Epidemic dynamics and antigenic evolution in a single season of influenza A. Authors: M. F. Boni, J. R. Gog, V. Andreasen, M. W. Feldman.

Appendix A: Numerical Simulations. The system (9-13) in the main text was integrated forward in time with starting conditions $i(0) = 1$, $i_k(0) = 0$ for $k > 0$, $I(0) = N^{-1}$, $Q(0) = 1 - N^{-1}$. The tested parameter sets were:

$$\begin{aligned}
\beta &\in \{0.24\ ,\ 0.32\ ,\ 0.40\ ,\ 0.48\ ,\ 0.56\ ,\ 0.64\ ,\ 0.72\ ,\ 0.80\ ,\ 0.88\ ,\ 0.96\ ,\ 1.04\ ,\ 1.12\ ,\ 1.20\} \\
\theta &\in \{0.1\ ,\ 0.2\ ,\ 0.3\ ,\ 0.4\ ,\ 0.5\ ,\ 0.6\ ,\ 0.7\ ,\ 0.8\ ,\ 0.9\} \\
a &\in \{0.01\ ,\ 0.02\ ,\ 0.03\ ,\ 0.06\ ,\ 0.09\ ,\ 0.12\ ,\ 0.15\ ,\ 0.20\ ,\ 0.30\ ,\ 0.40\ ,\ 0.50\} \\
N &\in \{10^3\ ,\ 10^4\ ,\ 10^5\ ,\ 10^6\ ,\ 10^7\ ,\ 10^8\ ,\ 10^9\} \\
\mu &\in \{0.001\ ,\ 0.002\ ,\ 0.003\ ,\ 0.004\ ,\ 0.005\ ,\ 0.006\ ,\ 0.007\ ,\ 0.008\ ,\ 0.009\ , \\
&\qquad 0.01\ ,\ 0.02\ ,\ 0.03\ ,\ 0.04\ ,\ 0.05\},
\end{aligned}$$

and in all simulations $\nu = 0.2$ and $n = 60$.

The above parameter ranges indicate that $1 \leq \mathcal{R}_0 \leq 6$. However, the *effective* basic reproduction ratio of influenza in a host population takes the population immunity into account. In our case, effective $\mathcal{R}_0$ is $\beta(1-\theta)/\nu$. In the above tested parameter ranges 71% of the parameter sets fall inside the range $1 \leq \beta(1-\theta)/\nu \leq 3$, while 44% of the parameter sets fall in the range $1 \leq \beta(1-\theta)/\nu \leq 2$. This is consistent with recent opinion of influenza's low $\mathcal{R}_0$ (Mills et al. 2004). Empirical studies necessarily measure the effective basic reproduction ratio of influenza.

Our parameter ranges yield a total of $13 \times 9 \times 11 \times 7 \times 14 = 126,126$ parameter sets, 84,084 of which had $\mathcal{R}_0$-values greater than one. 90 of these parameter sets had epidemic lengths greater than 1000 days — these parameter sets were discarded. The simulation results can be summarized by a $83,994 \times 10$ matrix where the first five columns are the parameter values $\beta, \theta, a, \log N$, and μ respectively. The last five columns represent dynamic quantities of interest, namely, $1-Q(t_f)$, $1-\theta-S(t_f)$, t_f, D_S, and δ, respectively. We label the columns from 0 to 9, so that columns 0 to 4 correspond to parameters and columns 5 to 9 correspond to dynamic quantities. Partial correlation coefficients between a parameter and a dynamic quantity can be calculated, keeping the remaining four parameters constant. For example, we would denote the partial correlation coefficient between immunity (θ) and epidemic length (t_f) by

$$r_{17\cdot 0234},$$

and we calculate it using the recursions in Cramér (1945) (pp.305-308). The partial correlations mentioned in the last paragraph of the RESULTS section were calculated as

$$r_{59\cdot 01234} = -0.15\ (+0.21) \quad , \quad r_{69\cdot 01234} = -0.06\ (+0.21),$$

$$r_{57\cdot 01234} = -0.61\ (-0.37) \quad , \quad r_{67\cdot 01234} = -0.62\ (-0.36),$$

$$r_{79\cdot 01234} = +0.60\ (+0.25).$$

The parenthetical correlations have outliers (all epidemics longer than 250 days) removed. In all simulations, we also noted whether the covariance curve peaked before the force of infection (I) curve. The covariance curve peaked first in 97% of simulations.

