}^{n} \left( \theta^i_{\text{pri},t} - \bar{\theta}_{\text{pri},t} \right)^2}{n-1} ;
\]

\[
\sigma_{\sigma^2,\text{pri},t}^{2} = \frac{\sum_{i=1}^{n} \left\{ \mu^i_{c,\text{pri}}(t) - \bar{\sigma}_{\text{pri},t} \right\}^2}{n-1} .
\]

Then we set

\[
\bar{z}_t^p = (\bar{\theta}_{\text{pri},t}, \bar{\sigma}_{\text{pri},t})^T
\]

and

\[
\Sigma_t^p = \begin{bmatrix}
\sigma_{\theta,\mu_c(t),\text{pri}}^{2} & \sigma_{\theta,\mu_c(t),\text{pri}}^{\text{cov}} \\
\sigma_{\theta,\mu_c(t),\text{pri}}^{\text{cov}} & \sigma_{\sigma^2,\text{pri},t}^{2}
\end{bmatrix},
\]

These values characterize the prior distribution for \( Z_t \) identified in Equation (S.22). Furthermore, we express the quantities specified in Equations (S.23) and (S.24) in terms of the observations and the prior values in Equations (S.25)-(S.29) as follows:

\[
\bar{z}_t^p = \bar{z}_t^p \begin{bmatrix}
\bar{\theta}_{\text{pri},t} + \frac{y_t \times \sigma_{\sigma^2,\mu_c(t),\text{pri}}^{\text{cov}} - \bar{\sigma}_{\text{pri},t} \times \sigma_{\sigma^2,\mu_c(t),\text{pri}}^{\text{cov}}}{\sigma_{\theta,\mu_c(t),\text{pri}}^{2} + \sigma_{\sigma^2,\text{pri},t}^{2}} \\
\frac{\sigma_{\sigma^2,\text{pri},t}^{2} + y_t \times \sigma_{\sigma^2,\text{pri},t}^{2}}{\sigma_{\sigma^2,\text{pri},t}^{2} + \sigma_{\sigma^2,\text{pri},t}^{2}}
\end{bmatrix} ;
\]
\[
\Sigma_t^u = \frac{\sigma_t^2}{\sigma_t^2 + \sigma_{\text{pri},t}^2} \times \left[ \frac{\sigma_{\theta,\text{pri},t}^2}{\sigma_{\theta,\theta,\text{pri},t}^{\text{cov}}} + \frac{\sigma_{\mu_c,\text{pri},t}^2}{\sigma_{\mu_c,\mu_c,\text{pri},t}^{\text{cov}}} \right].
\]  
(S.32)

To describe the generation of the posterior values of \(\theta(t)\) and \(\mu_c(t)\) at time \(t\) within each iteration, for \(i = 1, \ldots, n\), let \(z_{i,t}^p = (\theta_{\text{pri},t}^i, \mu_{c,\text{pri}}(t))^T\) denote prior values of \(\theta(t)\) and \(\mu_c(t)\), and let \(z_{i,t}^u = (\theta_{\text{post},t}^i, \mu_{c,\text{post}}(t))^T\) denote posterior values of \(\theta(t)\) and \(\mu_c(t)\) to be generated. Then the EAKF algorithm (Anderson, 2001, p.2887) calculates posterior values of \(Z_t\) as follows:

\[
z_{i,t}^u = D^T(z_{i,t}^p - \overline{z}_t^p) + \overline{z}_t^u \quad \text{for } i = 1, \ldots, n,
\]  
(S.33)

where \(\overline{z}_t^p\) and \(\overline{z}_t^u\) are given by Equations (S.30) and (S.31), respectively, and \(D\) can be any matrix such that the covariance matrix of \(\{z_{i,t}^u : i = 1, \ldots, n\}\) computed from Equation (S.33) equals the value computed via Equation (S.32). The existence of such a matrix \(D\) is justified in Appendix A of Anderson (2001). For instance, \(D\) can be

\[
\begin{bmatrix}
\frac{\sigma_{\theta,\theta,\text{pri},t}^{\text{cov}}}{\sigma_{\theta,\theta,\text{pri},t}^2} & \frac{\sigma_{\mu_c,\mu_c,\text{pri},t}^{\text{cov}}}{\sigma_{\mu_c,\mu_c,\text{pri},t}^2} \\
\frac{\sigma_{\theta,\mu_c,\text{pri},t}^2}{\sigma_{\theta,\theta,\text{pri},t}^2 + \sigma_{\mu_c,\mu_c,\text{pri},t}^2} & \frac{\sigma_{\mu_c,\mu_c,\text{pri},t}^2}{\sigma_{\theta,\theta,\text{pri},t}^2 + \sigma_{\mu_c,\mu_c,\text{pri},t}^2}
\end{bmatrix} \times \left( \sqrt{\frac{\sigma_t^2}{\sigma_t^2 + \sigma_{\text{pri},t}^2}} - 1 \right) \sqrt{\frac{\sigma_{\theta,\theta,\text{pri},t}^2}{\sigma_{\theta,\theta,\text{pri},t}^2 + \sigma_{\mu_c,\mu_c,\text{pri},t}^2}}
\]  
(S.34)

as we show in Appendix F below.

Finally, substituting this value for \(D\) identified in Equation (S.34) for \(D\) in Equation (S.33), we obtain the following explicit expressions for posterior values of \(\theta(t)\) and \(\mu_c(t)\) at time \(t\) for \(i = 1, \ldots, n\), namely

\[
m_{u_{c,\text{post}}^i}(t) = \frac{\sigma_t^2}{\sigma_t^2 + \sigma_{\text{pri},t}^2} \overline{\sigma}_{\text{pri},t} + \frac{\sigma_{\text{pri},t}^2}{\sigma_t^2 + \sigma_{\text{pri},t}^2} y_t
\]
\[
+ \sqrt{\frac{\sigma_t^2}{\sigma_t^2 + \sigma_{\text{pri},t}^2} \left\{ \mu_{c,\text{pri}}^i(t) - \overline{\sigma}_{\text{pri},t} \right\}^2;}
\]

\[
\theta_{\text{post},t}^i = \theta_{\text{pri},t}^i + \frac{\sigma_{\theta,\mu_c,\text{pri},t}^{\text{cov}}}{\sigma_{\theta,\theta,\text{pri},t}^2} \left\{ \mu_{c,\text{post}}^i(t) - \mu_{c,\text{pri}}^i(t) \right\}.
\]  
(S.35)
Setting \( t = \tau_0 \), Equations (S.35) and (S.36) equal the expressions specified in Equations (9) and (10), respectively, in the primary article.

