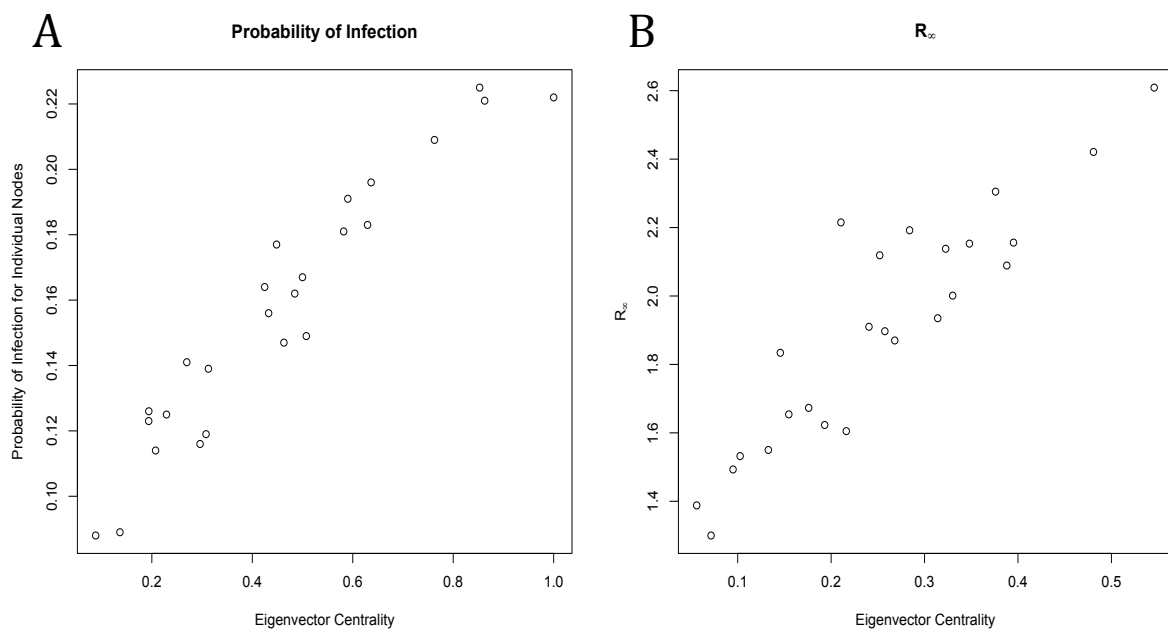


**Supplement D:** Understanding the dependency of eigenvector centralization index ( $C$ ) on the transmission probability ( $\beta$ ) in our simulations.

Within networks, the eigenvector centrality of a node is a strong predictor of the probability that the individual becomes infected in a given simulation, and of the average outbreak size ( $R_\infty$ ) when the infection starts in that individual:



The probability of becoming infected (A) and the average size of the outbreak when the infection starts in a given node (B) is strongly dependent on eigenvector centrality.

Networks with high  $C$  have a greater proportion of nodes with both very high *and* very low centrality scores, as illustrated by Fig. C on the following page.

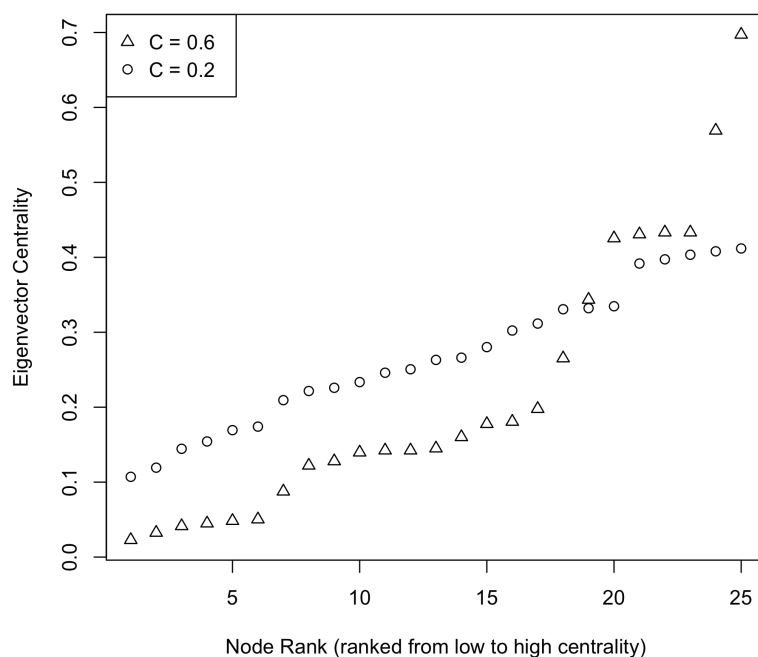
At low values of  $\beta$ , the vast majority of simulations result in small outbreaks. For instance, the median number of individuals infected in simulations when  $\beta = 0.1$  is nearly always 1 or 2. However, in networks with high  $C$ , rare, but larger outbreaks are possible when the

infection begins in a highly central node. The result is a positive effect of  $C$  on  $R_\infty$  at low values of  $\beta$ . At high values of  $\beta$ , the majority of simulations result in the infection spreading throughout most (and often the entire) network. However, nodes with very low centrality scores (i.e. very isolated nodes) are able to evade infections with a reasonably high probability, and when the initial infection begins in very isolated nodes the infection is less likely to spread. The result is a negative effect of  $C$  on  $R_\infty$  at higher values of  $\beta$ .

To summarize, networks with high  $C$  exhibit wider variation in centrality scores across their individual nodes, including a greater proportion of super-central *and* super-isolated nodes than networks with low  $C$ . Thus,  $C$  has a positive effect on  $R_\infty$  when large outbreaks are rare (i.e. low  $\beta$ ), because super-central nodes contribute to rare, larger outbreaks. Conversely,  $C$  has a negative effect on  $R_\infty$  when large outbreaks are common (i.e. high  $\beta$ ), because super-isolated nodes frequently avoid infection and allow for occasional rapid extinction.

## C

**Distribution of Centrality Scores in High C vs. Low C Networks**



We point out that these results fit well with the conclusions of Lloyd-Smith et al. (2005), who conclude that the probability of stochastic extinction rises with increasing heterogeneity in social ties when the average number of secondary infections for the population,  $R_0$ , is greater than 1. The key insight of Lloyd-Smith et al. is that in reality, there is variation in  $R_0$  across the population and this heterogeneity can have a dramatic impact on outbreak dynamics. In agreement with Lloyd-Smith et al., we also find that at higher values of  $R_0$  (which corresponds to higher values of  $\beta$  in our models), increased heterogeneity has a negative impact on the average success of an infectious disease. However, we add to this that when  $R_0$  is low, increased heterogeneity can actually increase the average success of an infectious disease because highly central individuals are able to facilitate occasional larger outbreaks in spite of a pathogen having a low  $\beta$ . Lloyd-Smith et al. may not have observed this effect because they set a relatively high threshold for a spreading event to be considered an “outbreak”, and at lower transmission probabilities this threshold would almost never be reached. In contrast, we did not set a threshold for counting spreading events as outbreaks.

- Lloyd-Smith JO, Schreiber SJ, Kopp PE et al. (2005) Superspreading and the effect of individual variation on disease emergence. *Nature* 438(17): 355-359.