

# Supplementary file:

## Age-structure and transient dynamics in epidemiological systems

### S1 Plotting Figures 2–3

The trajectory shown in Figure 2 is that of a model with two age classes ( $M = 2$ ) with birth and death rate  $\mu = \frac{1}{50}\text{yr}^{-1}$ , incubation rate  $\sigma = \frac{365}{8}\text{yr}^{-1}$ , recovery rate  $\gamma = \frac{365}{5}\text{yr}^{-1}$ , aging rate of  $\nu_1 = \frac{1}{12}$  and infection from outside  $\iota = 10^{-4}\text{yr}^{-1}$ . Using these values, the younger age class  $N_1$  has an average age of 9.7 years old. We set the transmission rate within the  $N_1$  class to be high and seasonal due to school-term forcing,  $\beta_{1,1}(t) = bT(t)$  where  $T(t)$  is the corrected term-time forcing function (Keeling and Rohani, 2008), given by

$$T(t) = \begin{cases} \frac{1+s}{(1+s)0.773+(1-s)(1-0.773)}, & \text{if } \text{mod}(t, 1) \in \left(\frac{7}{365}, \frac{100}{365}\right] \cup \left(\frac{115}{365}, \frac{199}{365}\right] \cup \left(\frac{252}{365}, \frac{300}{365}\right] \cup \left(\frac{308}{365}, \frac{365}{365}\right] \\ \frac{1-s}{(1+s)0.773+(1-s)(1-0.773)}, & \text{otherwise.} \end{cases}$$

Here  $t$  is assumed to be in years and  $\text{mod}(t, 1)$  is the decimal part of  $t$ . From this definition, the average of  $bT(t)$  over a one-year period is  $b$ .

The parameter values for Figures 2–3 were chosen so that the pre-vaccine era dynamics display a biennial cycle similar to what we see for measles in London. In particular, we chose parameters of the homogeneous model used in Figure 3 that are close to the values of the best homogeneous model fit that we found for the 1945–1990 data in Section 4. The homogeneous model in Figure 3 has a transmission rate magnitude of  $b = b_{\text{hom}} = 1750\text{yr}^{-1}$  and amplitude of seasonality of  $s = s_{\text{hom}} = 0.23$ . We assumed that the age-structured model had contact rates of the form  $\beta_{1,1} = bT(t)$  and  $\beta_{2,2} = \beta_{1,2} = \beta_{2,1} = \frac{1}{2}b$ . We then calculated the values of  $b$  and  $s$  so that the age-structured model has a magnitude and amplitude that coincides with the homogeneous model in Figure 3 before vaccination. We found that this required  $b = b_{\text{age}} = 435$  and the amplitude of seasonality for children to be  $s = s_{\text{age}} = 0.257$ . Using the next-generation matrix method (Heesterbeek and Diekmann, 2000; van den Driessche and Watmough, 2008), we computed that the basic reproduction numbers of the models are about 7.8 for the age-structured model and about 24 for the homogeneous model (computed under the simplifying assumption that school-aged children transmission is constant at  $b$ ).

The mean age of the infected class in Figure 2 was calculated by first finding the two-year-averaged forces of infection for both the younger and older age classes. The reciprocals of these were estimated to be the two-year-averaged mean age at which these different classes get infected. The overall mean age of infection were then calculated from these by multiplying age classes' mean age of infected by its two-year-averaged fraction of contributions to the total infections, and adding the results.

### S2 Calculation of $s_{\text{age}}$

The age-structured model does not have the exact same bifurcation diagram as a homogeneous model due to the overall effective transmission rate  $\hat{\beta}(t)$  not necessarily having the same form of the term-

time forcing function. However our tests using simulations of the model show that the age-structured model often has similar dynamics as a homogeneous model when the magnitude and amplitude of the effective transmission rate  $\hat{\beta}(t)$  is about the same. Thus it is informative, if not exact, to plot the change in the magnitude and amplitude of  $\hat{\beta}(t)$  on the two-dimensional bifurcation diagram of the homogeneous model (as in Figure 3).

The effective transmission rate  $\hat{\beta}(t)$  magnitude for the age-structured model was computed by taking the moving average value of  $\hat{\beta}(t)$  over two years. The amplitude is more complicated due to the form of the term-time forcing function. We approximate it in the following manner: Let  $p = 0.773$  (the fraction of time that school is in session) and let  $q(t)$  be the absolute value of the difference between  $\hat{\beta}(t)$  and the magnitude. Then, the amplitude can be calculated from:

$$q(t) = p \left( \frac{1 + s_{\text{age}}}{(1 + s_{\text{age}})p + (1 - s_{\text{age}})(1 - p)} - 1 \right) + (1 - p) \left( 1 - \frac{1 - s_{\text{age}}}{(1 + s_{\text{age}})p + (1 - s)(1 - p)} \right). \quad (\text{S2.1})$$

This yields,

$$s_{\text{age}}(t) = \frac{q(t)}{1 - (2p - 1)(q + 2p - 1)}. \quad (\text{S2.2})$$

We computed the changing amplitude  $s_{\text{age}}(t)$  from the two-year moving average of  $q(t)$ . We then plotted the two-year moving average of the amplitude against the magnitude (using the two-year moving average value of  $\hat{\beta}(t)$ ) in Figure 3.

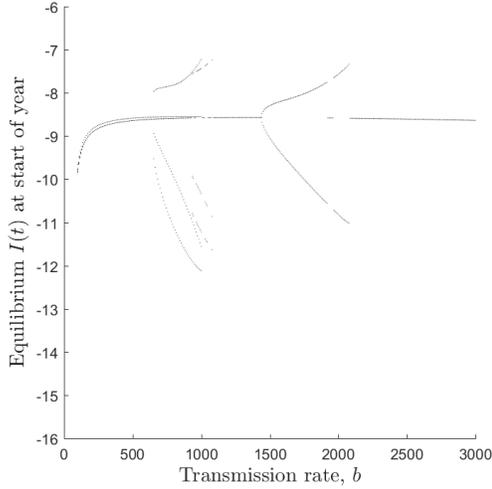
### S3 Bifurcation diagrams

We first conducted a search for different periodic solutions (one-year period to 16-year periods) of the homogeneous and age-structured models using long-term numerical integration. We did this by solving the systems of equations in the natural log scale and using MATLAB's `ode23`, starting and stopping the numerical integrator at all time points for which the transmission rates changed. This was followed by using numerical continuation to solve for both the periodic solutions, the monodromy matrix and the Floquet multipliers to determine the stability of periodic solutions with periods of one, two and three. The results are shown in Figure S3.1. Plots of some of these periodic solutions are shown in Figure S3.2. A comparison of the pre-vaccine era dynamics is shown in Figure S3.3 and a comparison of the vaccine era dynamics is shown in Figure S3.4. A comparison of the behaviour of the different Floquet multipliers are shown in Figure S3.5.

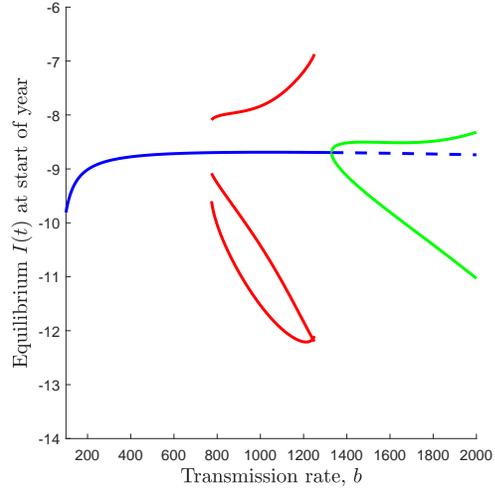
### S4 Details on fitting the models to the London data

All data and code required to reproduce the results of Section 5 are available via the Dryad Digital Repository (DOI:10.5061/dryad.vj645q8).