**Appendix F: Verification of the Validity of Matrix \( D \) in Appendix E**

Substituting the matrix \( D \) identified in Equation (S.34) in Equation (S.33) gives us

\[
\begin{bmatrix}
\theta_{\text{post},t}^i \\
\mu_{\text{c,post}(t)}^i
\end{bmatrix} =
\begin{bmatrix}
\frac{\sigma_{\theta,\mu_c(t),\text{pri}}^2}{\sigma_{\text{pri},t}^2} \\
0
\end{bmatrix}
\times \left( \frac{\sqrt{\frac{\sigma_t^2}{\sigma_t^2 + \sigma_{\text{pri},t}^2}} - 1}{\frac{\sigma_t^2}{\sigma_t^2 + \sigma_{\text{pri},t}^2}} \right)
\times \begin{bmatrix}
\theta_{\text{pri},t}^i - \bar{\theta}_{\text{pri},t} \\
\mu_{\text{c,\text{pri}(t)}}^i - \bar{\mu}_{\text{c,\text{pri}(t)}}^i
\end{bmatrix}
+ \begin{bmatrix}
\bar{\theta}_{\text{pri},t} \\
\bar{\theta}_{\text{pri},t} + \frac{y_t \times \sigma_{\text{cov}}^2(\theta_{\mu_c(t)},\text{pri} - \sigma_{\theta,\mu_c(t),\text{pri}})}{\sigma_t^2 + \sigma_{\text{pri},t}^2}
\end{bmatrix},
\]

so that

\[
\theta_{\text{post},t}^i = \theta_{\text{pri},t}^i - \bar{\theta}_{\text{pri},t} + \{ \mu_{\text{c,\text{pri}(t)}}^i - \bar{\mu}_{\text{c,\text{pri}(t)}}^i \} \times \frac{\sigma_{\theta,\mu_c(t),\text{pri}}^2}{\sigma_{\text{pri},t}^2} \times \left( \frac{\sqrt{\frac{\sigma_t^2}{\sigma_t^2 + \sigma_{\text{pri},t}^2}} - 1}{\frac{\sigma_t^2}{\sigma_t^2 + \sigma_{\text{pri},t}^2}} \right)
+ \bar{\theta}_{\text{pri},t} + \frac{y_t \times \sigma_{\text{cov}}^2(\theta_{\mu_c(t)},\text{pri} - \sigma_{\theta,\mu_c(t),\text{pri}})}{\sigma_t^2 + \sigma_{\text{pri},t}^2}
\]  
(S.37)

and

\[
\mu_{\text{c,post}(t)}^i = \{ \mu_{\text{c,\text{pri}(t)}}^i - \bar{\mu}_{\text{c,\text{pri}(t)}}^i \} \times \sqrt{\frac{\sigma_t^2}{\sigma_t^2 + \sigma_{\text{pri},t}^2}}
+ \bar{\theta}_{\text{pri},t} \times \sigma_t^2 + \sigma_{\text{pri},t}^2
\]  
(S.38)

Thus, Equations (S.37) and (S.38) provide a sequence of posterior values of \( \theta \), \( \{ \theta_{\text{post},t}^i : i = 1, \cdots, n \} \), and a sequence of posterior values of \( \mu_c(t) \), \( \{ \mu_{\text{c,post}(t)}^i : i = 1, \cdots, n \} \).

Next, we calculate the variance of the posterior values \( \{ \theta_{\text{post},t}^i : i = 1, \cdots, n \} \) using Equation (S.37), namely

\[
\sigma_{\theta,\text{post},t}^2 = \frac{\sum_{i=1}^{n} (\theta_{\text{post},t}^i - \frac{\theta_{\text{post},t}}{n})^2}{n - 1}
= \frac{\sigma_t^2}{\sigma_t^2 + \sigma_{\text{pri},t}^2} \times \left( \frac{\sigma_{\theta,\text{pri},t}^2 + \sigma_{\text{pri},t}^2 - \sigma_{\theta,\text{cov},\mu_c(t),\text{pri}}^2}{\sigma_t^2} \right),
\]

13
which is identical to the $(1, 1)$ element of $\Sigma^u_t$ found in Equation (S.32).

Using Equation (S.38), we calculate the variance of the posterior values $\{\mu^i_{c,\text{post}}(t) : i = 1, \cdots, n\}$ of $\mu_c(t)$, as

$$\sigma^2_{\text{post},t} = \sum_{i=1}^{n} \left\{ \frac{\mu^i_{c,\text{post}}(t) - \mu^i_{c,\text{post}}(t)}{n} \right\}^2$$

$$= \frac{1}{n-1} \sum_{i=1}^{n} \left\{ \mu^i_{c,\text{pr}},t - \bar{\theta}_{\text{pr},t} \right\}^2 \times \frac{\sigma^2_t}{\sigma^2_t + \sigma^2_{\text{pr},t}}$$

$$= \frac{\sigma^2_t \times \sigma^2_{\text{pr},t}}{\sigma^2_t + \sigma^2_{\text{pr},t}},$$

which is identical to the $(2, 2)$ element of $\Sigma^u_t$ identified in Equation (S.32).