Appendix B: Reverse-sigmoidal cross-immunity function. We consider an alternate cross-immunity function

$$\tau^*(k) = \frac{b+1}{b+e^{ak}} \tag{B1}$$

which is sigmoidal in shape (but reversed) so that the initial part of the cross-immunity curve τ^* is concave, as opposed to convex in the exponential case. To compare τ (from the main text) to τ^* (above), we fix the number of amino-acid changes at which a host retains 50% immunity. For example, in the top graph of Figure B1, the exponential cross-immunity curve, $\tau(k) = e^{-ak}$, has $a = 0.05$; this means that a host retains 50% immunity if the challenging strain is 13.86 substitutions (amino-acid changes) from the immunizing strain. The reverse-sigmoidal (RS) curves marked by the asterisks are likewise chosen for 50% immune-loss after 13.86 substitutions. The curve marked by one asterisk (*) has $a = 0.3$ and $b = 62$, while the curve marked by two asterisks (**) has $a = 0.5$ and $b = 1022$.

The bottom panel in Figure B1 is set up similarly. The unlabeled exponential curve has $a = 0.1$. The less concave curve (*) has $a = 0.3$ and $b = 6$, while the more concave curve (**) has $a = 0.5$ and $b = 30$. In all three curves, a virus escapes 50% immunity after 6.93 amino-acid changes.

Our alternate cross-immunity function τ^* is meant to mimic a situation where the first several substitutions do not benefit the virus with very much immune escape. This is motivated by the fact that an observable amount of immune escape occurs only after the virus has acquired substitutions in at least two different epitopes (Wilson and Cox 1990). If this is indeed the case, then the first several amino-acid changes in the HA1 may not confer much of a fitness benefit on the virus. In the top panel of Figure B1, the first 4 or 5 substitutions in the two τ^*-curves maintain the virus's fitness at more or less the same level as the

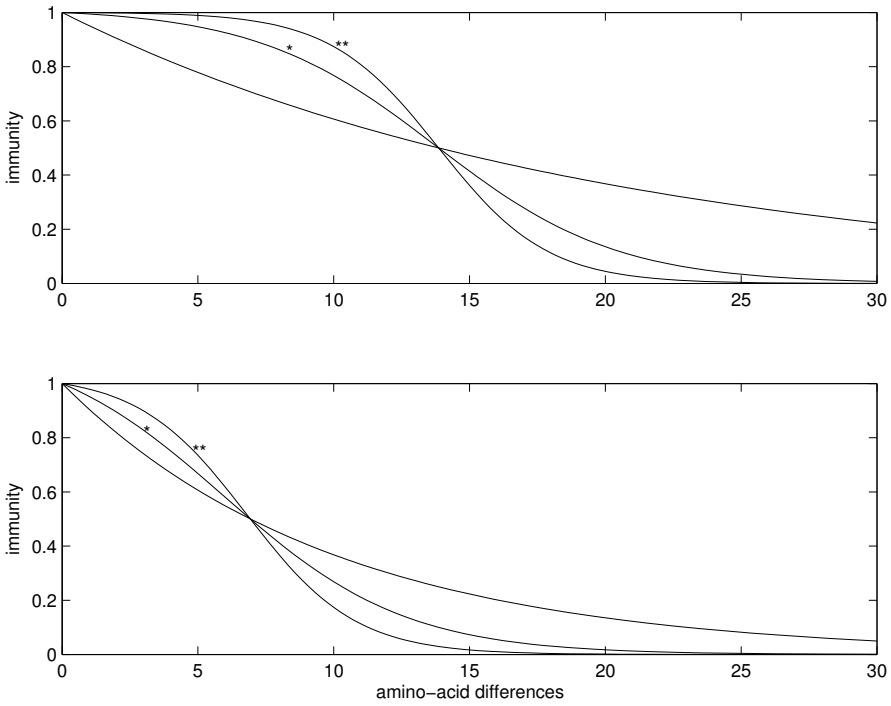

Figure B1: Reverse-sigmoidal cross immunity function as defined by equation (B1)

.