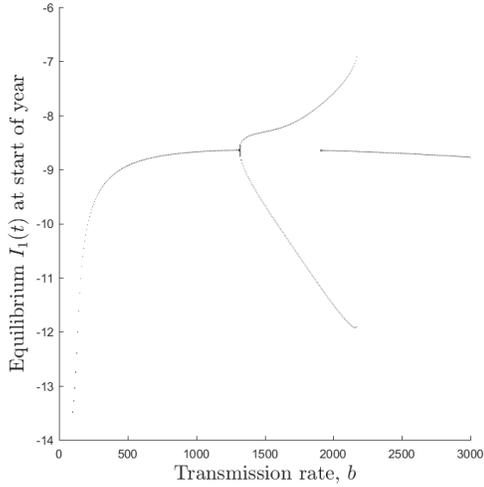
We obtained the annual number of births and population sizes for greater London from 1944 to 1994 (Grenfell et al., 2001). By adding up the births and population sizes from for all the boroughs that comprise inner London, we also obtained the corresponding demographic information from inner London from borough-specific data from 1939 to 1964 (Grenfell et al., 2001). These two sets of data were compared during the overlapping period from 1944–1964. We found that during this period, the ratio of births from inner London to that from greater London ranged from 0.373 to 0.445, and the ratio of population sizes ranged from 0.386 to 0.411. Thus we made the assumption that we could use the complete data from greater London to infer the demographics in inner London, by multiplying both the number of births and population sizes by 0.4.



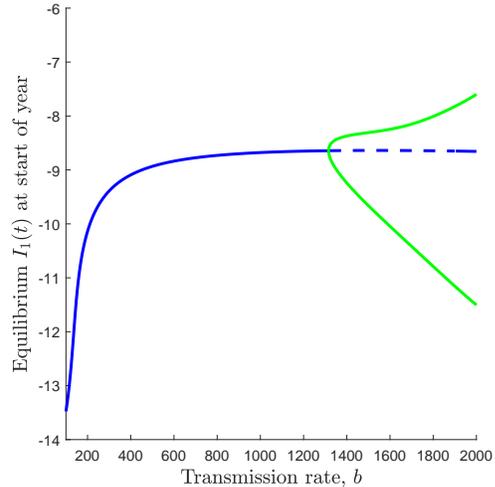
(a) Homogeneous model, computed using long-term numerical simulations



(b) Homogeneous model, computed using numerical continuation

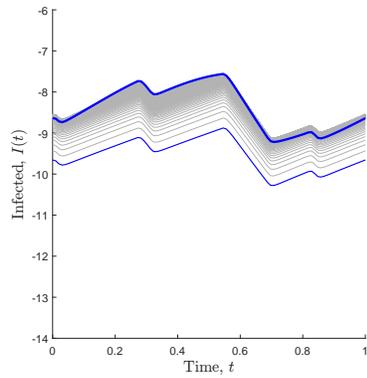


(c) Age-structured model, computed using long-term numerical simulation

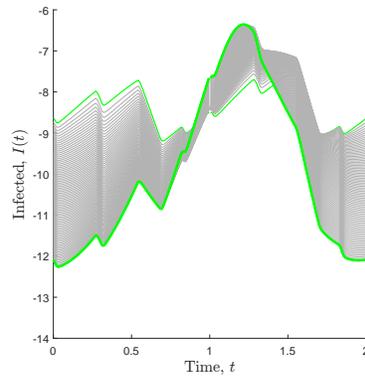


(d) Age-structured model, computed using numerical continuation

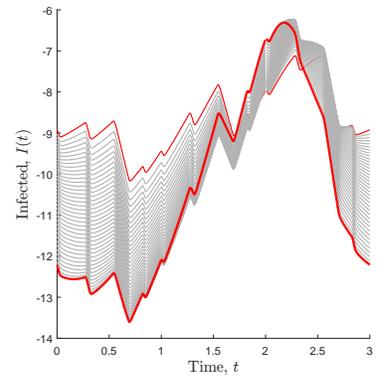
Figure S3.1: Bifurcation diagrams showing the existence of one-year, two-year and three-year periodic solutions for fixed amplitudes ( $s = 0.23$  for the homogeneous case). The vertical axes shows the log-transformed values of the states. In (b) and (d), blue represents a one-year periodic solution, green represents a two-year periodic solution and red represents a three-year periodic solution. Solid lines indicate stable and dashed lines indicate unstable solutions (determined using numerically computed Floquet multipliers).



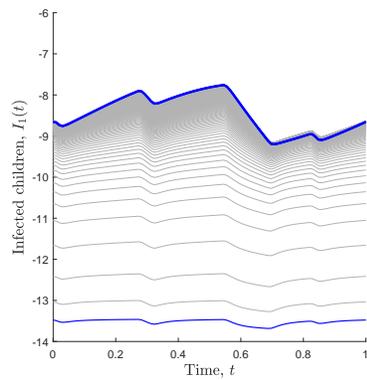
(a) Homogeneous, Period 1



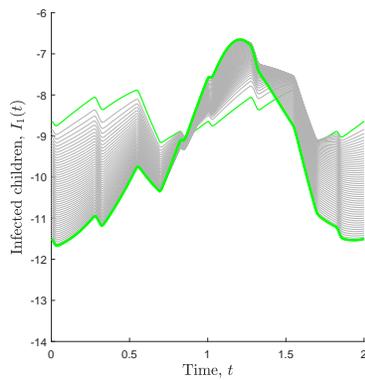
(b) Homogeneous, Period 2



(c) Homogeneous, Period 3



(d) Age-structured, Period 1



(e) Age-structured, Period 2

Figure S3.2: Plots of some of the stable periodic orbits from Figure S3.1 as  $b$  changes from its minimum (thin coloured line) to maximum value (thick coloured line). The vertical axes show the log-transformed values of the indicated infected class.

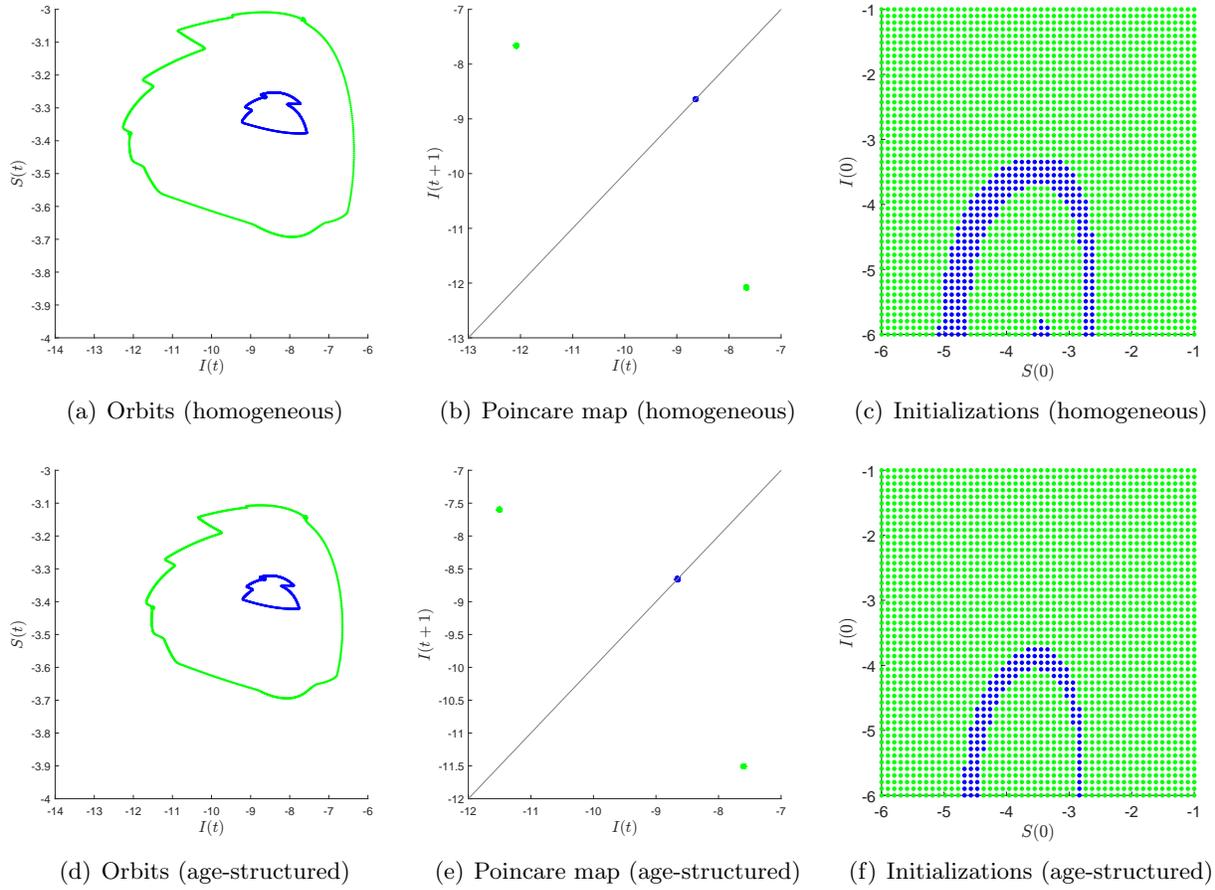


Figure S3.3: Pre-vaccine era dynamics. The first row shows dynamics from the homogeneous model. The second row shows dynamics from the age-structured model. All figures show the log-transformed values of the states.