Furthermore, using Equations (S.37) and (S.38), we can calculate the covariance of $\{\theta^i_{\text{post},t} : i = 1, \cdots, n\}$ and $\{\mu^i_{c,\text{post}}(t) : i = 1, \cdots, n\}$, as

$$\sigma^2_{\text{cov},\theta,\mu_c(t),\text{post}} = \sum_{i=1}^{n} \left\{ \theta^i_{\text{post},t} - \frac{\sum_{i=1}^{n} \theta^i_{\text{post},t}}{n} \right\} \left\{ \mu^i_{c,\text{post}}(t) - \frac{\sum_{i=1}^{n} \mu^i_{c,\text{post}}(t)}{n} \right\}$$

$$= \frac{1}{n-1} \sum_{i=1}^{n} \left( \theta^i_{\text{pr},t} - \bar{\theta}_{\text{pr},t} \right) \times \left\{ \mu^i_{c,\text{pr}}(t) - \bar{\theta}_{\text{pr},t} \right\} \times \sqrt{\frac{\sigma^2_t}{\sigma^2_t + \sigma^2_{\text{pr},t}}}$$

$$+ \left\{ \mu^i_{c,\text{pr}}(t) - \bar{\theta}_{\text{pr},t} \right\}^2 \times \frac{\sigma^2_{\text{cov},\theta,\mu_c(t),\text{pr}}}{\sigma^2_{\text{pr},t}} \times \left( \frac{\sigma^2_t}{\sigma^2_t + \sigma^2_{\text{pr},t}} - \frac{\sqrt{\sigma^2_t}}{\sqrt{\sigma^2_t + \sigma^2_{\text{pr},t}}} \right)$$

$$= \frac{\sigma^2_t}{\sigma^2_t + \sigma^2_{\text{pr},t}} \times \sigma^2_{\text{cov},\theta,\mu_c(t),\text{pr}},$$

which is identical to two off-diagonal elements of $\Sigma^u_t$ found in Equation (S.32). Therefore, the matrix $D$ satisfies the condition required by the EAKF algorithm.
Section S3: Additional Analysis for the Quebec COVID-19 Data

Appendix G: Additional Figures

Figure S.2: Quebec data from April 2, 2020 to April 30, 2020: the reported daily net number (left), weekly net number (median), and daily cumulative number (right) of confirmed cases.
Appendix H: Sensitivity Analysis of the Quebec COVID-19 Data

To further assess the performance of the proposed SEASAR model, we conducted twelve sensitivity analyses of the Quebec data that we analyzed in Section 5 to evaluate the sensitivity of our analysis results to our specification of the initial values. We used the same analysis framework as we reported in Section 5.1 except that in each of the 12 following settings we changed one element. Thus, in the first three settings, we investigated different ways of specifying the OEV $\sigma_t^2$ which is identified in Equation (7). Instead of using Equation (8), we set $\sigma_t^2$ equal to one of the following three forms: (S1): $\sigma_t^2 = max\left(1, \frac{\sqrt{y_{t-1}}}{20}\right)$; (S2): $\sigma_t^2 = max(c, y_{t-1})$; (S3): $\sigma_t^2 = max(c, y_{t-1}, y_{t-2})$, where $c$ equalled the average of the $y_s$ for $s \in T$, and $y_{T_0-2}$ and $y_{T_0-1}$ are the number of confirmed cases on March 31, 2020 and April 1, 2020, respectively.

Next, considered different values for the latent period by setting $Z$ to be one of the following...
three values that were estimated by different authors: (S4): $Z = 4.0$ days (Guan et al., 2020); (S5): $Z = 5.08$ days (He et al., 2020); (S6): $Z = 6.4$ days (Backer et al., 2020). Finally, we adopted one of the following settings involving different values of $r_1$, $r_2$ and $r_3$: (S7): $r_3 = 1$; (S8): $r_3 = 3$; (S9): $r_2 = 2$; (S10): $r_2 = 0.5$; (S11): $r_1 = 0.1$; (S12): $r_1 = 3$, to assess the effect of these model parameters.

The results of these sensitivity analyses are reported in 12 separate panels of Figure S.4 which display the results of model fitting (in green) and prediction (in blue) for the daily net number of cases, as opposed to the reported daily net number of cases (in red). These plots clearly indicate that the performance of the SEASAR model is sensitive to the specification of $\sigma^2_t$, $Z$ and $\{r_1, r_2, r_3\}$, and the disparity of the fitted values from the reported values varies from setting to setting. To see how these different estimation results may indicate the severity of the pandemic, in Table S.2 we report the estimates of the basic reproduction number $R_0$ that were obtained from the twelve sensitivity analyses, where Analysis $k$ represents the results associated with setting $(S_k)$ for $k = 1, \cdots, 12$.

Finally, we comment that the applicability of the SEASAR model hinges on the appropriateness of the two conditions: (1) the population is homogeneous; and (2) the population is closed to immigration and emigration. On April 1, 2020, the Quebec provincial government set up checkpoints to block all non-essential travel into the province and advised all Quebec residents to stay home. After May 10, 2020, Quebec and other provinces in Canada gradually relaxed these restrictions on migration. While the prohibition on immigration and emigration is difficult to satisfy in modelling real world data, it is reasonable to suppose that immigration and emigration during the period from April 2, 2020 to May 10, 2020 in Quebec was at a nadir. On the other hand, we need to be aware that the condition concerning a homogeneous population is almost never true in practical applications; consequently, care should be taken when interpreting the analysis results. Despite these issues, the analyses reported here demonstrate the utility of
Figure S.4: The model fitting to the daily net number of cases (in green) in the period of April 2, 2020 to April 30, 2020, and the model prediction to the daily net number of cases (in blue) in the period of May 1, 2020 to May 10, 2020, as opposed to the reported daily net number of cases (in red) in the period of April 2, 2020 to May 10, 2020.

Table S.2: Estimates of $R_0$ from twelve sensitivity analyses of the Quebec data under the SEASAR model

| Analysis | 1       | 2       | 3       | 4       | 5       | 6       | 7       | 8       | 9       | 10      | 11      | 12      |
|----------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| Estimate of $R_0$ | $3.20 \times 10^{-3}$ | 0.75    | 0.66    | 0.94    | 1.06    | 0.94    | 1.06    | 1.05    | 1.06    | 0.97    | 0.61    | 0.48    |

The SEASAR model. They may also shed some light on the underlying trajectory of COVID-19 in Quebec during the study period.

