PNAS RESEARCH ARTICLE | SOCIAL SCIENCES OPEN ACCESS

Endogenous competition and the underrealized reproduction of infectious diseases

Peng Huang[a,b], Zack W. Almquist[c,d], and Carter T. Butts[e,f,g,h,1]

Affiliations are included on p. 6.

Edited by Simon Levin, Princeton University, Princeton, NJ; received February 4, 2025; accepted April 29, 2025

Scientific inquiry about the transmissibility of infectious diseases is largely based on the basic reproduction number (R_0) and its derivations. This paper describes a mechanism overlooked in most conventional analyses, in which a disease can endogenously "compete" with itself when multiple infectious individuals race to infect the same susceptible individual, thereby reducing the effective reproductive rate. Utilizing an empirically calibrated network epidemiological model of wild-type COVID-19 diffusion in its early pandemic, we show that the mechanism would be expected to reduce its reproductive rate by an average of 39%. Simulation experiments further identify different types of endogenous competition mechanisms and their relative effect sizes. We highlight the incorporation of endogenous competition mechanism as a necessary step in realistically modeling the reproduction process of infectious diseases.

social network | reproduction number | diffusion | COVID-19

Significance

Our understanding and policy decisions regarding evolving epidemics are often driven by measures of transmissibility, particularly reproduction numbers. By contextualizing the disease diffusion process with multiple reproduction numbers that leverage heterogeneous and clustered social network properties, we identify influential mechanisms that have not yet been incorporated in extant metrics and analyses. Specifically, we demonstrate that a disease can endogenously "compete" with itself when multiple infectious individuals race to infect the same susceptible, undermining its transmissibility. An empirically calibrated model shows that this mechanism reduces the reproductive rate of early-pandemic COVID-19 by an average of 39%. We further characterize and quantify the effect sizes of different variations of endogenous competition mechanisms, paving the way to a more realistic disease reproduction model.

A core goal in the study of infectious diseases is to understand their transmissibility in populations, a question that receives substantial attention from scientists, policy-makers, and the general public—a goal whose importance was underscored by the recent COVID-19 pandemic (1, 2). The starting point for studying disease transmission is typically the basic reproduction number (R_0), which indicates "the number of secondary cases which one case would produce in a completely susceptible population" (3). An insightful and mathematically simple concept, R_0 has been usefully employed to study a wide range of infectious diseases (4–8). Nevertheless, R_0 is a frequently misinterpreted concept (9): It does not inform us about the number of secondary cases that are actually produced by one infectious case in the disease diffusion trajectory, but its expectation in a hypothetical world where all but one individual in the population is susceptible, and where the one infective is selected uniformly at random. In practice, neither condition holds for long (if at all), placing limitations on the appropriate use of R_0 in an evolving outbreak. The gap between R_0 and realized reproductive rates is illustrative of the more general gap between transmissibility as treated in highly stylized, mass-action models and in mechanistic models that can better account for local processes that impact the course of diffusion over time. These latter models demonstrate the presence of often nonobvious mechanisms that can lead to substantial effects on real-world reproductive rates; in this paper, we elucidate several of these mechanisms.

As the above suggests, considerable effort has gone into making the analysis of the reproduction process more "realistic." The literature has focused on two major approaches. The first direction loosens the "everyone-is-susceptible" assumption of R_0, accounting for the fact that the proportion of the susceptible population is a time-varying and space-varying property. This measurement, named the effective reproduction number (R_e), has become the gold standard in assessing the epidemic growth over time in the epidemiological community (1) (though R_0 continues to be widely used elsewhere). A second more ambitious approach is extending reproduction analysis from mass-action or compartment models to network models that account for the discreteness and heterogeneity of real-world contact patterns. Classic compartment models divide the population into various statuses (e.g., susceptible, exposed, infectious, and recovered in SEIR models), with population members of one compartment moving to another at fixed rates; in typical models, compartments are further segmented by demographic and/or geographical features, with contacts among pairs of individuals within and between compartments being treated as uniform random events with constant mixing rates. Like particles in solution, individuals are presumed to enjoy random "collisions," which may (when an infective bumps into a susceptible) lead to transmission. Neither the individuals nor the transmission events are modeled directly, being instead treated as a bulk process.

Author contributions: P.H., Z.W.A., and C.T.B. designed research; P.H., Z.W.A., and C.T.B. performed research; P.H. and C.T.B. analyzed data; and P.H., Z.W.A., and C.T.B. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: buttsc@uci.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2502676122/-/DCSupplemental.

Published June 6, 2025.

This allows for great simplicity and computational efficiency and can work well where populations are large and random mixing holds. In other cases, however, both the number of contacts per individual and the network structure of those contacts can vary widely within the population, creating strong deviations from bulk mixing models. Moreover, when contact networks are persistent on the timescale of infection, local "exhaustion" of available susceptibles becomes possible (even when others are present in the population). These factors can greatly alter the diffusion process (10–13). Network models account for these effects by explicitly modeling individuals within the population and their pattern of contacts, allowing for both heterogeneity and exhaustion (12–14).

