

# Estimating the impact of school closures on COVID-19 disease burden

## Supplementary Appendix

————— Add Authors List here —————

### Contents

<b>1</b>	<b>Model description</b>	<b>2</b>
1.1	General approach . . . . .	2
1.2	Transmission model . . . . .	2
1.2.1	Compartment types and sequence . . . . .	2
1.2.2	Model stratification by age . . . . .	3
1.2.3	Capturing the effects of vaccination . . . . .	4
1.2.4	Modelling multiple viral strains . . . . .	4
1.2.5	Dynamic social mixing . . . . .	5
1.2.6	Random transmission adjustment . . . . .	6
1.2.7	Ordinary differential equations . . . . .	7
1.2.8	Model parameters . . . . .	7
1.3	Estimation of COVID-19-related hospital pressure and deaths . . . . .	7
1.3.1	COVID-19-related hospital pressure . . . . .	8
1.3.2	COVID-19 deaths . . . . .	10
<b>2</b>	<b>Software and code used to conduct the analyses</b>	<b>10</b>
2.1	Code . . . . .	10
2.2	Software implementation . . . . .	10
2.2.1	Application Programming Interface (API) . . . . .	10
2.2.2	Optimising compiler . . . . .	10
<b>3</b>	<b>Model calibration and uncertainty propagation</b>	<b>10</b>
3.1	Parameters varied during calibration . . . . .	11
3.2	Calibration targets . . . . .	11
3.3	Likelihood definition . . . . .	13

# 1 Model description

## 1.1 General approach

We use a semi-mechanistic compartmental model of COVID-19 transmission governed by ordinary differential equations (ODEs). Our model captures important factors of COVID-19 transmission and disease such as age-specific characteristics, heterogeneous mixing, vaccination and the emergence of different variants of concern. The ODE-based model is used to capture only states relevant to transmission, whereas hospitalisations and deaths are estimated through a convolution process applied to the ODE-based model's outputs. This process combines the model-estimated disease incidence with statistical distributions modelling the time to hospitalisation, the hospital stay duration and the time to death. This approach presents two main advantages. First it reduces the complexity of the dynamic system relying on numerical solving of ODEs, which is computationally expensive. Second, the convolution approach allows for more flexibility and produces more realistic assumptions regarding the timings of hospitalisations and deaths, compared to what could be achieved with a simple compartmental approach. The following sections describe the model in details and Figure S1 summarises the overall approach used in our analysis.

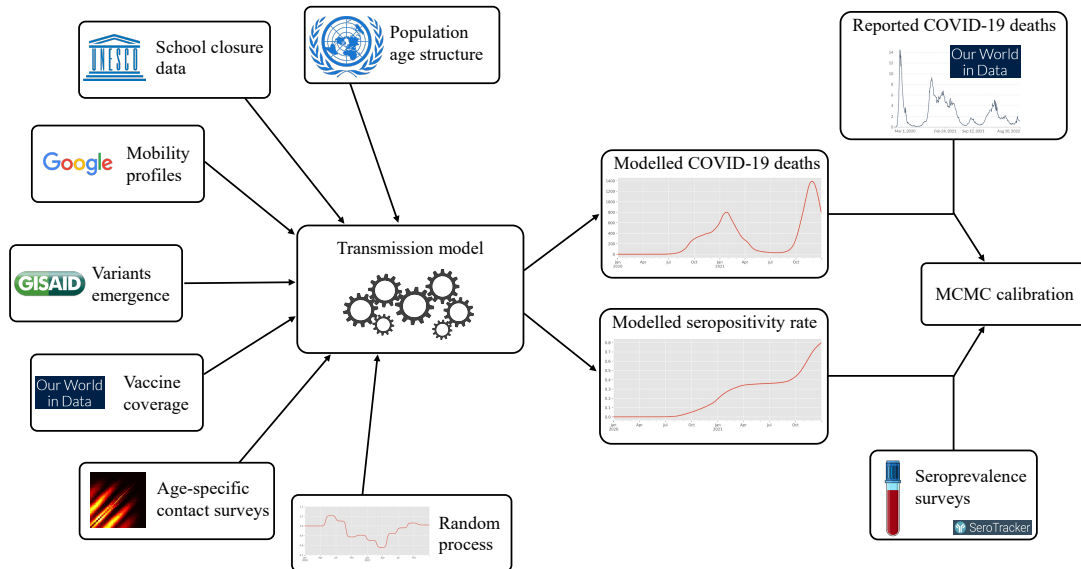


Figure S1: Conceptual approach used to build and calibrate the models

## 1.2 Transmission model

### 1.2.1 Compartment types and sequence

Model compartments represent sequential progressions through the processes of infection with, progression through, and recovery from the phases of SARS-CoV-2 infection and COVID-19 disease. The following types of compartments are implemented:

- Susceptible
  - Persons not previously infected with SARS-CoV-2 during or before the model simulation period
- Latent
  - Persons recently infected with SARS-CoV-2, but not in the active phase of the disease yet.
  - These individuals may still be infectious (see details in next paragraph).
- Active

- Persons with active COVID-19 who are currently infectious.
- Recovered
  - Persons recovered from COVID-19 during the model simulation period
  - Reinfection from these compartments is permitted through exposure to a different strain than the one that most recently infected the individual (see strain stratification section for details).

The base model structure consists of a sequence of one susceptible compartment ( $S$ ), four latent compartments ( $E_1, \dots, E_4$ ), four active disease compartments ( $I^1, \dots, I^4$ ) and one recovered compartment ( $R$ ) (Figure S2).

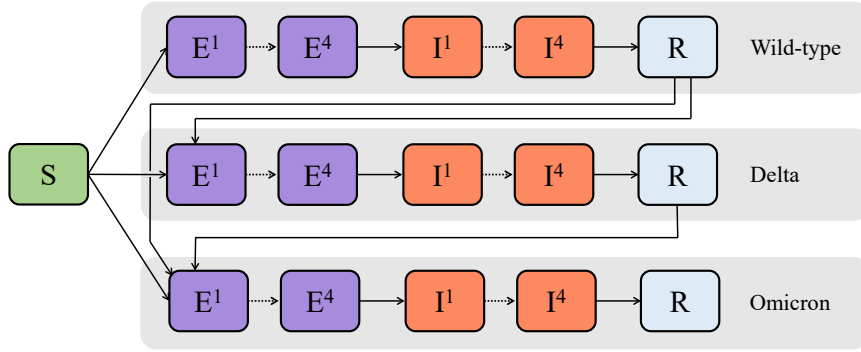


Figure S2: Compartmental model structure.  $S$  = Susceptible,  $E$  = Exposed / Latent,  $I$  = Active disease,  $R$  = Recovered. Stratification by age and vaccination status are not shown here.

