

Supporting Information

S2 Appendix. Approximate Bayesian analysis. We apply Bayesian inference to quantify uncertainty in the model parameters associated vaccine hesitancy effects, $\theta = [\nu, n_n, w_C, w_D, w_V]$, for each synthetic dataset i that represents the structure of data sets available from Johns Hopkins University or Our World in Data dashboards, that is, $\mathcal{D}_i = [\{C_{t,i}^*, D_{t,i}^*, V_{1,t,i}^*, V_{2,t,i}^*\}_{T \geq t \geq 0}]$ where T is the total number of days available of data. The task is to sample the posterior distribution with probability density given by Bayes' Theorem,

$$p(\theta | \mathcal{D}_i) = \frac{p(\mathcal{D}_i | \theta)p(\theta)}{p(\mathcal{D}_i)},$$

where $p(\theta)$ is the prior, $p(\mathcal{D}_i | \theta)$ is the likelihood and $p(\mathcal{D}_i)$ is the evidence. For the remainder of this section, we omit the country index i for notational convenience.

Since the full model state vector is only partially observable, the data model is no longer Markovian and is therefore computationally intractable. To deal with this likelihood intractability, we apply approximate Bayesian computation (ABC) [1–3], that samples from an approximation to the posterior for each country,

$$\begin{aligned} p(\theta | \mathcal{D}) &\approx p(\theta | \rho(\mathcal{D}, \mathcal{D}_s) \leq \epsilon) \propto \mathbb{P}(\rho(\mathcal{D}, \mathcal{D}_s) \leq \epsilon | \theta)p(\theta) \\ &= p(\theta) \int \mathbb{1}_{(0, \epsilon]}(\rho(\mathcal{D}, \mathcal{D}_s)) s(\mathcal{D}_s | \theta) d\mathcal{D}_s, \end{aligned}$$

where \mathcal{D} is the synthetic data that will align to one of the hesitancy scenarios scenario, $\mathcal{D}_s \sim s(\cdot | \theta)$ is simulated data from the full model, $\rho(\mathcal{D}, \mathcal{D}_s)$ is a discrepancy metric, ϵ is the discrepancy threshold and $\mathbb{1}_{(0, \epsilon]}(\rho(\mathcal{D}, \mathcal{D}_s)) = 1$ if $\rho(\mathcal{D}, \mathcal{D}_s) \leq \epsilon$, and $\mathbb{1}_{(0, \epsilon]}(\rho(\mathcal{D}, \mathcal{D}_s)) = 0$ otherwise. For our implementation, we apply the discrepancy metric,

$$\rho(\mathcal{D}, \mathcal{D}_s) = \left(\sum_{t=1}^T (C_t^* - C_{t,s}^*)^2 + (D_t^* - D_{t,s}^*)^2 + (V_{1,t}^* - V_{1,t,s}^*)^2 + (V_{2,t}^* - V_{2,t,s}^*)^2 \right)^{1/2}$$

where $\mathcal{D} = [\{A_t^*, D_t^*, V_{1,t}^*, V_{2,t}^*\}_{t \geq 0}]$ is the synthetic data for a given scenario and $\mathcal{D}_s = [\{A_{t,s}^*, D_{t,s}^*, V_{1,t,s}^*, V_{2,t,s}^*\}_{t \geq 0}]$ is simulated data form the full model.

Sequential Monte Carlo sampling

We apply a sequential Monte Carlo (SMC) scheme [?, ?] to move an initial set of N_p samples from the prior through a sequence of ABC approximations defined by a decreasing sequence of T discrepancy thresholds, $\epsilon_1 > \epsilon_2 > \dots > \epsilon_T = \epsilon$. Our particular implementation (Algorithm 1), based on the work of Drovandi and Pettit [?], adaptively selects the acceptance thresholds and utilises MCMC steps using tuned Gaussian random walk proposals. For all model calibrations we apply adaptive SMC with $N_p = 1000$ particles, tuning parameters $c = 0.01$, $a = 0.5$ and terminate sampling when the MCMC acceptance probability p_{acc} drops below $p_{\text{min}} = 0.01$. The acceptance probability is estimated using $R_{\text{trial}} = 5$ initial MCMC iterations per particle.