Top panel: exponential τ-curve has $a = 0.05$. RS-curve marked by (*) has $a = 0.3, b = 62$, while RS-curve marked by (**) has $a = 0.5, b = 1022$. All 3 curves exhibit 50% immune loss after 13.86 amino-acid changes. Bottom panel: exponential τ-curve has $a = 0.10$. RS-curve marked by (*) has $a = 0.3, b = 6$, while RS-curve marked by (**) has $a = 0.5, b = 30$. All 3 curves exhibit 50% immune loss after 6.93 amino-acid changes.

epidemic strain, while the first 4 or 5 substitutions under the exponential cross-immunity function (τ) allow the virus to escape more than 20% of the immune response. A similar but dampened effect is seen in the bottom panel of Figure B1.

Though it seems that creating this "hump" in the cross-immunity function would inhibit antigenic drift, this is not always the case. If the host population does not have much immunity, then the virus does not benefit very much from its initial mutations, new variants do not spread more effectively than the original strain, and influenza evolution is slowed. However, if the host population has a significant amount of immunity then the epidemic takes much longer to unfold under an RS-function than under an exponential cross-immunity function, since the flu population needs to get over the "hump" in the RS-function before an epidemic can take off; once the virus population gets over the hump, drift occurs more quickly since marginal fitness benefits of amino-acid substitutions are greater than under the the exponential cross-immunity function.

Figures B2 and B3 illustrate this effect. In the left column of both figures ($\theta = 0.3$) the host population is not very immune, and the epidemic takes off after about ten days regardless of the initial mean fitness of

the virus population. As the concavity of the RS-function is increased, there is less fitness variation in the initial population of mutants, and evolution proceeds more slowly. There is less drift and less excess drift. In the right-hand column of these figures, host-immunity drives the viral population over the hump in the RS-functions, after which the virus population is easily able to escape immunity with additional substitutions. For example, in the top-right panel of Figure B2, an exponential cross-immunity function is used, and the virus escapes 46% of host immunity. But, in the middle panel of the right column, the virus population escapes 93% of host immunity — after the first 10 substitutions, immune escape happens very rapidly.

Two important conclusions can be drawn from this simple numerical experiment with our alternate cross-immunity function τ^*. First, the δ–θ relationship described in the main text may be more extreme than under the standard exponential cross-immunity function. Figure B4 shows not only that immunity can drive antigenic drift, but that the dynamics may exhibit a threshold behavior where the host population becomes a "drift-friendly" environment once it crosses an immune threshold. In Figure B4, the range $0.5 \leq \theta \leq 0.8$, is amenable to much antigenic drift while immunity values outside this range are either too weak to drive antigenic drift or so strong that they prevent epidemics. Second, sensitivity to the mutation rate μ may be quite strong if the true cross-immunity function indeed has a reverse-sigmoidal shape. One of the key determinants in getting over the hump in the τ^*-form of the cross-immunity function is the mutation rate: the more often mutations occur, the quicker the virus population will have variants with enough mutations to escape immunity. In Figure B4, a population-wide immunity $\theta = 0.5$ is the approximate threshold beyond which the epidemic undergoes significant drift; for slower mutation ($\mu = 0.005$), more immunity is needed ($\theta = 0.7$) to significantly drive influenza evolution. If an intermediate range of θ truly drives antigenic drift, an accurate measurement of the neutral mutation rate (μ) would be necessary to determine the lower boundary of this range.

References

Cramér, H. 1945 *Mathematical Methods of Statistics*. Princeton, NJ: Princeton University Press, 9th edn.

Mills, C. E., Robins, J. M. and Lipsitch, M. 2004 Transmissibility of the 1918 pandemic influenza. *Nature*, **432**, 904–906.

Wilson, I. A. and Cox, N. J. 1990 Structural basis of immune recognition of influenza virus hemagglutinin. *Annu. Rev. Immunol.*, **8**, 737–771.