The main rationale for using four serial compartments for both the latent and active states is to achieve an Erlang distribution for the time spent in each of these states. This distribution is more realistic than the exponential distribution which is the consequence of a single compartment assumption, because the Erlang distribution does not have a large density mass around 0 and is not heavy-tailed. Figure S3 illustrates the modelled distributions of the incubation period and the active disease period.

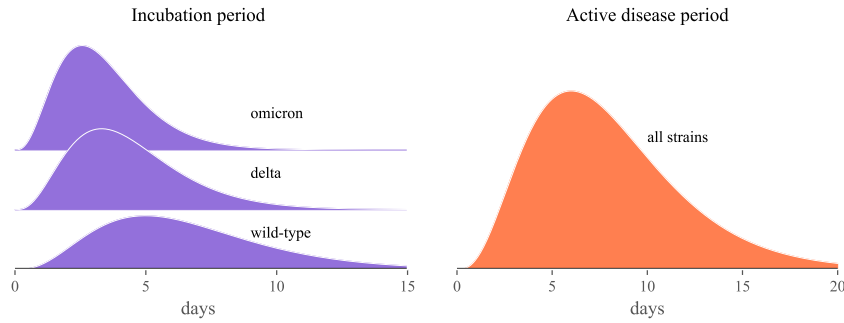


Figure S3: Modelled distributions of the incubation and active disease periods. Active disease period shown for an average duration of 8 days, but this parameter is varied during calibration.

The four active disease compartments all have identical characteristics. However, the last two latent compartments ( $E^3$  and  $E^4$ ) are infectious whereas the first two ( $E^1$  and  $E^2$ ) are not. We further assume that the infectious latent compartments are half as infectious as the active disease compartments.

### 1.2.2 Model stratification by age

All compartments of the base compartmental structure were stratified by age into the following age bands: zero to 14 years / 15 to 24 years / 25 to 49 years / 50 to 69 years / 70 years and above.

The initial population was distributed between the different age bands to reflect the age distributions reported by the United Nations Population Division. Demographic processes, including births, ageing and non-infection-related deaths are not simulated, given the timeframes considered in this simulation.

We assumed heterogeneous mixing between the different age groups to account for the assortative nature of social interactions by age (see Section 1.2.5). Age was assumed to affect:

- The susceptibility to infection
- The risk of COVID-19 hospitalisation
- The risk of COVID-19 death
- The vaccination rate (see Section 1.2.3)

### 1.2.3 Capturing the effects of vaccination

History of vaccination was captured by stratifying all model compartments by vaccination status. Two vaccination strata were included to represent those who have received at least two doses of a COVID-19 vaccine, and those who have not.

We used data from *Our World in Data* to inform the modelled dynamic vaccination coverage. In particular, we used the reported proportion of people “fully vaccinated” to specify the time-variant proportion of vaccinated people in our model. We assumed that older individuals are vaccinated first by prioritising the modelled age groups in descending order. That is, the oldest age group receives all available vaccines until a saturation coverage of 80% is reached for this group. Then the next oldest category starts receiving vaccines and we repeat this process until all available vaccines are allocated. Note that in the event that the population-level vaccine coverage exceeds 80%, the saturation coverage was set equal to the population-level coverage.

Let us consider two successive time points  $t_i$  and  $t_{i+1}$  for which vaccination data are available. Let us denote  $r_{a,i}$  and  $r_{a,i+1}$  the associated vaccine coverage for age group  $a$ . The time-variant and age-specific vaccination rate per capita  $w_a(t)$  verifies:

$$1 - r_{a,i+1} = (1 - r_{a,i})e^{-w_a(t)(t_{i+1}-t_i)} \quad , \forall t \in [t_i, t_{i+1}). \quad (1)$$

Then,

$$w_a(t) = \frac{\ln(1 - r_{a,i}) - \ln(1 - r_{a,i+1})}{t_{i+1} - t_i} \quad , \forall t \in [t_i, t_{i+1}), \quad (2)$$

where  $\ln(x)$  represents the natural logarithm of  $x$ .

Figure S4 shows the modelled vaccination coverage over time against the reported data for the analysed countries.

The effect of vaccination on transmission is to reduce the rate of infection partially for all persons at-risk of infection in the vaccinated stratum. This includes both fully susceptible (never previously infected) persons, as well as recovered persons who are at risk of reinfection. The model allows for hybrid immunity in the sense that the vaccination-induced relative reduction of transmission risk is multiplied with that induced by previous infection. Vaccination is also assumed to reduce the risk of hospitalisation and death.

Emerging variants of concern (VoCs) may partially escape vaccine-induced (as well as infection-induced) immunity, as described further below (Table S2 and Section 1.2.4).

### 1.2.4 Modelling multiple viral strains

The model was stratified by “strain” to simulate the emergence of multiple variants of concern (VoC). This approach explicitly represents multiple competing strains, each with an independent force of infection calculation. We assumed that VoCs can have different levels of transmissibility, incubation period and disease severity (hospitalisation and death risks) compared to the ancestral COVID-19 strain. In addition, VoCs were assumed to escape immunity partially for both vaccination- and infection-related immunity.

We assumed that individuals previously infected with the wild-type strain could only be reinfected with the delta or omicron strains. However, such individuals have a reduced risk of infection with these variants compared to infection-naïve individuals (68% and 45% reduction for delta and omicron,

respectively) [REF DOI: 10.1503/cmaj.211248]. We assumed that individuals previously infected with the delta variant could only be reinfected with the omicron variant, with an infection risk reduced by 45% compared to infection-naïve individuals [REF DOI: 10.1503/cmaj.211248]. The other parameters used to represent strain-specific characteristics are presented in Table S2.

Seeding of each new strain into the model was achieved through the importation of a small number (10 per million population) of new infectious persons with the relevant strain into the model. The seeding process was done over a ten-day period and the start of this period was set to the emergence date reported by GISAID.

### 1.2.5 Dynamic social mixing

The model captures changes in social interactions over time through a dynamic age-specific mixing matrix. The following sections describe how this matrix was defined and how it captures the different non-pharmaceutical interventions implemented in the analysed countries, including school closures. The overall approach is also illustrated by Figure S6.