**Appendix I: Implementing the IF-EAKF Algorithm for the SIR Model**

In the SIR model, $\mu_{SIR}^c(t) := \lambda S_{SIR}(t)I_{SIR}(t)/N$ equals the number of infected cases that occur on day $t$. Since the SIR model does not involve the confirmation process, $\mu_{SIR}^c(t)$ represents the number of confirmed cases that occur on day $t$. Like our treatment of the SEASAR model, we assume the distribution for $Y_t$ specified in Equation (7) of the primary article, where $\mu_{c}(t)$ is replaced by $\mu_{SIR}^c(t)$, and $\sigma_t^2$ is equal to the expression found in Equation (8).
Let $\phi_{\text{SIR}}(t) = (S_{\text{SIR}}(t), I_{\text{SIR}}(t), R_{\text{SIR}}(t))^T$, $\eta_{\text{SIR}} = (\lambda, \xi)^T$ and $g_{\text{SIR}}(\cdot, \cdot)$ represent the right hand side of Equations (13)-(15) in the article, which can be written in the compact form
\[
\frac{d\phi_{\text{SIR}}(t)}{dt} = g(\phi_{\text{SIR}}(t), \eta_{\text{SIR}}).
\]

Following our implementation of IF-EAFK for the SEASAR model, we add the argument $t$ to $\eta_{\text{SIR}}$, denoted $\eta_{\text{SIR}}(t) = (\lambda(t), \xi(t))^T$.

Similarly, let the initial size $\phi_{\text{SIR}}(\tau_0)$ equal $I_{\text{SIR}}(\tau_0) = I_a(\tau_0) + I_s(\tau_0), R_{\text{SIR}}(\tau_0) = R(\tau_0)$, and $S_{\text{SIR}}(\tau_0) = N - I_{\text{SIR}}(\tau_0) - R_{\text{SIR}}(\tau_0)$. The prior distribution for $\eta_{\text{SIR}}(\tau_0)$, denoted $\pi_{\eta_{\text{SIR}}}$, is assumed to be independent and uniform for $\lambda$ and $\xi$ with the following ranges: $10^{-6} \leq \lambda \leq 7; 10^{-6} \leq \xi \leq 0.8$.

We now repeat the algorithm outlined in Section 4.2 with the following modifications to accommodate the SIR model. Specifically, for $t \in \mathcal{T}$, we replace $\eta(t), \mu_c(t), \phi(t)$, and $g(\cdot, \cdot)$ for the SEASAR model with $\eta_{\text{SIR}}(t), \mu_{c_{\text{SIR}}}(t), \phi_{\text{SIR}}(t)$ and $g_{\text{SIR}}(\cdot, \cdot)$, respectively. In Step 1 of Stage 1 of Part 1, $\pi_{\eta}$ is replaced by the prior $\pi_{\eta_{\text{SIR}}}$ specified above. In Step 3 of Stage 1 of Part 1, $\mu_{c_{\text{pri}}}^i(\tau_0)$ is now re-defined as $\lambda_{\text{pri,}r_0}^i S_{\text{SIR}}(\tau_0) I_{\text{SIR}}(\tau_0)/N$; the symbol $X$ represents either $\lambda$ or $\xi$. In Step 2 of Stage 2 of Part 1, $\eta_{\text{post,}\tau_0}^i$ is replaced by $(\lambda_{\text{post,}\tau_0}^i, \xi_{\text{post,}\tau_0}^i)^T$ and $X$ is either $\lambda$ or $\xi$. We repeat Appendices D, E and F, replacing $\theta(t)$ by $\lambda(t)$ or $\xi(t)$ and replacing $\mu_c(t)$ by $\mu_{c_{\text{SIR}}}(t)$ to derive the corresponding expressions in Stage 2 of Part 1. In Step 1 of Stage 1 of Part 2, $\eta_{\text{pri,}\tau_0+1}^i = (\lambda_{\text{pri,}\tau_0+1}^i, \xi_{\text{pri,}\tau_0+1}^i)^T$. In Step 2 of Stage 1 of Part 2, $\phi_{\text{SIR,}\tau_0+1}^i = (S_{\text{SIR,}\tau_0+1}^i, I_{\text{SIR,}\tau_0+1}^i, R_{\text{SIR,}\tau_0+1}^i)^T$. In Step 3 of Stage 1 of Part 2, we set $\mu_{c_{\text{SIR,}\tau_0+1}}^i = \lambda_{c_{\text{SIR,}\tau_0+1}}^i S_{\text{SIR,}\tau_0+1}^i I_{\text{SIR,}\tau_0+1}^i R_{\text{SIR,}\tau_0+1}^i/N$, and $X$ is a symbol for each of $S_{\text{SIR}}, I_{\text{SIR}}$ or $R_{\text{SIR}}$. In Step 2 of Stage 2 of Part 2, $\phi_{\text{post,}\tau_0+1}^i = (S_{\text{SIR,}\tau_0+1}^i, I_{\text{SIR,}\tau_0+1}^i, R_{\text{SIR,}\tau_0+1}^i)^T$. 

19
and $X$ is a symbol for each of $S_{SIR}$, $I_{SIR}$ or $R_{SIR}$.

**Appendix J: Implementing the IF-EAKF Algorithm for the SEIR Model**

In the SEIR model, $\mu^{SEIR}_{c}(t) := E_{SEIR}(t)/Z$ equals the number of infected cases that occur on day $t$, with $Z = 5.2$. Since the SEIR model does not incorporate the confirmation process, $\mu^{SEIR}_{c}(t)$ is also regarded as the number of confirmed cases that occur on day $t$. Once more we assume the distribution for $Y_t$ specified in Equation (7) in the primary article, where $\mu_c(t)$ is replaced by $\mu^{SEIR}_{c}(t)$, and $\sigma^2_t$ equals the expression found in Equation (8).