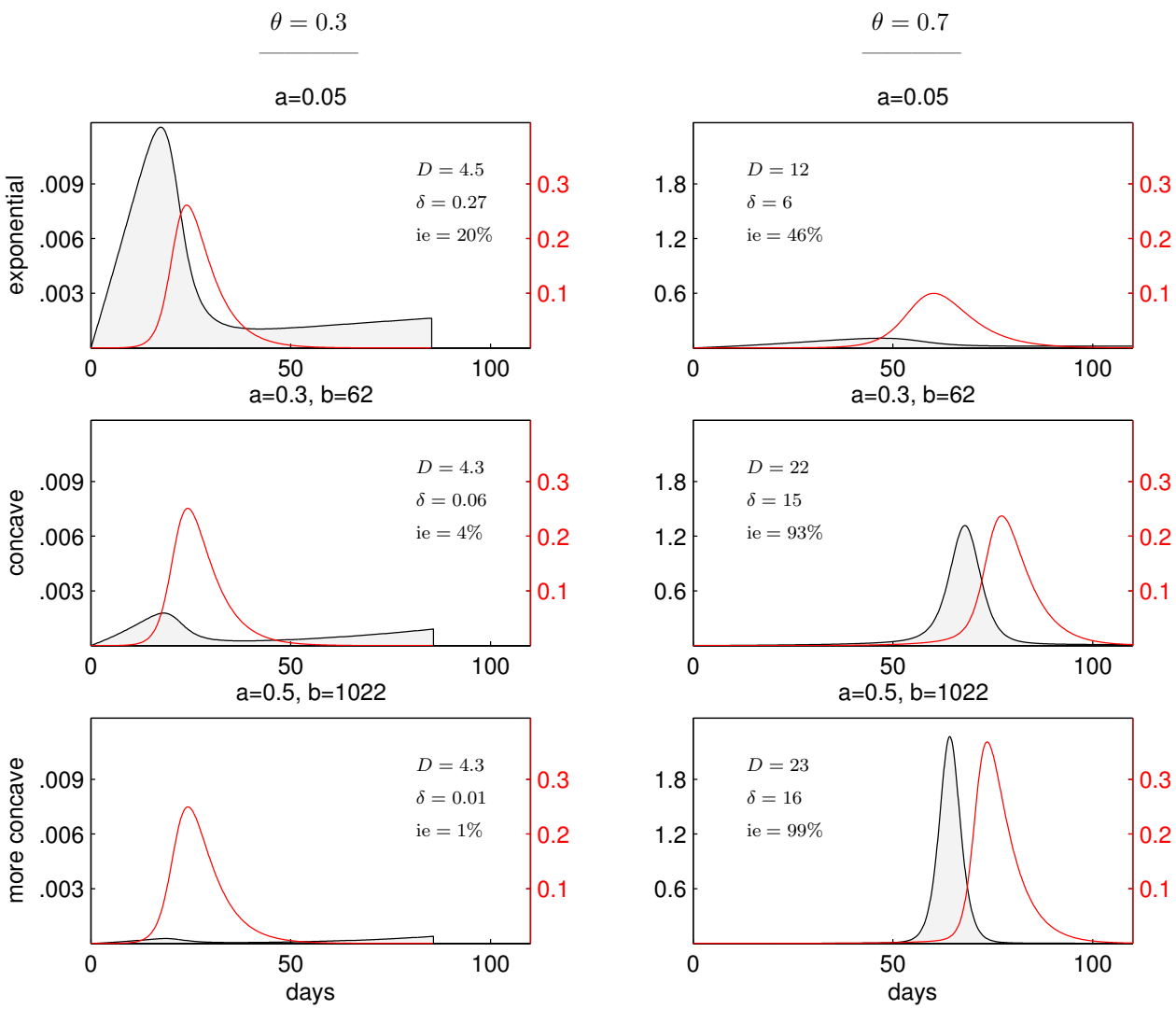

Figure B2: Effects of concavity on excess drift (δ).

In all plots, $\beta = 1.0, \nu = 0.2, \mu = 0.05, N = 10^5, n = 60$. The three plots in the left column have $\theta = 0.3$; the three plots in the right column have $\theta = 0.7$.

D is total antigenic drift, δ is excess antigenic drift, and 'ie' stands for immunity escaped.

Red line denotes the force of infection (I) and corresponds to the right-hand axis (in red). The filled curve is $\beta Q \cdot \text{Cov}$ from the Price equation (16) in the main text; it corresponds to the scale on the left-hand vertical axis. Note that this scale is 200 times greater in the right column than in the left column.

The left-hand column has $\theta = 0.3$, which is not enough immune pressure to drive the virus population over the "hump" in the concave cross-immunity function. The right-hand column has $\theta = 0.7$ which forces the virus population over the cross-immunity hump and subsequently leads to rapid immune escape. The top row uses the exponential cross-immunity function (τ) from the main text with $a = 0.05$. The middle row uses a concave cross-immunity function ($\tau^*, a = 0.3, b = 62$), and the last row uses an even more concave cross-immunity function ($\tau^*, a = 0.5, b = 1022$).