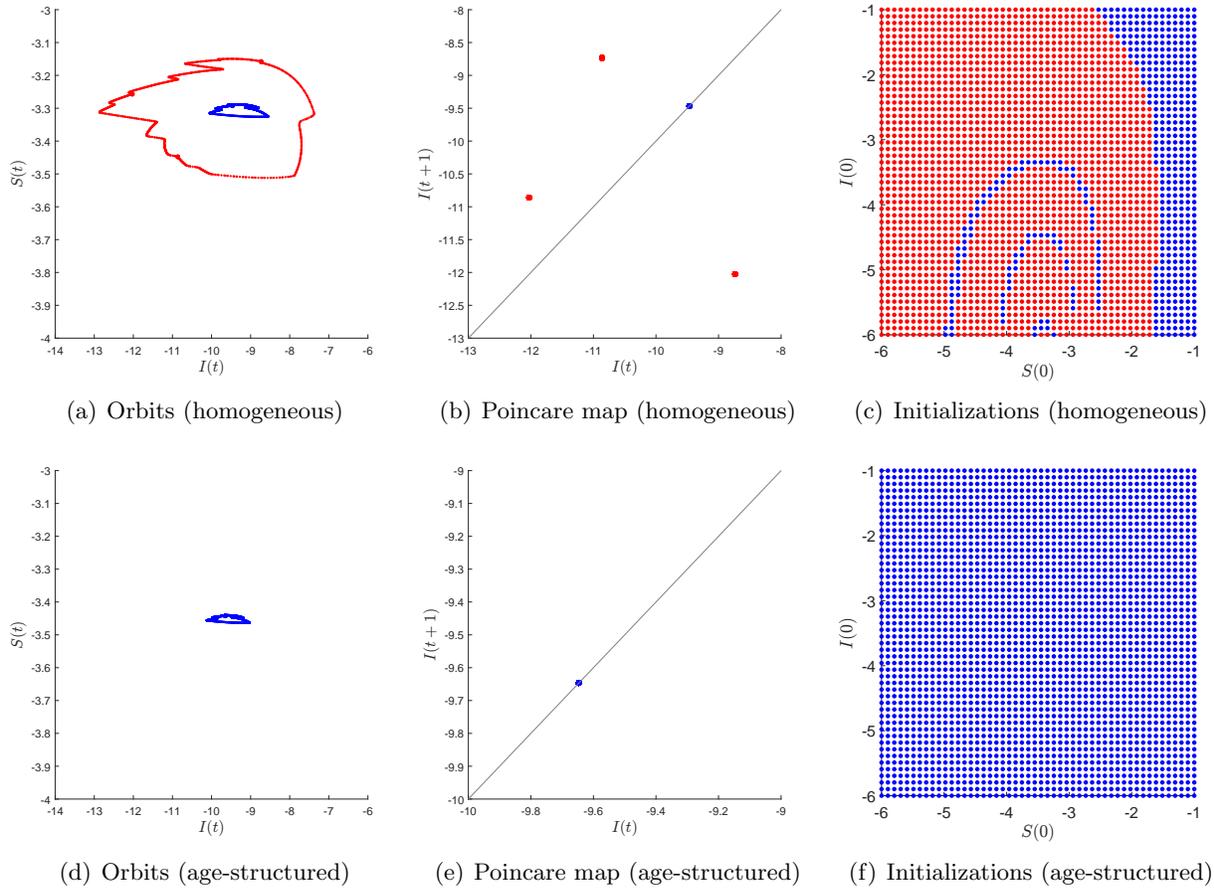


Figure S3.4: Steady-state dynamics at 60% vaccination coverage. The first row shows dynamics from the homogeneous model. The second row shows dynamics from the age-structured model. All figures show the log-transformed values of the states.

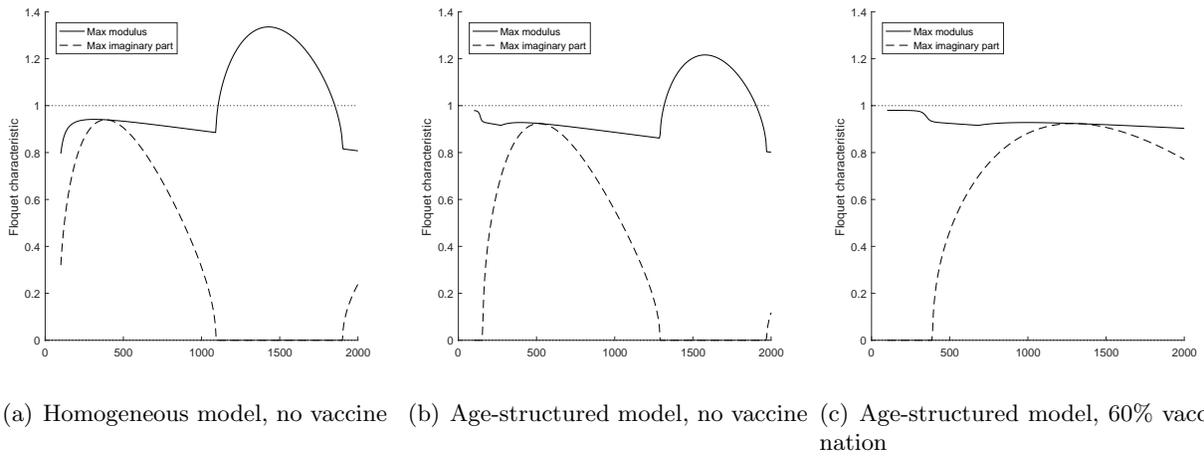


Figure S3.5: Characteristics of the Floquet multipliers for the different models. The imaginary part of the Floquet multipliers indicates the occurrence of a second frequency in the system, which explains the appearance of longer inter-epidemic periods after the start of vaccination.

To fit the models to data, we used time-varying functions  $\mu_B(t)$  and  $N(t)$  using the linear interpolations of the data on birth rates and total population size. A time-varying death and net migration rate  $\mu_d(t)$  was calculated using the birth rates and the population size (note that  $\mu_d(t)$  can be negative). The constant vaccine coverage  $p$  was also replaced by linear interpolations  $p(t)$  of the national vaccine estimates for England.

The age-structured model fitted to the London data consisted of three age classes ( $M = 3$ ). The first age class corresponded to very young children with aging rate of  $\nu_1 = \frac{1}{4}$  yr<sup>-1</sup>. The second age class corresponded to older children who are in school with aging rate of  $\nu_2 = \frac{1}{12}$  yr<sup>-1</sup>. The last age class corresponded to adults. Assuming a mean lifetime of 70 years ( $\mu = \frac{1}{70}$ ) and using (2.1), the mean age of individuals in the first age class is  $\frac{1}{\mu+\nu_1} \approx 3.8$  years and the mean age in the second age class is  $\frac{1}{\mu+\nu_1} + \frac{1}{\mu+\nu_2} \approx 14$  years.