#### 1.2.5.1 Reference mixing matrices

We extracted country-specific contact matrices using the *conmat* R package which derives social mixing matrices from contact survey data. These matrices provide the average numbers of contacts per day between different age groups, disaggregated by the following locations: home, school, work, other locations.

The overall contact matrix (before adjustments for mobility changes) results from the summation of the four location-specific contact matrices:  $C_0 = C_H + m_S C_S + C_W + C_L$ , where  $C_H$ ,  $C_S$ ,  $C_W$  and  $C_L$  are the age-specific contact matrices associated with households, schools, workplaces and other locations, respectively. Note that the school contribution  $C_S$  is multiplied by the factor  $m_S$  that was varied during model calibration (see Section 3), in order to account for uncertainty around the relative contribution of school contacts to overall mixing.

#### 1.2.5.2 Modifications of contact rates over time

To capture mobility changes over time, the contributions of the matrices  $C_S$ ,  $C_W$  and  $C_L$  vary with time such that the input contact matrix can be written as:

$$C(t) = h(t)^2 C_H + s(t)^2 m_S C_S + w(t)^2 C_W + l(t)^2 C_L \quad (3)$$

The modifying functions  $h$  (for households),  $s$  (for schools),  $w$  (for work) and  $l$  (for other-locations) were each squared to capture the effect of the mobility changes on both the infector and the infectee in any given interaction that could potentially result in transmission.

#### School closure/re-opening

Reduced attendance at schools was represented through the function  $s$ , which represents the proportion of all school students currently attending on-site teaching. If schools are fully closed at time  $t$ ,  $s(t) = 0$  and  $C_S$  does not contribute to the overall mixing matrix  $C(t)$ . The function  $s$  was derived from the UNESCO database on school closures from the start of the COVID-19 pandemic (ADD REF HERE). This database provides school opening status over time as a categorical variable taking the following values: “Fully open”, “Partially open”, “Academic break”, “Closed due to COVID-19”. Table S1 indicates how the different categorical values were converted into the numerical function  $s$ .

Table S1: Assumed percentage of students on-site for the different UNESCO school closure categories.

UNESCO category	Assumed proportion of students on-site ( $s(t)$ )
Fully open	100%
Partially open	10-30%
Academic break	0%
Closed due to COVID-19	0%

We included uncertainty around the value associated with the partial closure category, as there was no quantitative data available to inform this parameter. The partial closure periods are likely to be periods where only a small fraction of students such as children of “essential workers” were attending school. We assumed that between 10 and 30% of students attended on-site learning during these periods.

To model the counterfactual “no school closure” scenario, we assumed that the schools were “Fully open” during the periods reported as “Partially open” or “Closed due to COVID-19”.

Figure S5 summarises the UNESCO data on school closure for the analysed countries.

### Dynamic mobility outside of schools and homes

Changes to people’s mobility in places other than schools and homes were modelled using Google Mobility data, after applying a seven-day moving average smoothing. We used the “Workplace” category of the Google data to scale the work-related matrix contribution  $C_W$  to overall mixing over time, using the adjusting function  $w$ . The “other locations” matrix  $C_L$  was scaled through the adjusting function  $l$  which was defined as the average of the Google mobility indicators across the following Google categories: “Retail and recreation”, “Grocery and pharmacy” and “Transit stations”.

### Household contacts

In the base case analysis, the contribution of household contacts to the overall mixing matrix was fixed over time (i.e.  $h(t) = 1$  in Equation 3). Although Google provides mobility estimates for residential contacts, the nature of these data is different from that of each of the other Google mobility types. They represent the time spent in that location, as opposed to other categories, which measure a change in total visits rather than the duration. The daily frequency with which people attend their residence is likely to be close to one, and we considered that household members likely have a daily opportunity for infection with each other household member regardless of the background level of mobility. Therefore, we did not implement a function to scale the contribution of household contacts to the mixing matrix with time.

#### 1.2.5.3 Sensitivity analyses around dynamic social mixing

In addition to the Base Case analysis described above, we performed two sensitivity analyses considering different assumptions for the modelled social mixing.

**SA1: No Google mobility data** In a first sensitivity analysis (SA1), we removed the contribution of the Google mobility data to the modelled social mixing (see Figure S6). In this configuration, the calibrated random process ( $W(t)$ ) was used to capture mobility changes implicitly.

**SA2: School closures increase household contact rates** In another sensitivity analysis (SA2), we considered an alternative assumption under which the effective contact rates within households were increased during periods of school closure. In that case, the household component of the mixing matrix is modified by the following function:

$$h(t) = 1 + 0.20(1 - s(t)) \quad ,$$

where  $s$  is the function modifying school contacts as introduced in Section 1.2.5.2. This is equivalent to assuming that each individual has 20% more household contact potential when schools are fully closed.

#### 1.2.6 Random transmission adjustment

The risk of SARS-CoV-2 transmission per contact was adjusted by a time-variant random process, making the model semi-mechanistic. This random process reflects the fact that all the variations observed in the transmission risk in the real world cannot be explained solely by the factors that are explicitly captured through model inputs such as vaccination, dynamic mobility or new variants’ emergence. We therefore allowed for random perturbations to the risk of transmission over time, although these perturbations were highly auto-correlated to avoid unrealistic changes over a short period of time.

We used a random walk with Gaussian update defined by:

$$\begin{aligned} W(0) &= 0 \\ W(t+1) &\sim \mathcal{N}(W(t), 0.5) \quad , \end{aligned} \tag{4}$$

where  $\mathcal{N}$  denotes the normal distribution. The random process  $W$  was updated every two months and was transformed using the exponential function before being applied to the risk of transmission per

contact (see Equation 6). Finally, the contribution of the random process to the risk of transmission was squared in order to capture its effect on both the susceptible and the infectious individuals (see Equation 6).