Algorithm 1 Adaptive SMC sampler for approximate Bayesian computation

```
1: Initialise  $N_a = aN_p$ ,  $N_\ell = N_p - N_a$ 
2: for  $j \in [1, 2, \dots, N_p]$  do
3:   Sample prior,  $\theta^* \sim p(\cdot)$  and simulate model,  $\mathcal{D}_s \sim s(\cdot \mid \theta^*)$ ;
4:   Set  $\rho_j \leftarrow \rho(\mathcal{D}, \mathcal{D}_s)$ ;
5: end for
6: repeat
7:   Sort particles  $\{(\theta_j, \rho_j)\}_{j=1}^{N_p}$ , such that  $\rho_j \leq \rho_{j+1}$  for all  $j \in [1, 2, \dots, N_p - 1]$ ;
8:   Remove particles  $\{(\theta_j, \rho_j)\}_{j=N_\ell+1}^{N_p}$  and set  $\epsilon \leftarrow \rho_{N_\ell}$ ;
9:   Resample particles  $\{\theta_j\}_{j=1}^{N_p}$  from  $\{(\theta_j)\}_{j=1}^{N_\ell}$  with replacement;
10:  Estimate sample covariance,  $\hat{\Sigma}$ , of particles  $\{\theta_j\}_{j=1}^{N_p}$ .
11:  Adapt proposal kernel  $q(\mathbf{u} \mid \mathbf{v}) = \phi\left(\mathbf{u}; \mathbf{v}, \frac{2.38^2}{\dim(\theta)} \hat{\Sigma}\right)$ , where  $\phi(\cdot; \boldsymbol{\mu}, \boldsymbol{\Sigma})$  is a
    multivariate Gaussian density function and  $\dim(\theta)$  is the number of parameters;
12:  Set  $p_{\text{acc}} \leftarrow 0$ ;
13:  for  $j \in [N_\ell + 1, N_\ell + 2, \dots, N_p]$  do
14:    for  $k \in [1, 2, \dots, R_{\text{trial}}]$  do
15:      Generate proposal,  $\theta^* \sim q(\cdot \mid \theta_j)$  and sample  $u \sim \mathcal{U}(0, 1)$ ;
16:      if  $u \leq \min\left(1, \frac{p(\theta^*)q(\theta_j \mid \theta^*)}{p(\theta_j)q(\theta^* \mid \theta_j)}\right)$  then
17:        Simulate model  $\mathcal{D}_s \sim s(\cdot \mid \theta^*)$ ;
18:        if  $\rho(\mathcal{D}_i, \mathcal{D}_s) \leq \epsilon$  then
19:          Set  $\theta_j \leftarrow \theta^*$ ,  $\rho_j \leftarrow \rho(\mathcal{D}_i, \mathcal{D}_s)$ , and  $p_{\text{acc}} \leftarrow p_{\text{acc}} + (R_{\text{trial}}N_a)^{-1}$ ;
20:        end if
21:      end if
22:    end for
23:  end for
24:  Set  $R \leftarrow \log c / \log(1 - p_{\text{acc}})$ ;
25:  for  $j \in [N_\ell + 1, N_\ell + 2, \dots, N_p]$  do
26:    for  $k \in [1, 2, \dots, R - R_{\text{trial}}]$  do
27:      Generate proposal,  $\theta^* \sim q(\cdot \mid \theta_j)$  and sample  $u \sim \mathcal{U}(0, 1)$ ;
28:      if  $u \leq \min\left(1, \frac{p(\theta^*)q(\theta_j \mid \theta^*)}{p(\theta_j)q(\theta^* \mid \theta_j)}\right)$  then
29:        Simulate model  $\mathcal{D}_s \sim s(\cdot \mid \theta^*)$ ;
30:        if  $\rho(\mathcal{D}_i, \mathcal{D}_s) \leq \epsilon$  then
31:          Set  $\theta_j \leftarrow \theta^*$ ,  $\rho_j \leftarrow \rho(\mathcal{D}_i, \mathcal{D}_s)$ , and  $p_{\text{acc}} \leftarrow p_{\text{acc}} + (RN_a)^{-1}$ ;
32:        end if
33:      end if
34:    end for
35:  end for
36: until  $p_{\text{acc}} < p_{\text{min}}$ 
```

References

1. Sisson SA, Fan Y, Beaumont MA, editors. Handbook of Approximate Bayesian Computation. Chapman and Hall/CRC; 2018.
2. Sunnåker M, Busetto AG, Numminen E, Corander J, Foll M, Dessimoz C. Approximate Bayesian Computation. PLoS Computational Biology. 2013;9(1):e1002803. doi:10.1371/journal.pcbi.1002803.
3. Warne DJ, Baker RE, Simpson MJ. Simulation and inference algorithms for stochastic biochemical reaction networks: from basic concepts to state-of-the-art. Journal of The Royal Society Interface. 2019;16(151):20180943. doi:10.1098/rsif.2018.0943.