To initialize the models, we used the steady state age-distributions of the different compartments and age classes assuming a constant birth and death rate of  $\mu = \frac{1}{70}$  yr<sup>-1</sup>. Here we use the subscripts  $Y$ ,  $C$  and  $A$  to denote the young children, older children and adult age classes. The initial number of individuals in each age class values were initialized using an approximately steady state age-distribution for the young children ( $N_Y(a)$ ), children ( $N_C(a)$ ) and adults ( $N_A(a)$ ) where,

$$\begin{aligned} N_Y(a) &= \mu e^{-(\mu+\nu_1)a}, \\ N_C(a) &= \frac{\mu\nu_1}{\nu_1 - \nu_2} [e^{-(\mu+\nu_2)a} - e^{-(\mu+\nu_1)a}], \\ N_A(a) &= \frac{\mu\nu_1\nu_2}{\nu_1 - \nu_2} \left[ \frac{1 - e^{-\nu_2 a}}{\nu_2} - \frac{1 - e^{-\nu_1 a}}{\nu_1} \right] e^{-\mu a}. \end{aligned} \quad (\text{S4.1})$$

Under the assumption of constant birth and death rates, the equilibrium fractions of each age class is,

$$\begin{aligned} N_Y^* &= \int_0^\infty N_Y(a) da = \frac{\mu}{\mu + \nu_1}, \\ N_C^* &= \int_0^\infty N_C(a) da = \frac{\mu\nu_1}{(\mu + \nu_1)(\mu + \nu_2)}, \\ N_A^* &= \int_0^\infty N_A(a) da = \frac{\nu_1\nu_2}{(\mu + \nu_1)(\mu + \nu_2)}. \end{aligned} \quad (\text{S4.2})$$

To determine the fraction of susceptible  $S_i^*$  within each subpopulation  $i$  ( $i = Y, C, A$ ), we used the following formula,

$$S_i^* = \frac{\int_0^\infty (1 - F(a)) N_i(a) da}{\int_0^\infty N_i(a) da}.$$

The function  $F(a)$  is the age-specific susceptibility to measles which can be computed for different regions and time periods using either age-stratified measles notification data or sero-prevalence data. An estimate of the number of susceptibles  $S_i^*$  within each subpopulation  $N_i$  for different time periods is given in Table S4.1.

We note that Grenfell and Anderson (1985) included a susceptibility function for England and Wales based on age-stratified notification data from 1948–1968 (pre-vaccine era) in their paper. However, we were unable to use this to compute the susceptible fractions since this function yielded values much larger than one.

Since our model simulations start in the year 1945, we decided to initialize the fraction of susceptible young children at  $S_Y^* = 0.863$  and fraction of susceptible adults at  $S_A^* = 0.0173$  (using the function derived from sero-prevalence data from small families New Haven, CT in the pre-vaccine era). All other initial states of the model were fitted. We did not fix the susceptible fractions for older

Table S4.1: Comparison of susceptible fractions using different sources of susceptibility functions  $F(a)$  given in Grenfell and Anderson (1985)

Data used to produce susceptibility function	Young children susceptible fraction, $S_Y^*$	Older children susceptible fraction, $S_C^*$	Adult susceptible fraction, $S_A^*$
New Haven, CT (sero-prevalence in small families in 1959)	0.863	0.34	0.0173
New Haven, CT (sero-prevalence in big families in 1959)	0.723	0.214	0.00915
England and Wales (age-stratified case notification from 1969–1982)	0.666	0.198	0.0088

children in case this strongly affects the phase of the epidemics. For the same reason, we did not fix the initial susceptible fraction for the homogeneous model that was compared to the age-structured model.

All transmission rates except for the transmission rate among older children were set to be constant. Here we again use the subscripts  $Y$ ,  $C$  and  $A$  to denote the young children, older children and adult age classes. We set  $\beta_{1,1}(t) = \beta_{YY}$ ,  $\beta_{1,2}(t) = \beta_{2,1}(t) = \beta_{YC}$ ,  $\beta_{1,3}(t) = \beta_{3,1}(t) = \beta_{YA}$ ,  $\beta_{2,3}(t) = \beta_{3,2}(t) = \beta_{CA}$  and  $\beta_{3,3} = \beta_{AA}$ . The transmission rate with the age class of older children  $\beta_{2,2}(t)$  was set to be a seasonal B-spline with six basis functions (with six corresponding parameters to be fitted). These were generated using the `periodic.bspline.basis` function in `pomp` with degree equal to three and period of one year.

The exposed and infected compartments of each age class were each subdivided into  $K = 3$  subcompartments

$$E_i = \sum_{k=1}^K E_i^{(k)}, \quad I_i = \sum_{k=1}^K I_i^{(k)}$$

in order to allow for Erlang-distributed latent and infectious periods. The models were solved as discrete time, discrete state systems with small stepsize  $\Delta t = 0.005$  year.

We use the notation  $\Delta_{X,Y}(t)$  to denote a transition from compartment/subcompartment  $X$  to  $Y$  over time interval  $t$  to  $t + \Delta t$ . For  $i = 1, \dots, M$ , the we first calculate the transitions due to the disease process:

$$\begin{aligned}
 \Delta_{S_i, E_i^{(1)}}(t) &= \left(1 - \exp(-\lambda_i \Delta t)\right) S_i(t), \\
 \Delta_{E_i^{(k)}, E_i^{(k+1)}}(t) &= \left(1 - \exp(-K\sigma \Delta t)\right) E_i^{(k)}(t), \quad \text{for } k = 1, \dots, K-1, \\
 \Delta_{E_i^{(K)}, I_i^{(1)}}(t) &= \left(1 - \exp(-K\sigma \Delta t)\right) E_i^{(K)}(t), \\
 \Delta_{I_i^{(k)}, I_i^{(k+1)}}(t) &= \left(1 - \exp(-K\gamma \Delta t)\right) I_i^{(k)}(t), \quad \text{for } k = 1, \dots, K-1, \\
 \Delta_{I_i^{(K)}, R_i}(t) &= \left(1 - \exp(-K\gamma \Delta t)\right) I_i^{(K)}(t).
 \end{aligned} \tag{S4.3}$$

where the force of infection on age class  $i$  is given by,

$$\lambda_i(t) = \sum_{j=1}^M \beta_{i,j}(t) \frac{\sum_{k=1}^K I_j^{(k)}}{N_j} + \frac{\iota}{\sum_{j=1}^M N_j(t)} \quad (\text{S4.4})$$

and  $N_i = S_i + E_i + I_i + R_i$ . Each compartment were updated using these transitions. For  $i = 1, \dots, M$ ,

$$\begin{aligned} S_i(t + \Delta t) &= S_i(t) - \Delta_{S_i, E_i^{(1)}}(t), \\ E_i^{(1)}(t + \Delta t) &= E_i^{(1)}(t) + \Delta_{S_i, E_i^{(1)}}(t) - \Delta_{E_i^{(1)}, E_i^{(2)}}(t), \\ E_i^{(k)}(t + \Delta t) &= E_i^{(k)}(t) + \Delta_{E_i^{(k-1)}, E_i^{(k)}}(t) - \Delta_{E_i^{(k)}, E_i^{(k+1)}}(t), \quad \text{for } k = 2, \dots, K-1, \\ E_i^{(K)}(t + \Delta t) &= E_i^{(K)}(t) + \Delta_{E_i^{(K-1)}, E_i^{(K)}}(t) - \Delta_{E_i^{(K)}, I_i^{(1)}}(t), \\ I_i^{(1)}(t + \Delta t) &= I_i^{(1)}(t) + \Delta_{E_i^{(K)}, I_i^{(1)}}(t) - \Delta_{I_i^{(1)}, I_i^{(2)}}(t), \\ I_i^{(k)}(t + \Delta t) &= I_i^{(k)}(t) + \Delta_{I_i^{(k-1)}, I_i^{(k)}}(t) - \Delta_{I_i^{(k)}, I_i^{(k+1)}}(t), \quad \text{for } k = 2, \dots, K-1, \\ I_i^{(K)}(t + \Delta t) &= I_i^{(K)}(t) + \Delta_{I_i^{(K-1)}, I_i^{(K)}}(t) - \Delta_{I_i^{(K)}, R_i}(t), \\ R_i(t + \Delta t) &= R_i(t) + \Delta_{I_i^{(K)}, R_i}(t). \end{aligned} \quad (\text{S4.5})$$