### 1.2.7 Ordinary differential equations

Let us first introduce some new notations. Modelled age groups are indicated by the subscript  $a$ , and  $\mathcal{A}$  represents the set of all modelled age groups (see Section 1.2.2). Vaccination status is represented by the subscript  $v$ , and  $\mathcal{V}$  is the set of vaccination statuses (i.e.  $\mathcal{V} = \{“0”, “1”\}$ , where “0” represents unvaccinated people and “1” represents vaccinated people). The subscript  $s$  is used to represent the different viral strains, and  $\mathcal{S}$  is the set of all strains (i.e.  $\mathcal{S} = \{“wild-type”, “delta”, “omicron”\}$ ). The average incubation period duration associated with strain  $s$  is denoted  $q_s$  and the average duration of active disease is denoted  $w$ . The relative susceptibility to infection with strain  $s$  of individuals aged  $a$  with vaccination status  $v$  is denoted  $\rho_{a,v,s}$ . The term  $b_{a,v,s}(t)$  designates the introduction of individuals of age  $a$  and with vaccination status  $v$  that are infected with strain  $s$  (infection seeding). Vaccination is characterised by the age-specific and time-variant per-capita vaccination rate  $\omega_a$ . Finally,  $\chi_{s,\sigma}$  represents the relative susceptibility to infection with strain  $\sigma$  for individuals whose most recent infection episode was with strain  $s$ . Using this new notation combined with those previously introduced, we can describe the model with the following set of ordinary differential equations:

$$\begin{aligned}
\frac{dS_{a,v}}{dt} &= - \sum_{s \in \mathcal{S}} \lambda_{a,s}(t) \rho_{a,v,s} S_{a,v} + \Phi_v \omega_a(t) S_{a,v=0} \quad , \\
\frac{dE_{a,v,s}^1}{dt} &= \lambda_{a,s}(t) \rho_{a,v,s} \left( S_{a,v} + \sum_{\sigma \in \mathcal{S}} \chi_{\sigma,s} R_{a,v,\sigma} \right) - \frac{4}{q_s} E_{a,v,s}^1 + \Phi_v \omega_a(t) E_{a,v=0,s}^1 \quad , \\
\frac{dE_{a,v,s}^k}{dt} &= \frac{4}{q_s} E_{a,v,s}^{k-1} - \frac{4}{q_s} E_{a,v,s}^k + \Phi_v \omega_a(t) E_{a,v=0,s}^k \quad , \forall k \in \{2, 3, 4\}, \\
\frac{dI_{a,v,s}^1}{dt} &= \frac{4}{q_s} E_{a,v,s}^4 - \frac{4}{w} I_{a,v,s}^1 + b_{a,v,s}(t) + \Phi_v \omega_a(t) I_{a,v=0,s}^1 \quad , \\
\frac{dI_{a,v,s}^k}{dt} &= \frac{4}{w} I_{a,v,s}^{k-1} - \frac{4}{w} I_{a,v,s}^k + \Phi_v \omega_a(t) I_{a,v=0,s}^k \quad , \forall k \in \{2, 3, 4\}, \\
\frac{dR_{a,v,s}}{dt} &= \frac{4}{w} I_{a,v,s}^4 - \sum_{\sigma \in \mathcal{S}} \lambda_{a,\sigma}(t) \rho_{a,v,\sigma} \chi_{s,\sigma} R_{a,v,s} + \Phi_v \omega_a(t) R_{a,v=0,s} \quad ,
\end{aligned} \tag{5}$$

where  $\lambda_{a,s}$  represents the force of infection of strain  $s$  affecting individuals of age  $a$ . The quantity  $\Phi_v$  is a binary variable used to switch between a positive and negative multiplier depending on vaccination status. It is equal to 1 when  $v = “1”$  and  $-1$  when  $v = “0”$ . In other words,  $\Phi_v = 2\mathbb{1}_{v=“1”} - 1$ .

The force of infection was calculated as:

$$\lambda_{a,s}(t) = \beta e^{2W(t)} \psi_s \sum_{\alpha \in \mathcal{A}} \sum_{v \in \mathcal{V}} \frac{c_{a,\alpha}(t)}{N_{\alpha,v}} \left( 0.5 \sum_{k=3}^4 E_{\alpha,v,s}^k + \sum_{k=1}^4 I_{\alpha,v,s}^k \right) \quad . \tag{6}$$

In the previous equation,  $\beta$  represents the unadjusted risk of transmission per contact,  $\psi_s$  is the relative infectiousness of strain  $s$ , and  $W(t)$  is the random process introduced in Section 1.2.6. The size of the population of age  $\alpha$  with vaccination status  $v$  is denoted  $N_{\alpha,v}$ . The term  $c_{a,\alpha}(t)$  is a single element of the contact matrix  $C(t)$  introduced in Equation 3. It represents the average numbers of contacts per day that a individual of age  $a$  has with individuals of age  $\alpha$ .

### 1.2.8 Model parameters

## 1.3 Estimation of COVID-19-related hospital pressure and deaths

The transmission model described in Section 1.2 provides estimates of COVID-19 incidence over time, disaggregated by age, vaccination status and strain. We combine these incidence estimates with the age-, vaccination- and strain-specific risks of hospitalisation and deaths, as well as statistical distributions of time to events to compute COVID-19-related hospital pressure and deaths over time.

Table S2: Model parameters

Parameter	Value/Distribution	Evidence
Transmission probability per contact <sup>C</sup>	Uniform (0.01, 0.1)	-
Mean active disease period (days)	8.0	assumed
Country-specific IFR multiplier ( $m_C$ ) <sup>C</sup>	Uniform (0.5, 1.5)	-
Uncertainty multiplier for school contacts ( $m_S$ ) <sup>C</sup>	Uniform (0.8, 1.2)	assumed
Prop. students on-site during “Partially open” periods <sup>C</sup>	Uniform (0.1, 0.3)	assumed
VE against infection	0.5	assumed
VE against hospitalisation	0.9	assumed
VE against death	0.9	assumed
Time from symptom onset to hospitalisation (days)	Gamma (shape=5, mean=7.7)	ISARIC report 4th Oct 2020 (Mendeley citation key Pritchard2020)
Hospital stay duration (days)	Gamma (shape=5, mean=12.8)	ISARIC report 4th Oct 2020 (Mendeley citation key Pritchard2020)
Time from symptom onset to death (days)	Gamma (shape=10, mean=15.93)	Malahat Khalili et al. . <a href="https://doi.org/10.1017/S0950268820001430">https://doi.org/10.1017/S0950268820001430</a>
<b>Strain-specific parameters</b>		
<i>Wild-type strain</i>		
Infection seed time	Calibrated, with country-specific priors	Varied between 1 Jan 2020 and first date with reported deaths
Mean incubation period (days)	6.65	Wu et al. doi:10.1001/jamanetworkopen.2022.28008
<i>Delta variant</i>		
Infection seed time	Country-specific	GISAID database
Mean incubation period (days)	4.41	Wu et al. doi:10.1001/jamanetworkopen.2022.28008
Relative intrinsic transmissibility (ref. wild-type)	1.38	DOI: 10.1503/cmaj.211248
Relative risk of hospitalisation (ref. wild-type)	2.1	Fisman et al. doi: 10.1503/cmaj.211248
Relative risk of death (ref. wild-type)	2.3	Fisman et al. doi: 10.1503/cmaj.211248
Prop. escaping vaccine immunity against infection	0.32	DOI: 10.1503/cmaj.211248
<i>Omicron variant</i>		
Infection seed time	Country-specific	GISAID database
Mean incubation period (days)	3.42	Wu et al. doi:10.1001/jamanetworkopen.2022.28008
Relative intrinsic transmissibility (ref. wild-type)	1.92	DOI: 10.1503/cmaj.211248
Relative risk of hospitalisation (ref. wild-type)	0.861	Nyberg doi.org/10.1016/S0140-6736(22)00462-7
Relative risk of death (ref. wild-type)	0.713	Nyberg doi.org/10.1016/S0140-6736(22)00462-7
Prop. escaping vaccine immunity against infection	0.55	DOI: 10.1503/cmaj.211248

