Supporting Information

Lavine et al. 10.1073/pnas.1014394108

SI Materials and Methods

The Model. The age-specific differential equations are as follows:

$$\frac{dS_a}{dt} = \alpha_{a-1}S_{a-1}(1-\nu_a) - (\alpha_a + \mu + \lambda)S_a + 2\sigma W_{a,n} \quad [S1]$$

$$\frac{dI_a}{dt} = \lambda S_a + \alpha_{a-1}I_{a-1} - (\alpha_a + \mu + \gamma)I_a \quad [S2]$$

$$\frac{dR_{a,1}}{dt} = \kappa\lambda b_a + \gamma I_a + d_a\nu_a + \alpha_{a-1}R_{i-1,1} - (\alpha_a + \mu + 2\sigma)R_{a,1} \quad [S3]$$

$$\frac{dW_{a,1}}{dt} = 2\sigma R_{a,n} + \alpha_{a-1}W_{a-1,1}(1-\nu_a) - (\alpha_a + \mu + 2\sigma + \kappa\lambda)W_{a,1} \quad [S4]$$

$$\frac{dR_{a,i}}{dt} = 2\sigma R_{a,i-1} + \alpha_{a-1}R_{a-1,i}1-\nu_a - (\alpha_a + \mu + 2\sigma)R_{a,i} \quad [S5]$$

$$\frac{dW_{a,i}}{dt} = 2\sigma W_{a,i-1} + \alpha_{a-1}W_{a-1,i}(1-\nu_a) - (\alpha_a + \mu + 2\sigma + \kappa\lambda)W_{a,i}. \quad [S6]$$

The vector α contains the aging rates (set to 2 y^{-1} except $\alpha_{N_a} = 0$).

For integrating the system of ordinary differential equations, we used the R function `lsoda` in the `deSolve` package (1), an integrator that switches automatically between stiff (backward differentiation formula) and nonstiff (multistep predictor-corrector) methods. The Hopf bifurcations (Fig. 4) were computed semianalytically by identifying a point on the curve and then using continuation with local parametrization, tangent prediction, and Newton–Raphson correction (2) in Mathematica 6 (3). The code for all computations, along with a document detailing its use, is available upon request from the authors.

The Data. Some cases were tested at the State Laboratory Institutes, and others were identified at other laboratories or doctors' offices and reported to the Massachusetts Department of Public Health, because pertussis is reportable by law. Three data categories were used to calculate the age distributions: date of birth, date of diagnosis, and age in years. If both date of birth and date of diagnosis were present, the age was calculated from these. If not, the age in years was used.

Parameter Estimation. We used age-specific incidence data from pre-vaccine-era Massachusetts (4) to estimate the boosting coefficient, κ. We assumed that the prevaccine data represent a sample from the equilibrium distribution of the model. At the resulting parameter estimates, the model self-consistently predicts equilibrium dynamics. For the estimation procedure, we additionally assumed: (*i*) Everyone became infected at least once in the prevaccine-era (5), as is supported by serological data. (*ii*) The demographic age distribution was flat. In the estimation we model this by assuming that everyone died at the age of 60 years. The duration of infections, $\frac{1}{\gamma}$, was negligible in comparison with the other waiting times. (*iii*) The pre-vaccination-era age-specific incidence data consist of observations of first infections and second infections of people who have lost immunity induced by the first infection. This is a conservative assumption in that allowing subsequent infections will result in larger estimates of the boosting parameter κ. (*iv*) The age distribution is not strongly affected by the interannual fluctuations in incidence and force of infection present in the prevaccine-era time series data. (*v*) The age-specific incidence data are a representative sample of the age-specific incidence in the entire population; in particular, the probability of observing a case was independent of age. Assuming that there was a lower reporting rate for cases in adults than children in the prevaccine-era, possibly due to reduced disease severity, might lower the estimate for κ. However, the absence of highly symptomatic teenage cases in the prevaccine-era would still put a lower bound on κ. Let the random variables T_1 and T_2 represent the ages of individuals at their first and second infections respectively. Let U_j be a random variable representing the amount of time an individual spends in the susceptible class during his or her j-th visit to that class. By assumption, $U_j \sim \text{Gamma}(a,\lambda a)$; i.e., U_j is Gamma-distributed with shape parameter a and rate parameter λa. Thus U_j has expected value $1/\lambda$ for each j. Times of first infections are simply $T_1 = U_1$. In our model, a second infection can only occur when immunity engendered by the first infection has been lost. Depending on the ambient force of infection, λ, one or more boosting events may have occurred, each of which prolongs the period of immunity. Time to the j-th boosting event is modeled by a random variable $B_j \sim \text{Gamma}(a,\kappa\lambda a)$, time to the j-th waning from class R by another random variable $W_j \sim \text{Gamma}(n,\sigma)$, and time to the j-th waning from class W by the random variable $X_j \sim \text{Gamma}(n,\sigma)$. An individual visiting the W class for the j-th time, therefore, is boosted with probability $\mathbb{P}[B_j < X_j]$ and returns to the susceptible class with probability $\mathbb{P}[X_j < B_j]$. Because, in the model, boosting resets the immunity to its immediately post-infection or postvaccination state, the number, K, of boosting events an individual will undergo is a geometric random variable $K \sim \text{Geometric}(\mathbb{P}[B_j < X_j])$. U_j, B_j, and K are all dependent on the force of infection, λ, and therefore on the number of infections at any given time. When the model predicts equilibrium dynamics, the force of infection is predicted to be constant through time and the subsequent distribution of ages at second infection, T_2 is given by