The compartments were then updated to reflect aging from one age class to the next. For  $i = 1, \dots, M-1$ ,

$$\begin{aligned} \Delta_{V_i, V_{i+1}}(t) &= \left(1 - \exp(-\nu_i \Delta t)\right) V_i(t), \\ \Delta_{S_i, S_{i+1}}(t) &= \left(1 - \exp(-\nu_i \Delta t)\right) S_i(t), \\ \Delta_{E_i^{(k)}, E_{i+1}^{(k)}}(t) &= \left(1 - \exp(-\nu_i \Delta t)\right) E_i^{(k)}(t), \quad \text{for } k = 1, \dots, K, \\ \Delta_{I_i^{(k)}, I_{i+1}^{(k)}}(t) &= \left(1 - \exp(-\nu_i \Delta t)\right) I_i^{(k)}(t), \quad \text{for } k = 1, \dots, K, \\ \Delta_{R_i, R_{i+1}}(t) &= \left(1 - \exp(-\nu_i \Delta t)\right) R_i(t). \end{aligned} \quad (\text{S4.6})$$

For  $i = M$ , we set the transition terms in (S4.6) to zero. Then, for  $i = 1$ ,

$$\begin{aligned} V_i(t + \Delta t) &\leftarrow V_i(t + \Delta t) - \Delta_{V_i, V_{i+1}}(t) + p(t) \mu_B(t) \Delta t, \\ S_i(t + \Delta t) &\leftarrow S_i(t + \Delta t) - \Delta_{S_i, S_{i+1}}(t) + (1 - p(t)) \mu_B(t) \Delta t, \\ E_i^{(k)}(t + \Delta t) &\leftarrow E_i^{(k)}(t + \Delta t) - \Delta_{E_i^{(k)}, E_{i+1}^{(k)}}(t), \quad \text{for } k = 1, \dots, K, \\ I_i^{(k)}(t + \Delta t) &\leftarrow I_i^{(k)}(t + \Delta t) - \Delta_{I_i^{(k)}, I_{i+1}^{(k)}}(t), \quad \text{for } k = 1, \dots, K, \\ R_i(t + \Delta t) &\leftarrow R_i(t + \Delta t) - \Delta_{R_i, R_{i+1}}(t). \end{aligned} \quad (\text{S4.7})$$

and for  $i = 2, \dots, M$ ,

$$\begin{aligned} V_i(t + \Delta t) &\leftarrow V_i(t + \Delta t) - \Delta_{V_i, V_{i+1}}(t) + \Delta_{V_{i-1}, V_i}(t), \\ S_i(t + \Delta t) &\leftarrow S_i(t + \Delta t) - \Delta_{S_i, S_{i+1}}(t) + \Delta_{S_{i-1}, S_i}(t), \\ E_i^{(k)}(t + \Delta t) &\leftarrow E_i^{(k)}(t + \Delta t) - \Delta_{E_i^{(k)}, E_{i+1}^{(k)}}(t) + \Delta_{E_{i-1}^{(k)}, E_i^{(k)}}(t), \quad \text{for } k = 1, \dots, K, \\ I_i^{(k)}(t + \Delta t) &\leftarrow I_i^{(k)}(t + \Delta t) - \Delta_{I_i^{(k)}, I_{i+1}^{(k)}}(t) + \Delta_{I_{i-1}^{(k)}, I_i^{(k)}}(t), \quad \text{for } k = 1, \dots, K, \\ R_i(t + \Delta t) &\leftarrow R_i(t + \Delta t) - \Delta_{R_i, R_{i+1}}(t) + \Delta_{R_{i-1}, R_i}(t). \end{aligned} \quad (\text{S4.8})$$

After this, every compartment was multiplied by  $\exp(-\mu_d(t)\Delta t)$  (to correct for changes in population size due to death or migration) and then rounded off to the nearest integer value.

In order to fit this model to data, the number of transitions  $\Delta_{I_i^{(K)}, R_i}(t)$  were counted between each successive time points  $t_i$  where there is data on weekly measles reports. The number of true cases  $C_i$  is the number of transitions from the infected class to the recovered class from time  $t_{i-1}$  to  $t_i$ . We assumed a negative binomial reporting model (to allow for overdispersion) to get the number of reported cases from the number of true cases. The mean number of reported cases at time  $t_i$  is equal to a reporting probability  $\rho$  times  $C_i$ , and the size parameter is assumed to be equal to  $\frac{1}{\tau^2}$  where  $\tau$  is a parameter that is fitted. The full set of parameters and covariates used for the model is given in Tables S4.2–S4.3.

Table S4.2: Description of parameters and covariates used in fitting both the homogeneous and age-structured models to London data

Symbol	Parameter/Covariate	Fitted/Fixed
$M$	Number of age classes	$M = 1$ (homogeneous) or $M = 3$ (age-structured)
$p(t)$	Time-varying fraction of newborns vaccinated	Computed
$\mu_B(t)$	Time-varying <i>per capita</i> birth rates	Computed
$\mu_d(t)$	Time-varying <i>per capita</i> combined death and immigration rates	Computed
$\sigma$	Incubation rate	$\frac{365}{8} \text{ yr}^{-1}$
$\gamma$	Recovery rate	$\frac{365}{5} \text{ yr}^{-1}$
$\iota$	Constant transmission rate from outside the population	Fitted
$\rho$	Constant reporting probability	Fitted
$\tau$	Overdispersion parameter in reporting	Fitted
$b_1, \dots, b_6$	Coefficients of the seasonal transmission rate.	Fitted

We fit the model to different lengths of data within 1945–1990 on measles for London. The model trajectories were computed using (S4.5)–(S4.8) and parameter fitting was conducted using trajectory matching in the R package `pomp` (King et al., 2015).

A summary of the results, including the maximum likelihood estimates for the parameters, the maximum log-likelihood values and Akaike information criterion (AIC) values are given in Table 2. AIC is calculated using,

$$\text{AIC} = 2(\text{no. of estimated parameter}) - 2 \log(\mathcal{L}) \quad (\text{S4.9})$$

where  $\mathcal{L}$  is the likelihood of the model. From Tables S4.2–S4.3, we see that the homogeneous models in Section 5 have 12 free parameters (six spline coefficients for the seasonal transmission rates,  $\iota$ ,  $\rho$ ,  $\tau$  and three free initial conditions). The age-structured models have nine more parameters than the homogeneous models ( $\beta_{YY}, \beta_{YC}, \beta_{YA}, \beta_{CA}, \beta_{AA}$  and four additional initial conditions). Model selection is done by choosing the model with lowest AIC.

The maximum likelihood point estimates and confidence intervals of the model parameters are presented in Table 3 in the main text. The confidence intervals were computed using profile likelihood.

Table S4.3: Description of parameters and covariates used in fitting the homogeneous or age-structured models to London data