### 1.3.1 COVID-19-related hospital pressure

The risk of hospitalisation given infection was expected to vary markedly by setting. For example, different countries may have different criteria for whether or not a COVID-19 patient should be admitted to a hospital. This makes it difficult to provide accurate estimates of hospitalisation rates for multiple countries.

For this reason we introduced a universal indicator named “hospital pressure” in our analysis. This indicator was obtained by considering the age-specific risk of hospitalisation given infection observed in the first year of the pandemic in the Netherlands, adjusted for vaccination status and for the infecting strain (Table S2). The “hospital pressure” indicator can therefore be interpreted as the level of hospital occupancy that would be observed in the analysed country if the rates of hospitalisation given infection in this country were the same as for the Netherlands. This quantity is expected to vary proportionately with occupancy over time, providing an indicator of hospital pressure. Note that this indicator was used in order to make comparisons between scenarios, such that one should interpret the relative differences between scenarios rather than the absolute values of the indicator.

Let us denote  $i_{a,v,s}(t)$  the number of new disease episodes estimated to start at time  $t$  for people aged  $a$  with vaccination status  $v$  and infected with strain  $s$ . The number of new hospital admissions occurring at time  $t$  was calculated using the following convolution product:

$$\eta(t) = \sum_{a,v,s} \kappa_{a,v,s} \int_{u \geq 0} i_{a,v,s}(t-u) g_h(u) du \quad , \quad (7)$$

where  $\kappa_{a,v,s}$  is the risk of hospitalisation given infection for age  $a$ , vaccination status  $v$  and strain  $s$  based on the Netherlands data, and  $g_h$  is the probability density function of the statistical distribution chosen to represent the time from symptom onset to hospitalisation (Table S2).

We then computed the “hospital pressure” quantity  $h$ , which is an indicator of hospital occupancy level, by combining the number of new hospital admissions  $\eta$  with the statistical distribution used to model hospital stay duration:

$$h(t) = \int_{u \geq 0} \eta(t-u) (1 - \tau(u)) du \quad , \quad (8)$$



Table S3: Age-specific parameters for wild-type COVID-19

Age group	Rel. suscepitbil- ity to infection (ref. 15-69 y.o.)[3]	Proportion symp- tomatic[2]	Proportion of symptomatic pa- tients hospitalised [CURRENTLY USING Dutch GGD report from 4th August 2020, Table 3]	Infection rate[1]	fatality
0-4	0.36	0.533	0.0777	0.00003	
5-9	0.36	0.533	0.0069	0.00001	
10-14	0.36	0.533	0.0034	0.00001	
15-19	1.00	0.533	0.0051	0.00003	
20-24	1.00	0.679	0.0068	0.00006	
25-29	1.00	0.679	0.0080	0.00013	
30-34	1.00	0.679	0.0124	0.00024	
35-39	1.00	0.679	0.0129	0.00040	
40-44	1.00	0.679	0.0190	0.00075	
45-49	1.00	0.679	0.0331	0.00121	
50-54	1.00	0.679	0.0383	0.00207	
55-59	1.00	0.679	0.0579	0.00323	
60-64	1.00	0.803	0.0617	0.00456	
65-69	1.00	0.803	0.1030	0.01075	
70-74	1.41	0.803	0.1072	0.01674	
75-79	1.41	0.803	0.0703	0.03203	
80 and above	1.41	0.803	0.0703	0.08292	

where  $\tau$  is the cumulative density function of the statistical distribution chosen to represent the hospital stay duration (Table S2).

### 1.3.2 COVID-19 deaths

We estimated the number of COVID-19 deaths over time using a similar approach as for the hospital pressure indicator. We used the age-specific infection fatality rates reported in O’Driscoll et al. [1], adjusted for vaccination status and for the infecting strain to estimate COVID-19 mortality. Using the same notations as in Section 1.3.1, the number of COVID-19 deaths observed at time  $t$  was obtained by:

$$\mu(t) = m_C \sum_{a,v,s} ifr_{a,v,s} \int_{u \geq 0} i_{a,v,s}(t-u) g_d(u) du \quad , \quad (9)$$

where  $ifr_{a,v,s}$  is the risk of death given infection for age  $a$ , vaccination status  $v$  and strain  $s$ , and  $g_d$  is the probability density function of the statistical distribution chosen to represent the time from symptom onset to death (Table S2). We used the country-specific adjuster  $m_C$  to capture the fact that the infection fatality ratio is expected to vary by country, in part due to differences in COVID-19 death definition and reporting standards. This adjustment was automatically calibrated by the MCMC (Section 3).

## 2 Software and code used to conduct the analyses

### 2.1 Code

The Python code and data used to perform the analyses is fully available on Github at the following link: <https://github.com/monash-emu/AuTuMN>. In particular, [LIST LINKS TO SOME SUB-MODULES HERE](#).

### 2.2 Software implementation

We used the *summer2* Python package (v 1.2.5) to implement the model. This is a domain specific library for compartmental epidemiological models that addresses a few of the key concerns as follows.

#### 2.2.1 Application Programming Interface (API)

The model specification was done via a simple yet expressive Python API, while the numerical implementation was largely autogenerated by the *summer2* package at runtime. This specification is composable via stratification classes and other reusable components, thus the complexity of the software is kept to a minimum, reducing cognitive overhead and greatly reducing the possibility for error.

#### 2.2.2 Optimising compiler

The *summer2* package uses the *jax* library as its computational backend, meaning that while the specification of models is done largely in Python, the model execution itself is transformed via an optimising compiler into fast native code. This brings the model runtime from several seconds (for a naive implementation) to under 50ms per iteration, which was necessary to perform the computationally-intensive calibration tasks described in Section 3.