Let $\phi^{SEIR}(t) = (S_{SEIR}(t), E_{SEIR}(t), I_{SEIR}(t), R_{SEIR}(t))^T$, $\eta^{SEIR} = (\lambda^*, \xi^*)^T$ and $g^{SEIR}(\cdot, \cdot)$ represent the right hand side of Equations (16)-(19) in the primary article. These equations can be written in the compact form

$$\frac{d\phi^{SEIR}(t)}{dt} = g(\phi^{SEIR}(t), \eta^{SEIR}).$$

Following our implementation of IF-EAFK for the SEASAR model, we add argument $t$ to $\eta^{SEIR}$, denoted $\eta^{SEIR}(t) = (\lambda^*(t), \xi^*(t))^T$.

Similarly, the initial size $\phi^{SEIR}()$ equals $E_{SEIR}(\tau_0) = E(\tau_0), I_{SEIR}(\tau_0) = I_a(\tau_0)+I_s(\tau_0), R_{SEIR}(\tau_0) = R(\tau_0)$, and $S_{SEIR}(\tau_0) = N - E_{SEIR}(\tau_0) - I_{SEIR}(\tau_0) - R_{SEIR}(\tau_0)$. The prior distribution for $\eta^{SEIR}(\tau_0)$, denoted $\pi_{\eta^{SEIR}}$, is assumed to be independent and uniform for $\lambda^*$ and $\xi^*$ with the following ranges: $10^{-6} \leq \lambda^* \leq 7; 10^{-6} \leq \xi^* \leq 0.8$.

We now repeat the algorithm outlined in Section 4.2 with the following modifications to accommodate the SEIR model. Specifically, for $t \in \mathcal{T}$, we replace $\eta(t)$, $\mu_c(t)$, $\phi(t)$, and $g(\cdot, \cdot)$ for the SEASAR model with $\eta^{SEIR}(t)$, $\mu^{SEIR}_{c}(t)$, $\phi^{SEIR}(t)$ and $g^{SEIR}(\cdot, \cdot)$, respectively. In Step 1 of Stage 1 of Part 1, $\pi_\eta$ is replaced by the prior $\pi_{\eta^{SEIR}}$ specified above. In Step 3 of Stage 1 of Part 1, $\mu^{\lambda, \xi}_{c, \text{pri}}(\tau_0)$ is now re-defined as $E_{SEIR}(\tau_0)/Z$ with $Z = 5.2$; the symbol $X$ represents either $\lambda^*$ or $\xi^*$. In Step 2 of Stage 2 of Part 1, $\eta^{i}_{\text{post}, \tau_0}$ is replaced with $(\lambda^{i \text{post}, \tau_0}, \xi^{i \text{post}, \tau_0})^T$ and
$X$ is either $\lambda^*$ or $\xi^*$. We repeat Appendices D, E, and F, replacing $\theta(t)$ by $\lambda^*(t)$ or $\xi^*(t)$ and replacing $\mu_c(t)$ by $\mu_c^{SEIR}(t)$ to derive the corresponding expressions in Stage 2 of Part 1. In Step 1 of Stage 1 of Part 2, $\eta_{pri,\tau_0+1}^{i} = (\lambda_{pri,\tau_0+1}^{i}, \xi_{pri,\tau_0+1}^{i})^T$. In Step 2 of Stage 1 of Part 2, $\phi_{pri}^{i}(\tau_0 + 1) = (S_{SEIR_{pri}}^{i}(\tau_0 + 1), E_{SEIR_{pri}}^{i}(\tau_0 + 1), I_{SEIR_{pri}}^{i}(\tau_0 + 1), R_{SEIR_{pri}}^{i}(\tau_0 + 1))^T$. In Step 3 of Stage 1 of Part 2, we set $\mu_{c_{pri}}^{i}(\tau_0 + 1) = E_{SEIR_{pri}}^{i}(\tau_0 + 1)/Z$ with $Z = 5.2$, and $X$ is a symbol for each of $S_{SEIR}$, $E_{SEIR}$, $I_{SEIR}$ or $R_{SEIR}$. In Step 2 of Stage 2 of Part 2, $\phi_{post}^{i}(\tau_0 + 1) = (S_{SEIR_{post}}^{i}(\tau_0 + 1), E_{SEIR_{post}}^{i}(\tau_0 + 1), I_{SEIR_{post}}^{i}(\tau_0 + 1), R_{SEIR_{post}}^{i}(\tau_0 + 1))^T$ and $X$ is a symbol for each of $S_{SEIR}$, $E_{SEIR}$, $I_{SEIR}$ or $R_{SEIR}$.

Appendix K: Simulation Studies with the Assumptions Violated

To test the robustness of our method against the two model assumptions, we considered three situations involving violations of the model assumptions. In Situation 1, only the first assumption is violated, whereas in Situation 2, only the second assumption is not met. In Situation 3 we allowed both model assumptions to be both untrue. In this appendix, we set $T = 49$ and $m = 1000$.

Data Generation:

In Situation 1, we considered a population that is not homogeneous. One way of generating such a population involves dividing the population considered in the framework for Section 6 into two approximately equal disjoint subpopulations. Thus, we set the sizes of both subpopulations to be 4216651 and 4216650, respectively. Then in each of those subpopulations, we separately used Equations (1)-(6) to generate the mean sizes $\phi(t)$ of the six compartments from $t = \tau_0$ to $t = \tau_0 + T$. To differentiate them, we used $\phi_1(t) = (S_1(t), E_1(t), I_{a1}(t), I_{s1}(t), A_1(t), R_1(t))^T$ and $\phi_2(t) = (S_2(t), E_2(t), I_{a2}(t), I_{s2}(t), A_2(t), R_2(t))^T$ to denote the average sizes of the six compartments.
compartment states on day \( t \) for the first and second subpopulations, respectively. Further, we let \( \eta_1 \) and \( \eta_2 \) denote the model parameters of Equations (1)-(6) for the first and second subpopulations, respectively.