$$T_2 = U_1 + W_1 + \sum_{j=1}^{K}((B_j|B_j < X_j) + W_{j+1}) + (X_{K+1}|X_{K+1} < B_{K+1}) + U_2.$$

Note that the terms in the randomly stopped sum are independent and identically distributed random variables. Now, for any continuous random variable Z, let $f_Z(z)$ be the probability density function of Z and $G_Z(\omega) = \mathbb{E}[e^{i\omega Z}]$ be the characteristic function of Z. Also, let

$$H(t) = \mathbb{E}[t^K] = \frac{\mathbb{P}[W < B]}{1 - t\mathbb{P}[B < W]}$$

denote the probability generating function of the geometric random variable K. Notice that we have dropped the subscripts from B and W; this introduces no ambiguity because these random variables are independent and identically distributed. It is then elementary to show that

$$G_{T_1}(\omega) = G_U(\omega), \qquad \text{and}$$
$$G_{T_2}(\omega) = G_U(\omega)G_W(\omega)H(G_{B|B<W}(\omega)G_W(\omega))G_{W|W<B}(\omega)G_U(\omega).$$

Each factor of these equations is easily computed using the facts that, when $Y \sim \text{Gamma}(a,c)$ and $Z \sim \text{Gamma}(b,d)$,

$$\mathbb{P}[Y < Z] = \mathbb{P}\left[Q < \frac{c}{c+d}\right], \qquad \text{where } Q \sim \text{Beta}(a,b),$$

and the conditional probability density function

$$f_{A|A<B}(t) = \frac{\mathbb{P}[Z > t]f_Y(t)}{\mathbb{P}[Y < Z]}.$$

For these calculations, we discretized time into two-wk intervals. We computed the model-predicted distributions via the characteristic functions G using the discrete Fourier transform. We calculated the likelihood of each proposed set of parameters by summing the probability densities over the age categories corresponding to the data. Age class width varied from one to five years in the prevaccine-era data. We maximized the likelihood of the observed age distribution using the Nelder–Mead optimization algorithm implemented in the *optim* function in the program R version 2.10.1 (6). We generated a likelihood profile over κ and the mean duration of immunity (Fig. S3), with the rate and shape of the force of infection, λ, and the shape parameter for the loss of immunity distribution, n, as nuisance parameters. Both shape parameters were constrained by a lower bound at 1 (i.e., exponential distributions were the limit). We calculated a 99% confidence interval for the parameter estimates using the likelihood ratio test. The average force of infection was estimated at 0.2 y^{-1} (corresponding to a Gamma distribution with shape parameter 1.8 and rate 0.37) that accords with prevaccine-era estimates from other locations (7). The shape parameter for loss of immunity, n, was not well identified and varied widely for parameters inside the 99% confidence interval. We show that the bifurcations are qualitatively the same for a variety of values of n (Fig. S4).