## 3 Model calibration and uncertainty propagation

The model was calibrated using a Bayesian approach. In particular, we used the Differential Evolution Metropolis (DEMetropolis) algorithm implemented with the *PyMC* Python package (v.5.2.0) to sample parameters from their posterior distributions. For each country, we ran 32 independent DEMetropolis chains of 20,000 iterations, each starting from a different starting point. The starting points were obtained after conducting 32 independent optimisation searches performed with the Covariance Matrix Adaptation Evolution Strategy (CMA-ES) method, implemented with the *nevergrad* Python package (v.0.6.0) and with a budget of 20,000 model evaluations per search. The 32 optimisation searches were initialised

randomly using Latin Hypercube Sampling based on the parameter priors, in order to promote broad dispersion across the parameter space. Our calibration approach required a total of 1,280,000 model evaluations per country, which were completed in about one hour on a machine with 32 cpus and 128-GiB memory.

For each country, the results presented in the manuscript are associated with 1000 parameter sets randomly sampled from the posterior distributions obtained from DEMetropolis sampling (after discarding the first 10,000 iterations for each chain). The definitions of the prior distributions and the likelihood are detailed in the following sections.

### 3.1 Parameters varied during calibration

The parameters varied during calibration along with their associated prior distributions are listed in Table S2 and indicated with the superscript <sup>C</sup>. We used uniform prior distributions for all calibrated parameters. The primary parameters varied during calibration are the unadjusted risk of transmission per contact ( $\beta$ ), the IFR multiplier ( $m_C$ ), the infection seeding times of each strain and the proportion of students on-site during “Partially open” periods. Note that the values of the random process  $W_t$  described in Section 1.2.6 are also treated as calibrated parameters by the MCMC. The Gaussian auto-regressive component described in Equation 4 is incorporated in the posterior likelihood computation (Section 3.3).

### 3.2 Calibration targets

Model calibration was performed independently for each country. All models were fitted to the reported number of COVID-19 deaths over time. We used the daily number of COVID-19 deaths reported by Our World in Data and applied a 7-day moving average to the observed data.

In addition, for countries where a nationally representative seroprevalence survey had been conducted ( $n=44$ ), we include seroprevalence data in the calibration likelihood. We used the online platform SeroTracker to extract country-specific seroprevalence estimates. The estimates had to verify the following conditions to be included in the analysis as calibration targets:

- Be aligned with the World Health Organization’s Unity protocol (WHO Unity) for general population seroepidemiological studies.
- Have a sample size greater than 599 (minimum sample size recommended in [WHO sero-surveys protocol](#))
- Have a sampling start date later than 1 May 2020 (to avoid very early surveys that may be less accurate)
- Be nationally representative (as classified by SeroTracker)

We further excluded studies that focused on specific population subgroups presenting a risk of selection bias for seroprevalence (e.g. pregnant women, slum population, healthcare workers, quarantine workers). Finally, to minimise the risk of interference with vaccination, we only included studies for which the vaccination coverage at the time of the survey (midpoint date) was lower than 10% of the measured seroprevalence.

All the criteria listed above were first verified systematically using the SeroTracker database, and the extracted studies were then analysed individually by one author (RR) to check that all inclusion criteria were verified.

When a seroprevalence study was restricted to a specific age-group, we matched the survey estimate to the modelled seroprevalence measured in the closest modelled age-group. For example, as the age-group reported in the Kenya survey was 16-64 years old, we used the modelled seroprevalence in the age-group 15-69 years-old to inform model calibration.

When multiple estimates were available for a country, we selected the highest ranked estimate after ordering by the following preference criteria (applied in the presented order):

1. Lowest risk of bias (according to SeroTracker),
2. Latest sampling start date (when a greater number of infections have occurred, and to avoid bias due to early geographic heterogeneity),

### 3. Largest sample size.

The estimates used to inform the models are summarised in Table S4, with the associated original reports accessible by clicking on the countries' names.

Table S4: Seroprevalence data extracted from SeroTracker (national surveys).