To generate data for each subpopulation using a process similar to that identified in Equation (12) in the primary article, we set the initial values for \( \phi_1(t) \) and \( \phi_2(t) \) to be half of the initial values \( \phi(\tau_0) \) reported in Section 6; that is, \( \phi_1(\tau_0) = \phi_2(\tau_0) = (4210574.647, 1737.536, 1009.158, 1009.158, 2274.5, 45.5)^T \). We fixed \( Z = 5.2 \) as in Section 6, but we specified different model parameters for the first and second subpopulations to allow for the possibility that the COVID-19 transmission may be different in the two subpopulations, thus, yielding a non-homogeneous mixed population. Specifically, for the first subpopulation, we used the same values of the model parameters as in Section 6, i.e., \( \eta_1 = (0.2, 0.45, 0.18, 0.08, 0.7, 9, 24, 34)^T \), for the second subpopulation, we set \( \eta_2 = (0.25, 0.5, 0.3, 0.4, 0.6, 4.5, 27, 30)^T \).

Next, we used Equation (12) separately for the two subpopulations and then sequentially generated \( \phi_1(t) \) and \( \phi_2(t) \) from \( t = \tau_0 \) to \( t = \tau_0 + T \). Similar to our discussion in Section 3.1, we set \( \mu_{c,1}(t) = \frac{I_s_1(t)}{5} \) and \( \mu_{c,2}(t) = \frac{I_s_2(t)}{4.5} \) to be the mean numbers of confirmed cases on day \( t \) for the first and second subpopulations, respectively. Let \( \mu_c(t) = \mu_{c,1}(t) + \mu_{c,2}(t) \). Then we generated \( m \) realizations, denoted \( D_{c,1}^j = \{y^j_{1,t} : t \in \mathcal{T}\} \) for \( j = 1, 2, \ldots, m \), of time series \( \{Y(t) : t = \tau_0, \tau_0 + 1, \ldots, \tau_0 + T\} \) using Equation (7), where we specify the OEV identified in Equation (8) with \( y_{t-1} \) replaced by \( y_{1,t-1}^j \) and \( y_{1,\tau_0-1}^j = 449 \).

Situation 2 involved violating the assumption that immigration and emigration may not occur. One way of generating such data is to allow the population size to vary the population size vary with time. To this end, we used the same data generation framework that we previously reported in Section 6 except that the population size \( N \) was now specified as \( N(t) \) for \( t \in \mathcal{T} \), where \( N(t) = 8433301 + 1000 \times (t - \tau_0) \) for \( t \in \mathcal{T} \). Then we used a version of \( \phi(t + 1) \) similar to that specified in Equation (12) to generate \( \phi(t) \) together with \( N(t) \) for \( t \in \mathcal{T} \). As in Section
6, we generated realized values for the process \( \{ Y(t) : t \in \mathcal{T} \} \) using Equations (7) and (8) sequentially, where \( y_{\tau_0} = 449 \). We repeated this procedure \( m \) times, and let \( \mathcal{D}_{c,2}^j \) denote the \( j \)th sample of realized values for the time series \( \{ Y(t) : t \in \mathcal{T} \} \), where \( j = 1, \ldots, m \).

Finally, to generate data representing Situation 3, we combined our notions of data generation used in Situations 1 and 2. Specifically, we used two subpopulations, where the size of the first subpopulation equalled \( 4216651 + 500 \times (t - \tau_0) \) and that of the second subpopulation was \( 4216650 + 500 \times (t - \tau_0) \) for \( t \in \mathcal{T} \). Then, using Equations (12), (7) and (8) as we described above, we repeatedly generated realizations of \( \{ Y(t) : t \in \mathcal{T} \} \) \( m \) times, and let \( \mathcal{D}_{c,3}^j \) denote the \( j \)th sample of realizations \( \{ Y(t) : t \in \mathcal{T} \} \).

**Assessment of the Performance:**

Next we used the IF-EAKF algorithm with \( n = 300, a = 0.9 \) and \( L = 50 \) to fit the generated data for each of the three situations outlined above. As we previously reported in Section 6.2, we assessed the performance of the SEASAR model with estimation or prediction results, where Situation 2 yielded values for both estimation of the model parameter \( \eta \) and the prediction error. However, Situations 1 and 3 yielded only the results for prediction.

In Table S.3 we report the average estimate bias, average relative estimate bias and sample standard deviation of \( m \) estimates of the model parameter \( \eta \) in Situation 2 for three cases. For Scenario-1-short, Scenario-1-long, Scenario-2-short, and Scenario-2-long, Tables S.4 and S.5 report the average and the sample standard deviation of \( m \) TAPEs and TRAPEs for predicting of the daily net or cumulative number of confirmed cases. Figure 5.4 displays the boxplots of those \( m \) TAPEs and TRAPEs. It is clear that the TRAPEs in Situation 2 are more variable than the corresponding values observed in Situations 1 and 3. These results indicate that the proposed SEASAR model seems reasonably robust to the violation of the homogeneous popu-
lation assumption, but it is sensitive to the violation of the assumption that neither immigration nor emigration occurs in the population.