Symbol	Parameter/Covariate	Fitted/Fixed
<i>Parameters specific to the homogeneous model</i>		
$S_1(0),$ $E_1(0),$ $I_1(0)$	Initial number of susceptible, exposed and infected individuals in the population respectively	Fitted
$R_1(0)$	Initial number of recovered individuals in the population	Computed using $R_1(0) = N(0) - S_1(0) - E_1(0) - I_1(0)$
$\beta_{1,1}(t)$	$\beta_{1,1}(t) = \exp(\sum_{\ell=1}^6 b_{\ell} s_{\ell}(t))$ where $s_{\ell}(t)$ are seasonal B-splines with period of one year.	Computed using $b_1, \dots, b_6$
<i>Parameters specific to the age-structured model</i>		
$\nu_1$	Aging rate from young children to older children	$\frac{1}{4} \text{ yr}^{-1}$
$\nu_2$	Aging rate from older children to adults	$\frac{1}{12} \text{ yr}^{-1}$
$S_1(0)$	Initial number of young children susceptible	$S_Y^* N_Y^* N(0)$
$S_2(0)$	Initial number of older children susceptible	Fitted
$S_3(0)$	Initial number of adults susceptible	$S_A^* N_A^* N(0)$
$E_i(0)$	Initial number of individuals in the exposed compartment of age class $i$ . Each subcompartment of $E_i$ was assumed to have $\frac{E_i(0)}{K}$ individuals (rounded off to the nearest integer).	Fitted
$I_i(0)$	Initial number of individuals in the infected compartment of age class $i$ . Each subcompartment of $I_i$ was assumed to have $\frac{I_i(0)}{K}$ individuals (rounded off to the nearest integer).	Fitted
$R_i(0)$	Initial number of individuals in the recovered compartment of age class $i$ .	Computed using $N_i(0) - S_i(0) - E_i(0) - I_i(0)$
$\beta_{1,1}(t)$	Transmission rate among young children, constant value of $\beta_{YY}$ ,	Fitted
$\beta_{1,2}(t),$ $\beta_{2,1}(t)$	Transmission rate between young and older children, constant value of $\beta_{YC}$	Fitted
$\beta_{1,3}(t),$ $\beta_{3,1}(t)$	Transmission rate between young children and adults, constant value of $\beta_{YA}$	Fitted
$\beta_{2,2}(t)$	$\beta_{2,2}(t) = \exp(\sum_{\ell=1}^6 b_{\ell} s_{\ell}(t))$ where $s_{\ell}(t)$ are seasonal B-splines with period of one year.	Computed using $b_1, \dots, b_6$
$\beta_{2,3}(t),$ $\beta_{3,2}(t)$	Transmission rate between older children and adults, constant value of $\beta_{CA}$	Fitted
$\beta_{3,3}(t)$	Transmission rate among adults, constant value of $\beta_{AA}$	Fitted

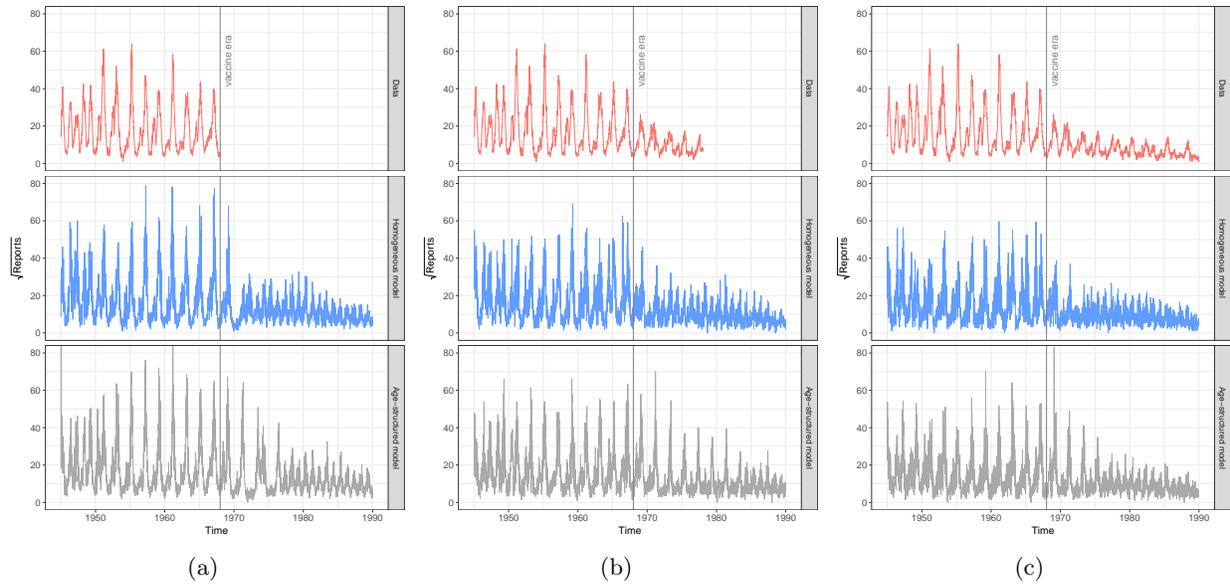
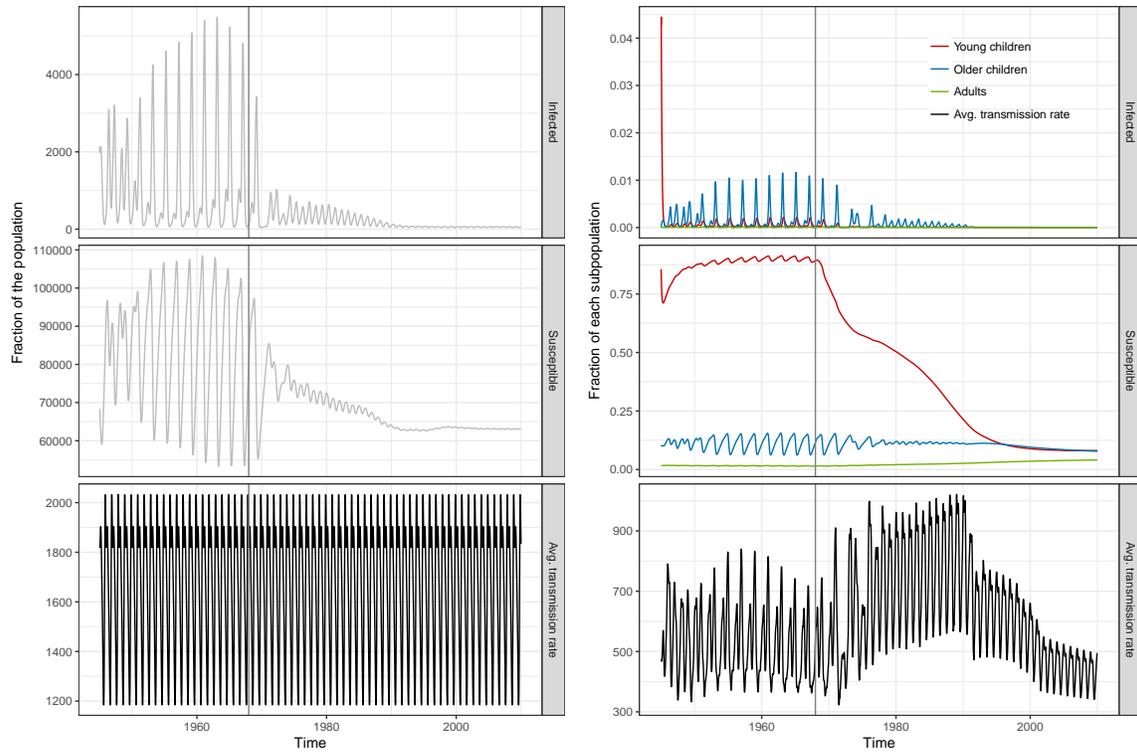
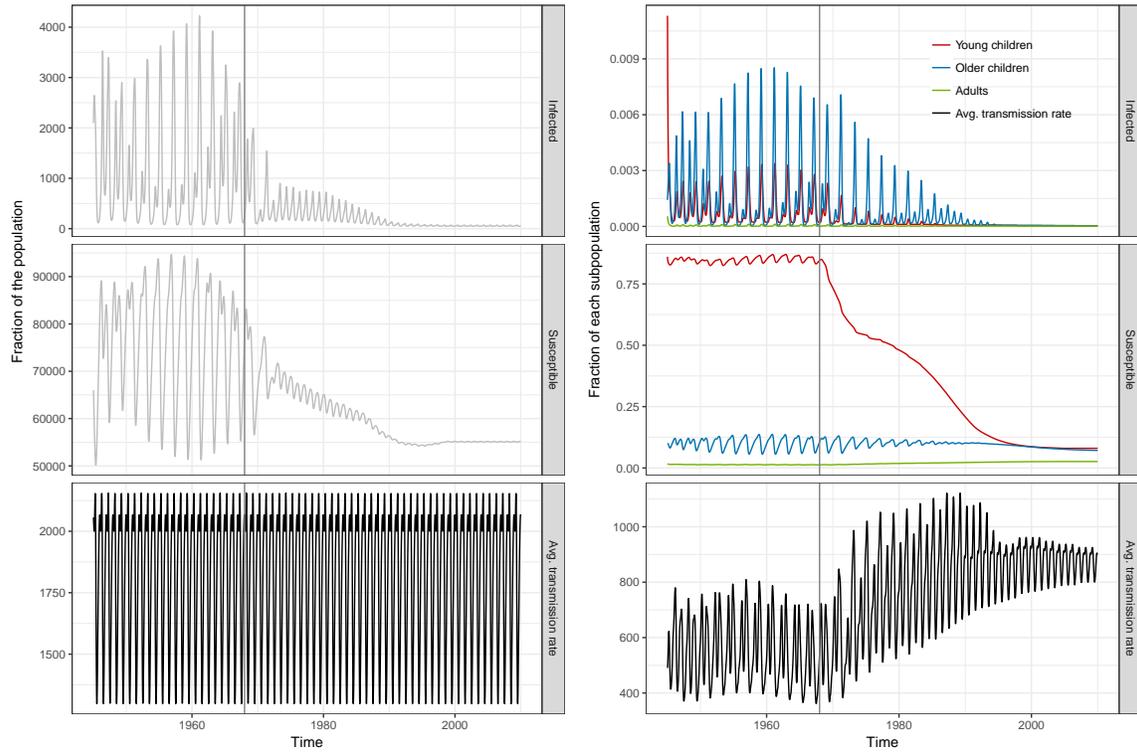