country	sampling start date	sampling end date	age min	age max	denom. value	serum pos prevalence	estimate grade	overall risk of bias
<a href="#">Australia</a>	2020-11-03	2021-03-12		19	1685	0.0023	National	High
<a href="#">Austria</a>	2020-06-05	2020-12-04	18	72	20228	0.025	National	High
<a href="#">Belgium</a>	2020-10-12	2020-10-17		101	2966	0.0418	National	Low
<a href="#">Brazil</a>	2020-05-14	2020-06-23			89362	0.023	National	Moderate
<a href="#">Canada</a>	2021-04-13	2021-04-30	17		16931	0.2692	National	Moderate
<a href="#">Chile</a>	2020-09-25	2020-11-25	7	94	2493	0.104	National	Low
<a href="#">Colombia</a>	2020-09-21	2020-12-11	5	80	17863	0.3253	National	Moderate
<a href="#">Croatia</a>	2020-12-15	2021-02-15			1436	0.251	National	High
<a href="#">Czechia</a>	2021-02-01	2021-03-31	18		19548	0.51	National	High
<a href="#">Denmark</a>	2020-12-01	2020-12-31	12		4044	0.043	National	Low
<a href="#">Ecuador</a>	2020-10-12	2020-10-19			1250	0.1168	National	Moderate
<a href="#">Egypt</a>	2021-01-15	2021-06-15			2360	0.463	National	Moderate
<a href="#">France</a>	2020-05-11	2020-05-17			3592	0.0493	National	Low
<a href="#">Germany</a>	2020-10-26	2020-11-18	18		9929	0.011	National	Moderate
<a href="#">Honduras</a>	2020-06-16	2020-06-23	5		792	0.062	National	Moderate
<a href="#">Hungary</a>	2020-05-01	2020-05-16	14		10474	0.0068	National	Low
<a href="#">India</a>	2020-12-18	2021-01-06	10		28598	0.241	National	Low
<a href="#">Iran</a>	2021-01-15	2021-03-15	10	90	7411	0.342	National	Low
<a href="#">Israel</a>	2020-06-28	2020-09-14			54357	0.046	National	Moderate
<a href="#">Italy</a>	2020-05-25	2020-07-15			64660	0.025	National	Moderate
<a href="#">Japan</a>	2020-06-01	2020-06-07	20		7950	0.001	National	Moderate
<a href="#">Jordan</a>	2020-12-27	2021-01-06			5044	0.342	National	Moderate
<a href="#">Kazakhstan</a>	2020-07-16	2021-07-07			85346	0.63	National	High
<a href="#">Kenya</a>	2021-01-03	2021-03-15	16	64	3018	0.485	National	Moderate
<a href="#">Lebanon</a>	2020-12-07	2021-01-15			2058	0.185	National	Low
<a href="#">Lithuania</a>	2020-08-10	2020-09-10	18	92	3089	0.014	National	Moderate
<a href="#">Malawi</a>	2020-10-14	2020-12-08			4261	0.078	National	Low
<a href="#">Mexico</a>	2020-08-15	2020-11-15	3	12	944	0.187	National	Low
<a href="#">Nepal</a>	2020-10-09	2020-10-22			3040	0.144	National	Low
<a href="#">Nigeria</a>	2021-06-29	2021-08-21			4904	0.789	National	Moderate
<a href="#">Norway</a>	2020-12-27	2021-02-13			1912	0.032	National	Moderate
<a href="#">Oman</a>	2020-11-08	2020-11-19	5		4064	0.22	National	Moderate
<a href="#">Pakistan</a>	2020-10-21	2020-11-08			4998	0.0702	National	Moderate
<a href="#">Portugal</a>	2020-09-08	2020-10-14			13398	0.022	National	Moderate
<a href="#">Senegal</a>	2020-06-15	2020-10-15			3231	0.204	National	High
<a href="#">Singapore</a>	2020-11-15	2020-12-15	23	83	937	0.0016	National	High
<a href="#">Slovenia</a>	2020-10-17	2020-11-10		99	1211	0.0429	National	Low
<a href="#">South Africa</a>	2021-01-15	2021-05-15	15	69	16762	0.474	National	Moderate
<a href="#">Rep. of Korea</a>	2020-09-24	2020-12-09	18	86	4085	0.0039	National	Moderate
<a href="#">Spain</a>	2020-06-08	2020-06-22			62167	0.052	National	Low
<a href="#">Sweden</a>	2020-11-23	2020-12-04			3183	0.07	National	Moderate
<a href="#">USA</a>	2020-08-09	2020-12-08	18		4654	0.0471	National	Low
<a href="#">UK</a>	2020-08-24	2020-09-18	17		8230	0.061	National	Low
<a href="#">Zambia</a>	2020-07-04	2020-07-27			2704	0.021	National	Low

### 3.3 Likelihood definition

Let  $d_w$  denote the rounded average daily number of COVID-19 deaths during week  $w$ , and  $\hat{d}_w^\theta$  the associated predicted number of deaths according to the model with parameter set  $\theta$ .

For countries with seroprevalence data, let us denote  $\pi$  the measured seroprevalence proportion extracted from SeroTracker (Section 3.2). Let  $\hat{\pi}^\theta$  denote the modelled age-matched proportion ever infected by the time the survey was conducted (using the midpoint date) associated with the parameter set  $\theta$ . The likelihood was defined as follows for countries with seroprevalence data:

$$\mathcal{L}(\theta) := f_\sigma(\pi|\hat{\pi}^\theta) \times \prod_w g_r(d_w|\hat{d}_w^\theta) \quad , \quad (10)$$

where  $f_\sigma(\cdot|\mu)$  is the probability density function of a  $[0, 1]$ -truncated normal distribution with mean  $\mu$  and standard deviation  $\sigma$ ; and  $g_r(\cdot|\mu)$  is the probability mass function of a negative binomial distribution with mean  $\mu$  and overdispersion parameter  $r$ . The overdispersion parameter  $r$  was automatically estimated by the MCMC algorithm, while the standard deviation  $\sigma$  was set to different values depending on the SeroTracker-reported risk of bias associated with the seroprevalence estimate ( $\sigma = 0.05$  if “Low”,  $\sigma = 0.1$  if “Moderate”,  $\sigma = 0.2$  if “High”).

For countries without seroprevalence data, the likelihood equation reduces to:

$$\mathcal{L}(\theta) := \prod_w g_r(d_w|\hat{d}_w^\theta) \quad . \quad (11)$$

The likelihood functions described above represent the goodness of fit of a particular model parameterisation with regards to the targeted data. This quantity needs to be adjusted for the prior likelihood of the parameter set in order to compute the MCMC acceptance quantity  $\mathcal{Q}(\theta)$ . As we used uniform priors for all the parameters, the inclusion of the individual parameters’ priors in the acceptance quantity is not necessary. Indeed, their respective contributions would cancel out as the same quantity would appear in the numerator and the denominator of the MCMC acceptance quantity ratio. However, the auto-regressive relationship described in Equation 4 must be accounted for as part of the combined prior likelihood of a parameter set. This prevents unrealistically large fluctuations of the random process. If  $W^\theta$  represents the random process associated with the parameter set  $\theta$ , the overall MCMC acceptance quantity is obtained by:

$$\mathcal{Q}(\theta) = \mathcal{L}(\theta) \times \prod_{i=1}^n z_{W_{i-1}^\theta, \epsilon}(W_i^\theta) \quad , \quad (12)$$

where  $z_{\mu, \epsilon}(\cdot)$  represents the probability density function of the normal distribution  $\mathcal{N}(\mu, \epsilon)$ , and  $n$  is the number of random process updates.

## References

- [1] M. O'Driscoll et al. "Age-specific mortality and immunity patterns of SARS-CoV-2". In: *Nature* 590.7844 (2021), pp. 140–145. ISSN: 1476-4687. DOI: 10.1038/s41586-020-2918-0. URL: <https://www.ncbi.nlm.nih.gov/pubmed/33137809>.
- [2] Pratha Sah et al. "Asymptomatic SARS-CoV-2 infection: A systematic review and meta-analysis". In: *Proceedings of the National Academy of Sciences* 118.34 (2021), e2109229118. DOI: 10.1073/pnas.2109229118. URL: <http://www.pnas.org/content/118/34/e2109229118.abstract>.
- [3] J. Zhang et al. "Changes in contact patterns shape the dynamics of the COVID-19 outbreak in China". In: *Science* 368.6498 (2020), pp. 1481–1486. ISSN: 1095-9203. DOI: 10.1126/science.abb8001. URL: <https://www.ncbi.nlm.nih.gov/pubmed/32350060>.



Figure S4: Modelled vaccine coverage (lines) against data (dots).