Table S.3: Simulation study with $n = 300$, $a = 0.9$ and $L = 50$: Estimation results for the SEASAR model parameters in Situation 2; the entries with * are the original values times $10^3$. BIAS and RBIAS represent the average bias of the estimates and the average relative bias of the estimates, respectively; and SSD stands for the sample standard deviation of the estimates.

| Parameter | Case 1 | Case 2 | Case 3 |
|-----------|--------|--------|--------|
|           | BIAS   | RBIAS  | SSD    | BIAS   | RBIAS  | SSD    | BIAS   | RBIAS  | SSD    |
| $\theta$  | -0.01  | -0.06  | 0.02   | -0.01  | -0.07  | 0.02   | -0.02  | -0.08  | 0.02   |
| $\mu$     | 0.17   | 0.38   | 0.09   | 0.20   | 0.45   | 0.09   | 0.22   | 0.50   | 0.09   |
| $\alpha$  | -0.01  | -0.08  | 0.02   | -0.02  | -0.10  | 0.02   | -0.02  | -0.11  | 0.02   |
| $\beta$   | 1.84*  | 0.02   | 0.02   | 3.21   | 0.04   | 0.02   | 1.08*  | 0.01   | 0.02   |
| $\gamma$  | -0.06  | -0.09  | 0.09   | -0.08  | -0.11  | 0.11   | -0.09  | -0.13  | 0.10   |
| $F$       | -7.94* | -0.88* | 0.16   | -0.01  | -1.29* | 0.16   | -0.02  | -1.83* | 0.16   |
| $B$       | 0.55   | 0.02   | 2.17   | 0.25   | 0.01   | 2.14   | 0.43   | 0.02   | 2.28   |
| $J$       | 1.03   | 0.03   | 2.49   | 1.10   | 0.03   | 2.80   | 0.96   | 0.03   | 2.66   |
Table S.4: Simulation Study: Prediction performance under Setting 1 for Scenario-1-short ($T_1 = 29; T_2 = 5$), Scenario-1-long ($T_1 = 29; T_2 = 10$), Scenario-2-short ($T_1 = 39; T_2 = 5$), and Scenario-2-long ($T_1 = 39; T_2 = 10$). Here $T_1$ and $T_2$ represent the size of the training and test data respectively. ATAPE and ATRAPE represent the averages of the total absolute prediction error and the total relative absolute prediction error, respectively.

| Situation 1 | ATAPE | SSD | ATRAPE | SSD |
|-------------|-------|-----|--------|-----|
| Scenario-1-short | 24.56 | 8.97 | 0.17 | 0.06 |
| Scenario-1-long | 49.47 | 14.16 | 0.37 | 0.11 |
| Scenario-2-short | 22.69 | 7.87 | 0.21 | 0.07 |
| Scenario-2-long | 45.94 | 11.39 | 0.47 | 0.12 |

| Situation 2 | ATAPE | SSD | ATRAPE | SSD |
|-------------|-------|-----|--------|-----|
| Scenario-1-short | 24.60 | 8.92 | 0.31 | 0.11 |
| Scenario-1-long | 49.36 | 13.99 | 0.68 | 0.20 |
| Scenario-2-short | 22.81 | 7.84 | 0.41 | 0.15 |
| Scenario-2-long | 46.07 | 11.30 | 0.91 | 0.23 |

| Situation 3 | ATAPE | SSD | ATRAPE | SSD |
|-------------|-------|-----|--------|-----|
| Scenario-1-short | 24.56 | 8.92 | 0.17 | 0.06 |
| Scenario-1-long | 49.33 | 14.11 | 0.37 | 0.11 |
| Scenario-2-short | 22.70 | 7.85 | 0.22 | 0.08 |
| Scenario-2-long | 45.96 | 11.40 | 0.48 | 0.12 |
Table S.5: Simulation Study: Prediction performance under Setting 2 for Scenario-1-short ($T_1 = 29; T_2 = 5$), Scenario-1-long ($T_1 = 29; T_2 = 10$), Scenario-2-short ($T_1 = 39; T_2 = 5$), and Scenario-2-long ($T_1 = 39; T_2 = 10$). Here $T_1$ and $T_2$ represent the size of the training and test data respectively. ATAPE and ATRAPE represent the averages of the total absolute prediction error and the total relative absolute prediction error, respectively. The entries with * are the original values times $10^3$.

| Situation   | TAPE | TRAPE |
|-------------|------|-------|
|             | ATAPE | SSD   | ATRAPE | SSD   |
| Scenario-1-short | 104.04 | 77.68 | 0.01   | 9.45$^*$ |
| Scenario-1-long  | 264.72 | 193.43 | 0.03   | 0.02   |
| Scenario-2-short | 103.66 | 74.43 | 0.01   | 7.89$^*$ |
| Scenario-2-long  | 245.44 | 172.44 | 0.03   | 0.02   |

| Situation 2  | TAPE | TRAPE |
|--------------|------|-------|
|              | ATAPE | SSD   | ATRAPE | SSD   |
| Scenario-1-short | 103.87 | 78.10 | 0.02   | 0.02   |
| Scenario-1-long  | 264.37 | 194.11 | 0.05   | 0.04   |
| Scenario-2-short | 109.30 | 80.56 | 0.02   | 0.01   |
| Scenario-2-long  | 257.94 | 185.05 | 0.04   | 0.03   |

| Situation 3  | TAPE | TRAPE |
|--------------|------|-------|
|              | ATAPE | SSD   | ATRAPE | SSD   |
| Scenario-1-short | 102.50 | 76.45 | 0.01   | 9.31$^*$ |
| Scenario-1-long  | 261.20 | 190.63 | 0.03   | 0.02   |
| Scenario-2-short | 102.18 | 75.01 | 0.01   | 7.96$^*$ |
| Scenario-2-long  | 242.60 | 172.37 | 0.03   | 0.02   |
Figure S.5: The boxplots of $M$ TAPEs (left) and TRAPEs (right) for Scenario-1-short ($T_1 = 29; T_2 = 5$), Scenario-1-long ($T_1 = 29; T_2 = 10$), Scenario-2-short ($T_1 = 39; T_2 = 5$), and Scenario-2-long ($T_1 = 39; T_2 = 10$) in three situations under the two settings. Here $T_1$ and $T_2$ represent the size of the training and test data, respectively.
Section S4: Additional Results for Simulation Studies

Figure S.6: The curves of \( S(t) \), \( E(t) \), \( I_a(t) \), \( I_s(t) \), \( A(t) \), and \( R(t) \).

