S10 Fig. Using a single simulated dataset to compare the exact probability distribution for the current epidemic size given idealized data, as used in our paper, with approximations of this distribution obtained using back-calculation (see Text S4). A. The dataset used for estimating E (the true value of E at the time of estimation is 2). B. Estimation of E using simple back-calculation. As can be seen, simple back-calculation does not discriminate well between different small values of E, which is important in estimating the probability of a major outbreak when the first few symptomatic cases occur. C. Estimation of E using extended back-calculation (with the full $I(t)$ curve observed, and population size and transmission parameters known). Extended back-calculation captures the exact distribution used in our manuscript.