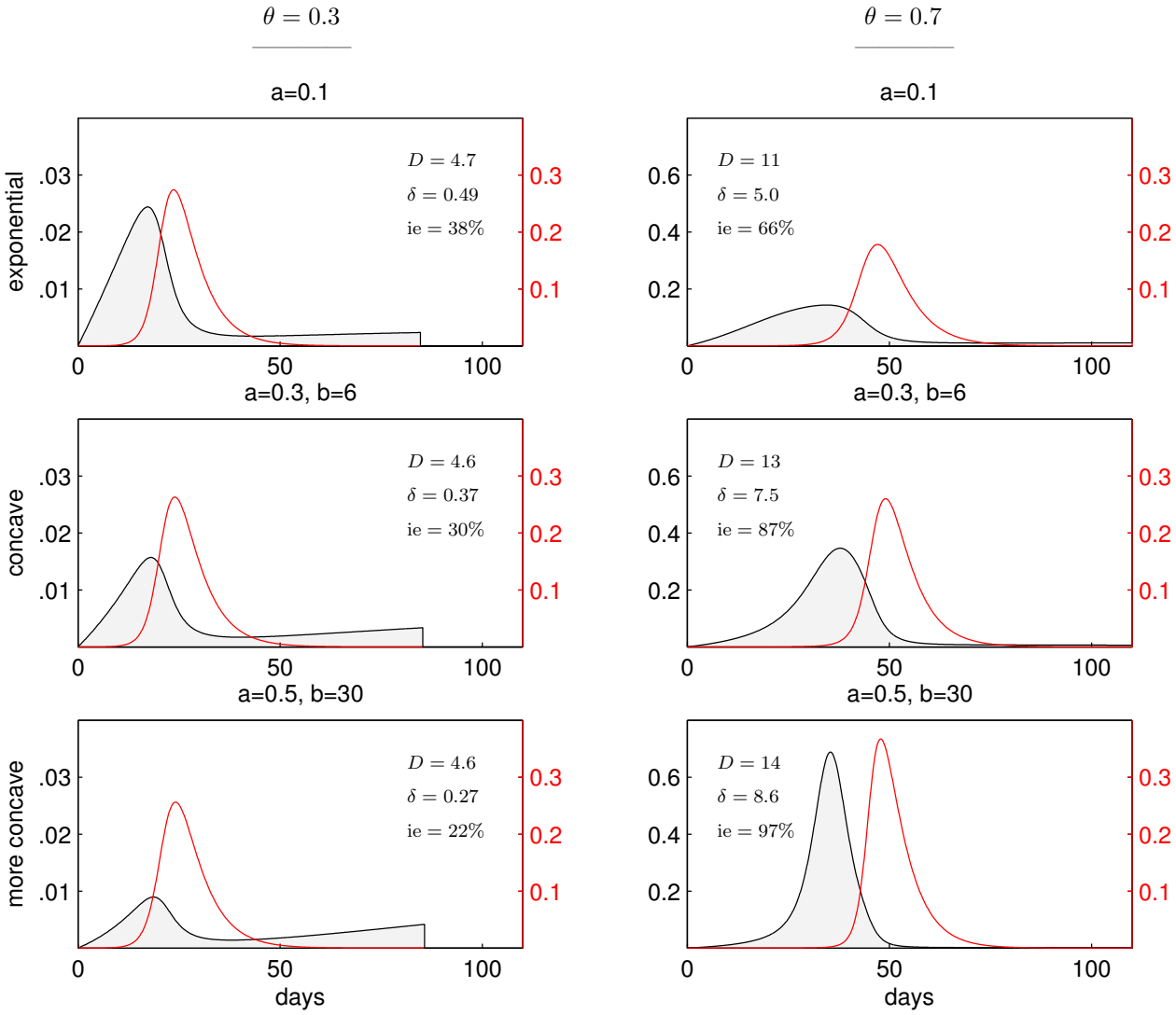

Figure B3: Effects of concavity on excess drift (δ).

In all plots, $\beta = 1.0, \nu = 0.2, \mu = 0.05, N = 10^5, n = 60$. The three plots in the left column have $\theta = 0.3$; the three plots in the right column have $\theta = 0.7$.

D is total antigenic drift, δ is excess antigenic drift, and 'ie' stands for immunity escaped.

Red line denotes the force of infection (I) and corresponds to the right-hand axis (in red). The filled curve is $\beta Q \cdot \text{Cov}$ from the Price equation (16) in the main text; it corresponds to the scale on the left-hand vertical axis. Note that this scale is 20 times greater in the right column than in the left column.