Figure S4.6: (a)–(c) Trajectories of the best fitting model to different lengths of data. The first row is the data. The second row shows sample simulations (deterministic model + measurement noise) for the homogeneous model and the third row shows a sample simulation for the age-structured model.



(a) Homogeneous model fitted to 1945–1968 data

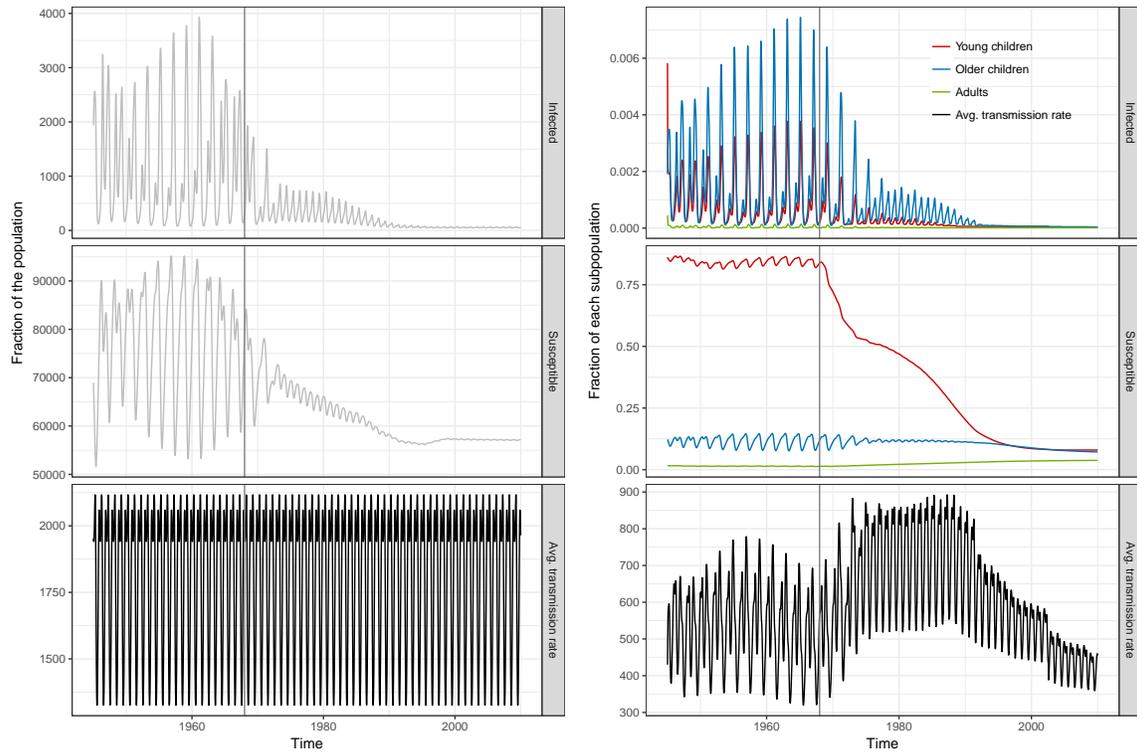
(b) Age-structured model fitted to 1945–1968 data



(c) Homogeneous model fitted to 1945–1978 data

(d) Age-structured model fitted to 1945–1978 data

Figure S4.7: The infected fraction, susceptible fraction and average transmission rate over time in the (a) homogeneous and (b) age-structured models at their maximum likelihood estimate parameter values when fitted to different lengths of data. Simulations of the models are shown until 1990.



(a) Homogeneous model fitted to 1945–1990 data      (b) Age-structured model fitted to 1945–1990 data

Figure S4.8: The infected fraction, susceptible fraction and average transmission rate over time in the (a) homogeneous and (b) age-structured models at their maximum likelihood estimate parameter values when fitted to 1945–1990 data. This is the same as Figure 9 in the main text and is added her for completeness and comparison with Figure S4.7.

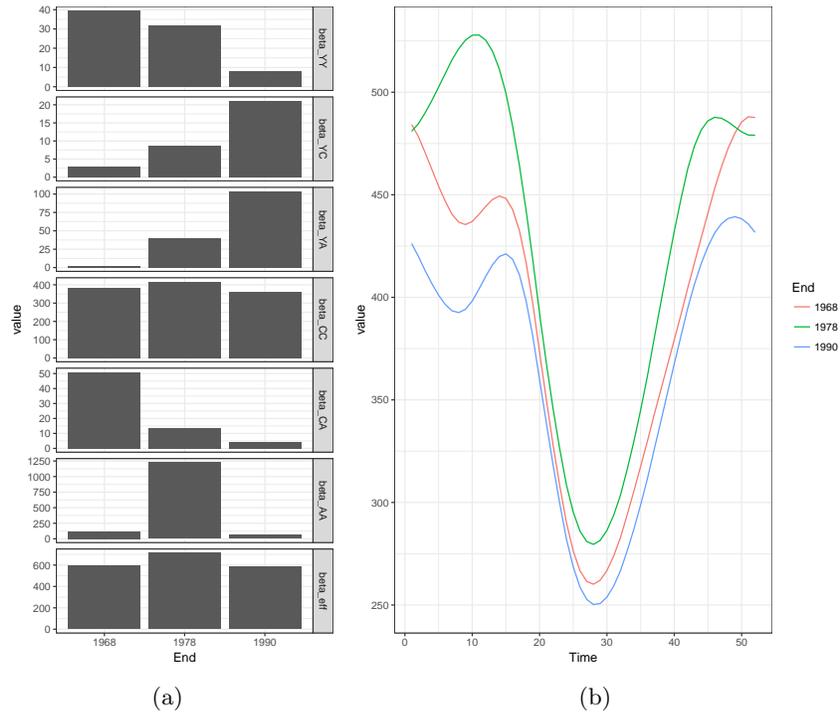


Figure S4.9: (a) Comparison of the best fit transmission rates between different age classes for the age-structured model fitted to 1945–1968 data, 1945–1978 and 1945–1990 data. Here betaCC represents the time-averaged transmission rate between older children. (b) The seasonal transmission rates between older children, plotted over a one-year period.