Figure S.7: The boxplots of MTAPES (left) and TRAPES (right) for Scenario-1-short (\( T_1 = 29; T_2 = 5 \)), Scenario-1-long (\( T_1 = 29; T_2 = 10 \)), Scenario-2-short (\( T_1 = 39; T_2 = 5 \)), and Scenario-2-long (\( T_1 = 39; T_2 = 10 \)) under the two settings. Here \( T_1 \) and \( T_2 \) represent the size of the training and test data, respectively.
Table S.6: Simulation study in Section 6: Estimate results of the SEASAR model parameters; the entries with * are the original values times $10^3$. BIAS and RBIAS represent the average bias of the estimates and the average relative bias of the estimates, respectively; and SSD stands for the sample standard deviation of the estimates.

| Parameter | Case 1 | Case 2 | Case 3 |
|-----------|--------|--------|--------|
| $\theta$  |        |        |        |
| $\mu$     | 0.13   | 0.29   | 0.13   |
| $\alpha$  | -0.01  | -0.06  | 0.03   |
| $\beta$   | 1.48*  | 0.92   | 0.02   |
| $\gamma$  | -0.05  | -0.07  | 0.12   |
| $\bar{E}$ | -7.40* | -0.82* | 0.17   |
| $B$       | 0.32   | 0.01   | 4.70   |
| $J$       | 0.92   | 0.03   | 5.54   |

| Parameter | Case 1 | Case 2 | Case 3 |
|-----------|--------|--------|--------|
| $\theta$  |        |        |        |
| $\mu$     | 0.15   | 0.34   | 0.11   |
| $\alpha$  | -0.01  | -0.08  | 0.02   |
| $\beta$   | 1.59   | 0.02   | 0.02   |
| $\gamma$  | -0.06  | -0.08  | 0.10   |
| $\bar{E}$ | -8.47* | -0.94* | 0.16   |
| $B$       | 0.33   | 0.01   | 3.66   |
| $J$       | 1.00   | 0.03   | 4.21   |

| Parameter | Case 1 | Case 2 | Case 3 |
|-----------|--------|--------|--------|
| $\theta$  |        |        |        |
| $\mu$     | 0.17   | 0.38   | 0.09   |
| $\alpha$  | -0.02  | -0.08  | 0.02   |
| $\beta$   | 1.92*  | 0.02   | 0.02   |
| $\gamma$  | -0.06  | -0.09  | 0.10   |
| $\bar{E}$ | -8.06* | -0.90* | 0.16   |
| $B$       | 0.46   | 0.02   | 2.13   |
| $J$       | 1.06   | 0.03   | 2.52   |

| Parameter | Case 1 | Case 2 | Case 3 |
|-----------|--------|--------|--------|
| $\theta$  |        |        |        |
| $\mu$     | 0.18   | 0.39   | 0.08   |
| $\alpha$  | -0.02  | -0.09  | 0.02   |
| $\beta$   | 1.90*  | 0.02   | 0.02   |
| $\gamma$  | -0.07  | -0.09  | 0.09   |
| $\bar{E}$ | -8.83* | -0.98* | 0.16   |
| $B$       | 0.44   | 0.02   | 1.68   |
| $J$       | 0.94   | 0.03   | 2.04   |

| Parameter | Case 1 | Case 2 | Case 3 |
|-----------|--------|--------|--------|
| $\theta$  |        |        |        |
| $\mu$     | 0.18   | 0.40   | 0.08   |
| $\alpha$  | -0.02  | -0.09  | 0.02   |
| $\beta$   | 1.51*  | 0.02   | 0.02   |
| $\gamma$  | -0.07  | -0.10  | 0.09   |
| $\bar{E}$ | -9.43* | -1.05* | 0.16   |
| $B$       | 0.48   | 0.02   | 1.21   |
| $J$       | 0.96   | 0.03   | 1.41   |
References

Anderson, J. L. (2001). An ensemble adjustment kalman filter for data assimilation. *Monthly Weather Review, 129*(12), 2884–2903.

Backer, J. A., Klinkenberg, D., & Wallinga, J. (2020). Incubation period of 2019 novel coronavirus (2019-nCov) infections among travellers from Wuhan, China, 20-28 January 2020. *Eurosurveillance, 25*(5), 1–6.

Bretó, C., He, D., Ionides, E. L., King, A. A., et al. (2009). Time series analysis via mechanistic models. *The Annals of Applied Statistics, 3*(1), 319–348.

Delamater, P. L., Street, E. J., Leslie, T. F., Yang, Y. T., & Jacobsen, K. H. (2019). Complexity of the basic reproduction number ($R_0$). *Emerging Infectious Diseases, 25*(1), 1–4.

Diekmann, O., Heesterbeek, J., & Roberts, M. G. (2010). The construction of next-generation matrices for compartmental epidemic models. *Journal of the Royal Society Interface, 7*(47), 873–885.

Diekmann, O., Heesterbeek, J. A. P., & Metz, J. A. (1990). On the definition and the computation of the basic reproduction ratio $R_0$ in models for infectious diseases in heterogeneous populations. *Journal of Mathematical Biology*, 28(4), 365–382.

Guan, W., Ni, Z., Hu, Y., Liang, W., Ou, C., He, J., Liu, L., Shan, H., Lei, C., Hui, D. S., et al. (2020). Clinical characteristics of coronavirus disease 2019 in China. *New England Journal of Medicine*, 382(18), 1708–1720.

He, D., Ionides, E. L., & King, A. A. (2010). Plug-and-play inference for disease dynamics: measles in large and small populations as a case study. *Journal of the Royal Society Interface, 7*(43), 271–283.
He, W., Yi, G. Y., & Zhu, Y. (2020). Estimation of the basic reproduction number, average incubation time, asymptomatic infection rate, and case fatality rate for COVID-19: Meta-analysis and sensitivity analysis. *Journal of Medical Virology, 92*(11), 2543–2550.

Ionides, E. L., Bretó, C., & King, A. A. (2006). Inference for nonlinear dynamical systems. *Proceedings of the National Academy of Sciences, 103*(49), 18438–18443.

Süli, E. & Mayers, D. F. (2003). *An Introduction to Numerical Analysis*. Cambridge University Press.