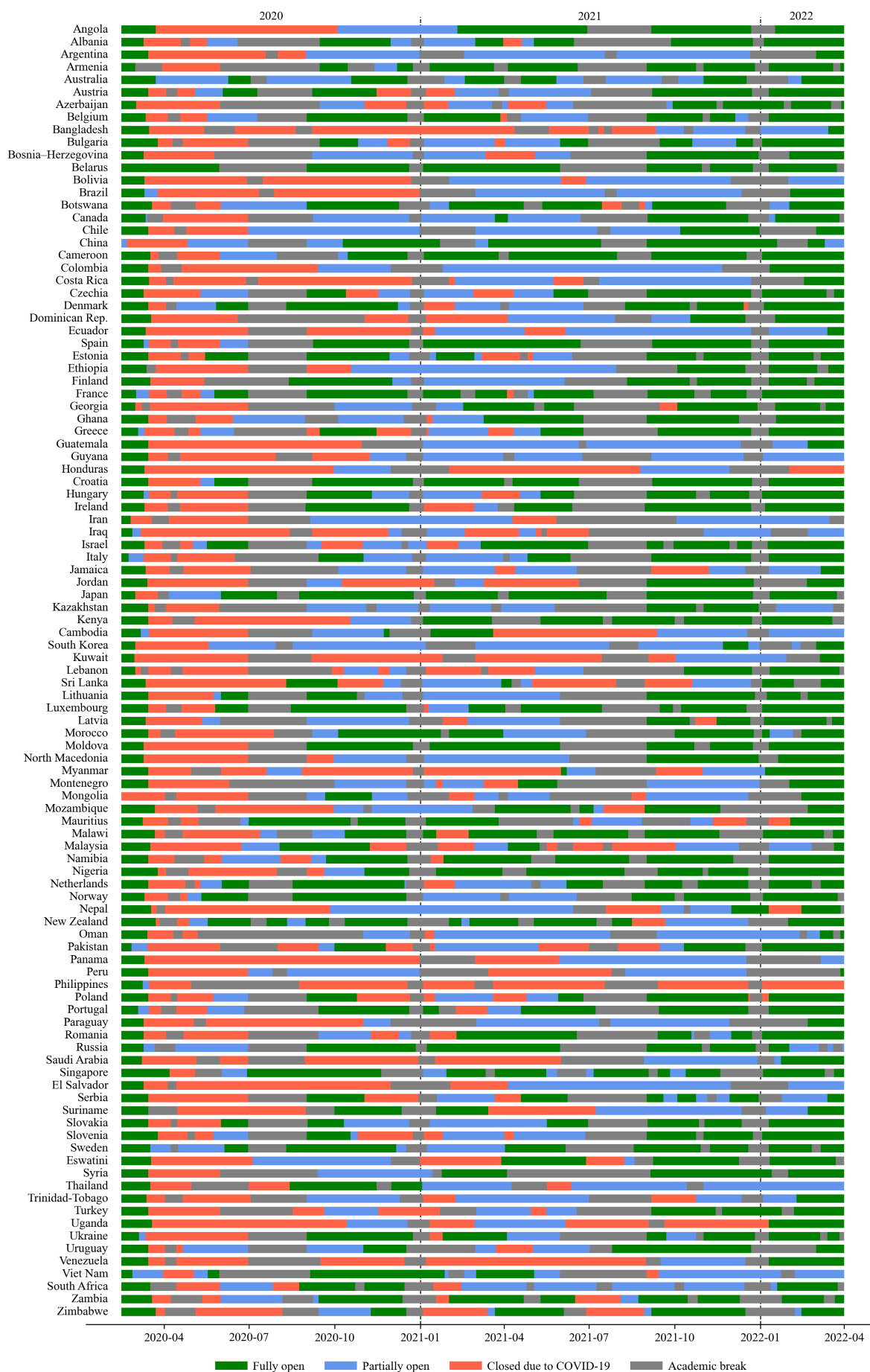


Figure S5: UNESCO data on school closures during the COVID-19 pandemic.



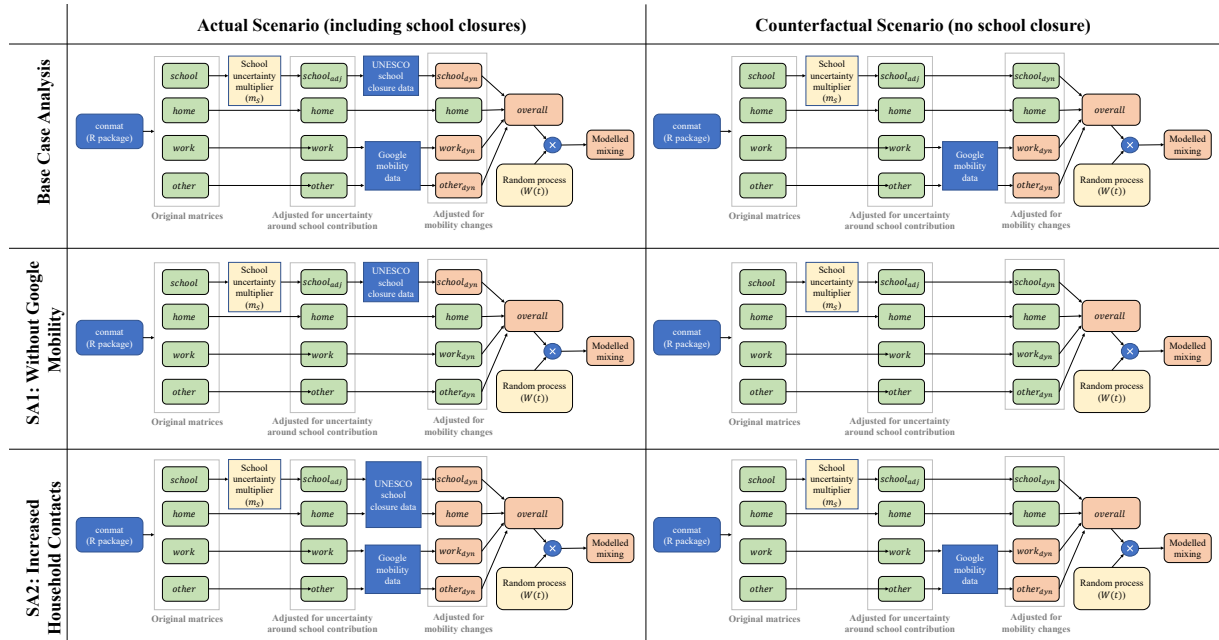


Figure S6: Modelled social mixing under Base Case and Sensitivity Analyses.