We did not use this method to estimate parameters from the current-era data because, in the presence of high vaccine coverage, the model predicts cyclic dynamics and therefore a time-fluctuating force of infection, violating the assumptions of our method. We instead show an example of the resultant age distributions from simulations from the age-specified dynamic model (Fig. 3).

1. Soetaert K, Petzoldt T, Setzer RW (2009) deSolve: General solvers for initial value problems of ordinary differential equations (ODE), partial differential equations (PDE) and differential algebraic equations (DAE)R package v 13.
2. Seydel R (1994) *Practical Bifurcation and Stability Analysis* (Springer, New York), 2nd Ed.
3. Wolfram Research, Inc. (2007) *Mathematica*, v 6.0 (Wolfram Research, Champaign, IL).
4. Gordon JE, Hood RI (1951) Whooping cough and its epidemiological anomalies. *Am J Med Sci* 222:333–61.
5. Anderson RM, May RM (1991) *Infectious Diseases of Humans: Dynamics and Control* (Oxford University Press, Oxford).
6. R Development Core Team (2009) *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna).
7. Grenfell BT, Anderson RM (1985) The estimation of age-related rates of infection from case notifications and serological data. *J Hyg-Lond* 95:419–436.

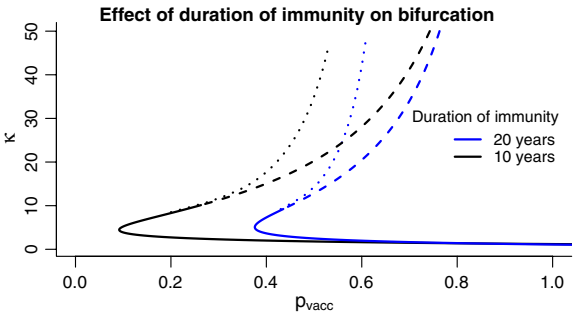

Fig. S1. Analogous to Fig. 4*A* but with an additional curve to show the effect of a longer mean duration of immunity.

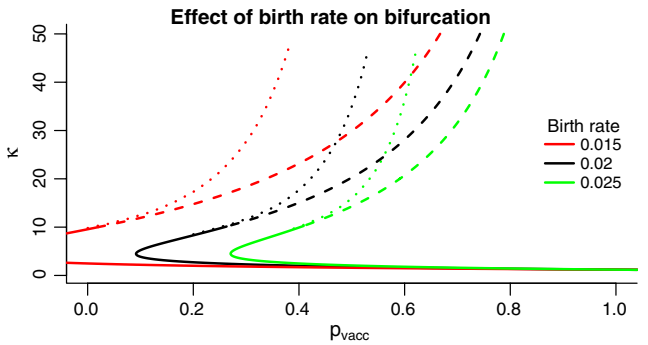

Fig. S2. Analogous to Fig. 4*A* but with two additional curves to show the effect of higher or lower birth rates.

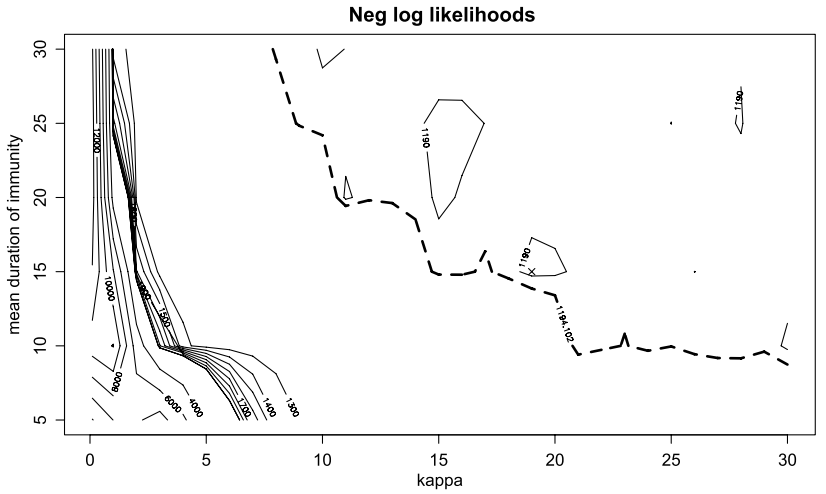

Fig. S3. Likelihood profile over κ and the mean duration of immunity. The "x" identifies the maximum likelihood estimate. The bold dashed line shows the 99% confidence interval.