The left-hand column has $\theta = 0.3$, which is not enough immune pressure to drive the virus population over the "hump" in the concave cross-immunity function. The right-hand column has $\theta = 0.7$ which forces the virus population over the cross-immunity hump and subsequently leads to rapid immune escape. The top row uses the exponential cross-immunity function (τ) from the main text with $a = 0.1$. The middle row uses a concave cross-immunity function ($\tau^*, a = 0.3, b = 6$), and the last row uses an even more concave cross-immunity function ($\tau^*, a = 0.5, b = 30$).

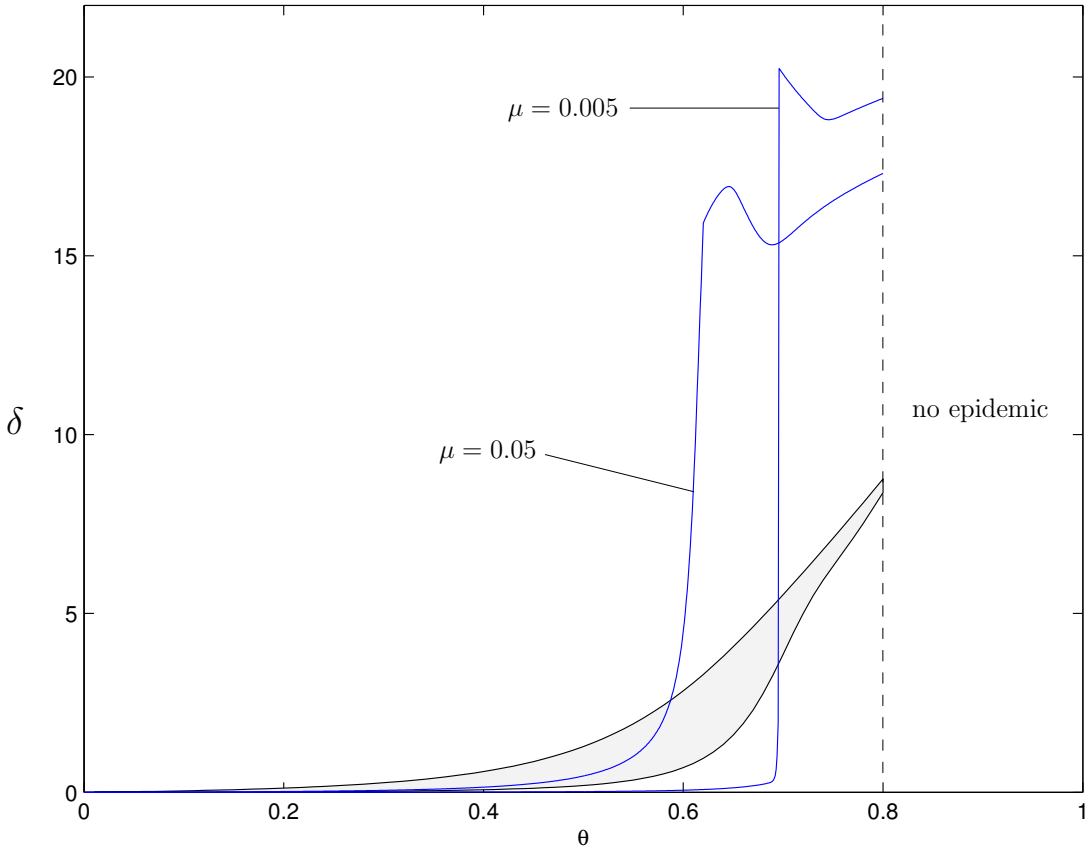

Figure B4: δ–θ graph for the reverse-sigmoidal cross-immunity function.

.

The solid black lines bounding the filled area are from the left panel of Figure 3 of the main text; they show excess drift (δ) as a function of immunity for $\beta = 1.0, \nu = 0.2, a = 0.05, N = 10^5$ with $\mu = 0.05$ and $\mu = 0.005$ corresponding to the top and bottom lines, respectively. Simulations were performed with $n = 60$.

The blue lines show excess drift for the same set of parameter values, but for a concave cross-immunity function τ^*, as defined by equation (B1), with $a = 0.3$ and $b = 62$. The mutation rates for the two curves are labeled in the figure. Simulations for the concave cross-immunity functions were performed with $n = 120$.