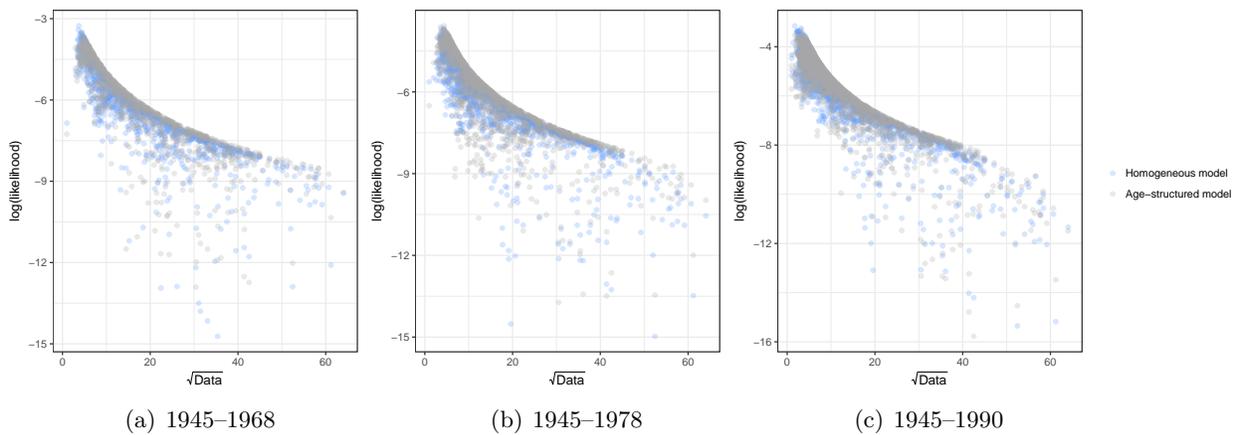


Figure S4.10: Comparison of the log-likelihoods corresponding to each data point (number of reported measles cases each week) for each model fitted to data from 1945–1990. We note that (c) is the same as Figure 7 in the main text and is included here for completeness.

## S5 Estimates of the basic reproduction numbers for measles

Here we present the basic reproduction numbers for the measles models that we fitted to the London data. These were all computed using the next-generation matrix method (Heesterbeek and Diekmann, 2000; van den Driessche and Watmough, 2008). To simplify the calculation, we assumed constant transmission rates. For the age-structured model with  $M = 3$ , we define the matrices  $F$  and  $V$  by

$$F = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{\beta_{YY}N_Y^*}{N_Y^*} & 0 & \frac{\beta_{YC}N_Y^*}{N_C^*} & 0 & \frac{\beta_{YA}N_Y^*}{N_A^*} \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{\beta_{YC}N_C^*}{N_Y^*} & 0 & \frac{\beta_{CC}N_C^*}{N_C^*} & 0 & \frac{\beta_{CA}N_C^*}{N_A^*} \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{\beta_{YA}N_A^*}{N_Y^*} & 0 & \frac{\beta_{CA}N_A^*}{N_C^*} & 0 & \frac{\beta_{AA}N_A^*}{N_A^*} \end{bmatrix}$$

and

$$V = \begin{bmatrix} \sigma + \mu + \nu_1 & 0 & 0 & 0 & 0 & 0 \\ -\sigma & \gamma + \mu + \nu_1 & 0 & 0 & 0 & 0 \\ -\nu_1 & 0 & \sigma + \mu + \nu_2 & 0 & 0 & 0 \\ 0 & -\nu_1 & -\sigma & \gamma + \mu + \nu_2 & 0 & 0 \\ 0 & 0 & -\nu_2 & 0 & \sigma + \mu & 0 \\ 0 & 0 & 0 & -\nu_2 & -\sigma & \gamma + \mu \end{bmatrix}$$

Using the maximum likelihood point estimates for each length of data (with  $\beta_{CC}$  being the geometric mean of  $\beta_{CC}(t)$ , same as in Table 3) and assuming  $\mu = \frac{1}{70} \text{ yr}^{-1}$ , the basic reproduction number can be found by computing the spectral radius of the next-generation matrix  $K = FV^{-1}$ ,

$$R_0 = \rho(FV^{-1}).$$

The computed values of  $R_0$  are shown in Table S5.4. We note that we see a big difference between the  $R_0$  values for the homogeneous and age-structured models, as has already been noted in Section S1. The values computed for the homogeneous models are smaller than those found for London in He et al. (2010) where the infectious and latent periods were not fixed. They are closer to the value of 10.2 found by Edmunds et al. (2000) from age-stratified force of infection estimates.

Table S5.4: Basic reproduction numbers at the maximum likelihood parameter estimates of the models fitted to the London data.

Model type	Years of data	Basic reproduction number at MLE
Age-structured	1945–1968	8.6
	1945–1978	26.9
	1945–1990	7.9
Homogeneous	1945–1968	22.6
	1945–1978	24.7
	1945–1990	24.5

## S6 Age-specific forces of infection

In Edmunds et al. (2000), the age-specific forces of infection of measles were estimated using notification data from 1956–1965 in the UK. They estimated  $0.1 \text{ yr}^{-1}$  (for 0–1 year olds),  $0.21$

$\text{yr}^{-1}$  for (2–4 year olds and 11–17 year olds),  $0.48 \text{ yr}^{-1}$  (for 5–10 year olds) and  $0.11 \text{ yr}^{-1}$  for adults (18+).

Although the age-structured model in this paper only had three age classes, it is possible to do a rough comparison of the forces of infection experienced by the three age classes at the MLEs of this model. The forces of infection of the age-structured models at its different MLEs (corresponding to fits using different lengths of data) were computed using (S4.4) and their values from 1956–1965 were averaged and plotted in Figure S6.11. Here we see that the forces of infection experienced by older children is much higher than that of the other age classes. Our models’ estimated values for the older children are also higher than the maximum values computed in Edmunds et al. (2000) while our values for young children are lower. Our estimated force of infection for adults are similar to those found in Edmunds et al. (2000).

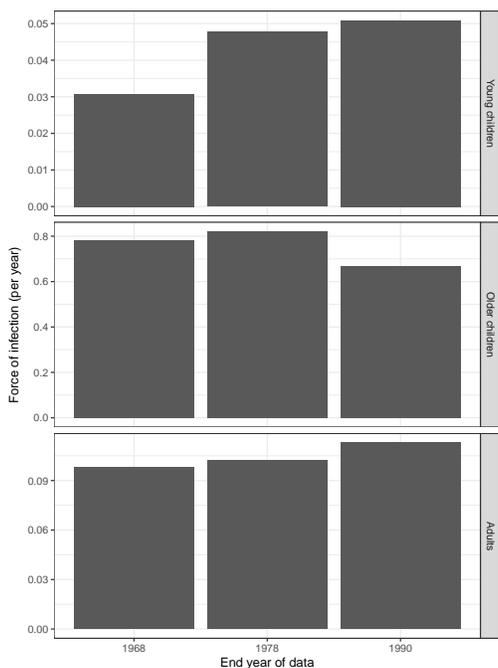


Figure S6.11: Comparison of the average forces of infection from 1956–1965 of the age-structured models at its different MLEs (corresponding to fits to different lengths of data).

## References

- W. J. Edmunds, N. J. Gay, M. Kretzschmar, R. G. Pebody, H. Wachmann, and ESEN Project. European Sero-epidemiology Network. The pre-vaccination epidemiology of measles, mumps and rubella in europe: implications for modelling studies. *Epidemiol. Infect.*, 2000.
- B. Grenfell, O. Bjørnstad, and J. Kappey. Travelling waves and spatial hierarchies in measles epidemics. *Nature*, 414:716–723, 2001. doi: 10.1038/414716a.
- B. T. Grenfell and R. M. Anderson. The estimation of age-related rates of infection from case notifications and serological data. *J. Hyg. (Lond.)*, 95:419–436, 1985.
- D. He, E. L. Ionides, and A. A. King. Plug-and-play inference for disease dynamics: measles in large and small populations as a case study. *J. Royal Soc. Interface*, 7(43):271–283, June 2010. doi: 10.1098/rsif.2009.0151.
- J A P Heesterbeek and O Diekmann. *Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis and Interpretation*. John Wiley & Sons, February 2000.
- M. J. Keeling and P. Rohani. *Modeling Infectious Diseases: In Humans and Animals*. Princeton University Press, 2008.
- A. A. King, E. L. Ionides, C. Bretó, S. P. Ellner, M. J. Ferrari, B. E. Kendall, M. Lavine, D. Nguyen, D. C. Reuman, H. Wearing, and S. N. Wood. *pomp: Statistical Inference for Partially Observed Markov Processes*, 2015. URL <http://pomp.r-forge.r-project.org>. R package, version 0.69-1.
- P. van den Driessche and J. Watmough. Further notes on the basic reproduction number. In F. Brauer, P. van den Driessche, and J. Wu, editors, *Mathematical Epidemiology*, chapter 6, pages 159–178. Springer, 2008.