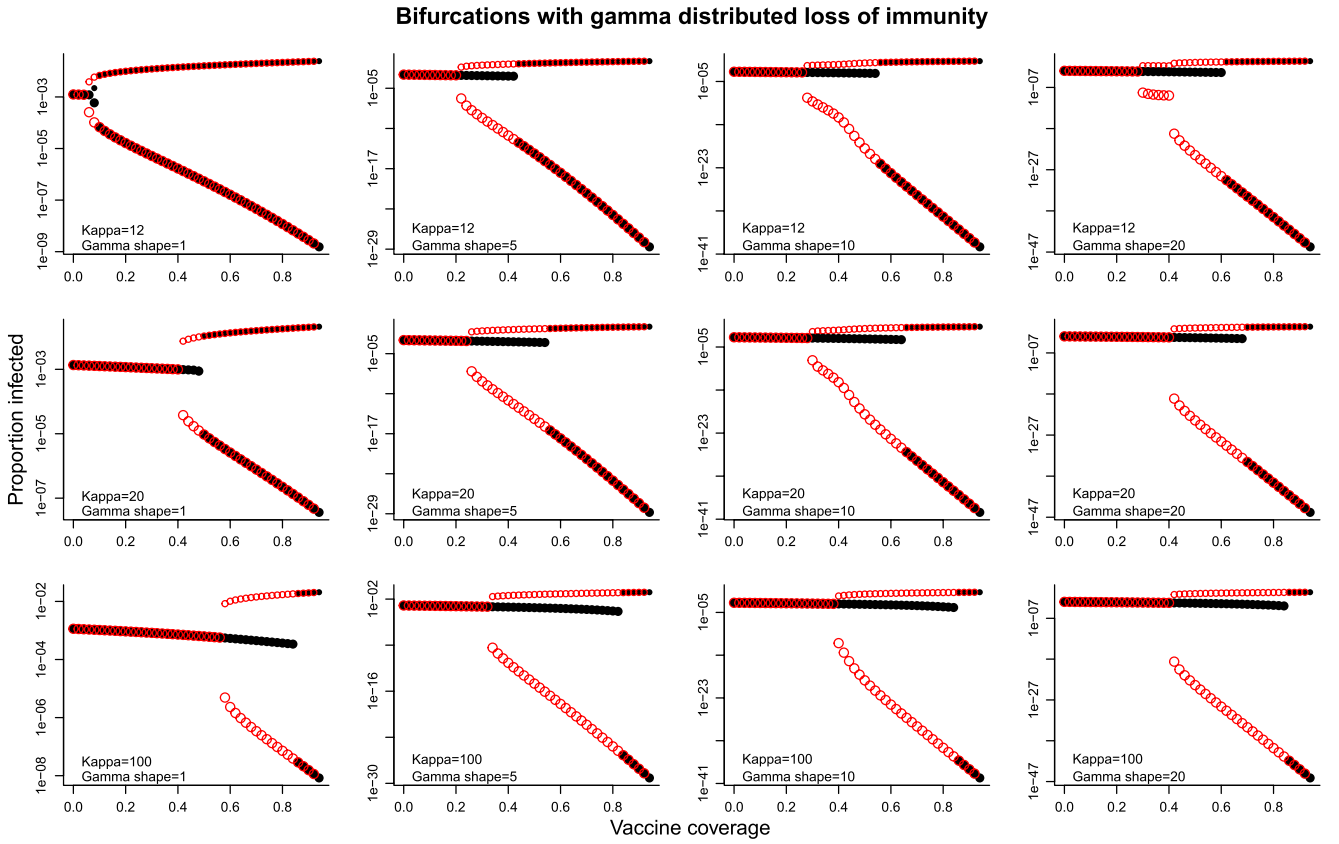

Fig. S4. One-dimensional bifurcations over vaccine coverage for three values of κ and four values of n, the shape parameter for loss of immunity. The solid black dots indicate the incidence with increasing vaccine coverage (forward bifurcation). The open red circles show the same, but as vaccine coverage is slowly decreased (backward bifurcation). The location of the bifurcation shifts to the right for larger shape parameters, but the qualitative results are the same: A stable fixed point gives rise to coexisting equilibrium and cyclic attractors. Then, for very high vaccine coverage, only cycles are predicted.