The use of network models to study disease diffusion has revealed numerous insights that are not evident in a compartmental view [e.g., the role of partnership concurrency in HIV diffusion (15, 16)]. Here, we examine a mechanism—endogenous competition—whose implications have to our knowledge not been fully characterized in existing work on disease transmission. The basic notion is simple: Endogenous competition occurs when multiple infectious hosts have the opportunity to transmit a pathogen to the same susceptible individual (thus forcing the disease to "compete" with itself for the transmission opportunity). Intuitively, we may suspect that the strength of reproduction will be reduced by the level of endogenous competition—which can vary over time and space, and across subpopulations. Further, it seems evident that measurements that do not adjust for endogenous competition (e.g., R_0 and R_e) will overstate transmission rates in real-world settings. As we show, both intuitions are correct. Further, multiple types of endogenous competition exist, with different relative effects on transmission rates. Accounting for them leads to very different expected diffusion rates (by a factor of approximately 40% in the case studied here) than would be expected under a less explicit treatment.

To examine the mechanisms of endogenous competition, we employ disease diffusion trajectories using a network epidemiological model empirically calibrated for wild-type (WT) SARS-CoV-2 in the early phase of the COVID-19 pandemic (12, 13). The model generates a spatially-embedded contact network among all residents in King County (encompassing Seattle), WA, USA, and then simulates SARS-CoV-2 diffusion in the network based on biometrics of the wild-type virus and calibrated by reported death cases. We then calculate a series of reproduction numbers from the basic R_0 to the actual number of realized secondary cases based on the diffusion trajectory. Comparing these numbers reveals how multiple mechanisms—including endogenous competition—influence the transmissibility of SARS-CoV-2. Expanding on this, we design and test metrics based on contact network topology to explain the trend and scale of these effects. Last, we utilize a simulation experiment of ideal-type network structures to compare two types of endogenous competition mechanisms. Analyses show that endogenous competition reduces the transmissibility of SARS-CoV-2 on average by 39% versus a naive baseline, and this impact is heterogeneous over time in contact networks. As a result, the realized reproduction number (R_r) is significantly lower than those described by R_0 and R_e, and R_r actually stays around 1 for the majority of the trajectory. We conclude by discussing the implications of the disease diffusion process, reflecting on existing metrics and methods, and suggesting directions for future data collection and empirical analysis work on infectious disease transmissibility.

Reproduction Metrics

To bridge the stylized model with a realistic disease diffusion process, we define the following four reproduction numbers, moving from the simplest basic reproduction number to the realized reproduction number, expressing each in network terms (Fig. 1). With the exception of the basic reproduction number, each of these quantities can be defined at the level of an individual infective, or (averaging over infectives) the level of the population. As we show, considerable insight can be obtained by considering reproduction at the individual level.

We begin with the well-known basic reproduction number, described in network terms as

$$R_0 = \bar{d} \cdot p_0, \qquad [1]$$

where $\bar{d}$ is mean degree (the average number of contacts, or in network terms, alters, of an individual, ego, in the contact network), and p_0 is the probability that an infection passes from an infected ego to a susceptible alter; we call this the "basic probability." In the calibrated study case, $\bar{d} = 10.3, p_0 = 26.8\%$, and $R_0 = 2.7$–that is, the average individual has just over 10 contacts, the marginal probability of infecting a contact is approximately 27%, and a randomly infected individual would thus be expected, on average, to produce about 2.7 new cases in a completely susceptible population.

Incorporating network structure and time dimension, the second metric, the dynamic basic reproduction number ($R_d(t)$),

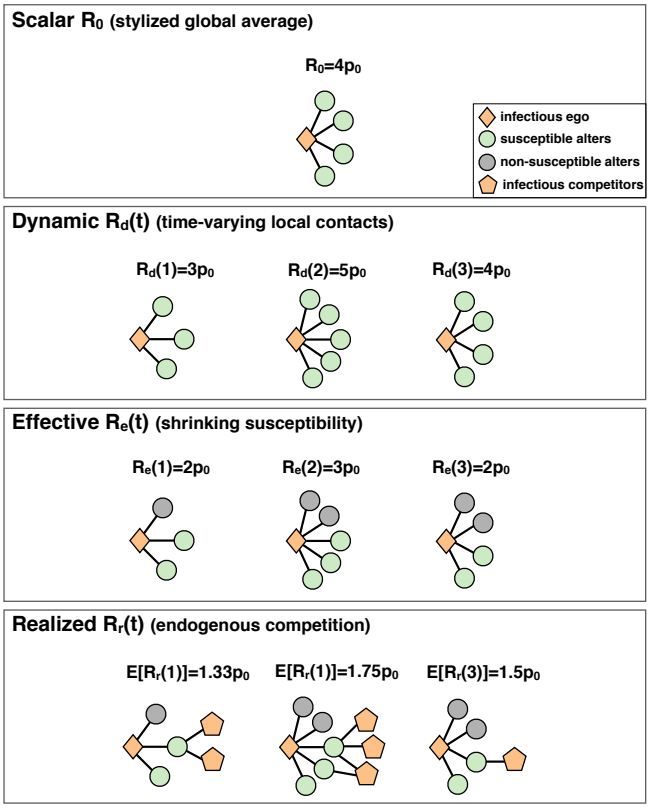

Fig. 1. Schematic illustration of four reproduction metrics and the mechanisms considered. Note that every metric except the realized reproduction number ($R_r(t)$) is an expectation, but $R_r(t)$ is an observed metric subject to diffusion stochasticity. Hence, we can only illustrate the expected value of $R_r(t)$ with the assumption that every infectious individual has the same chance of infecting their alter.

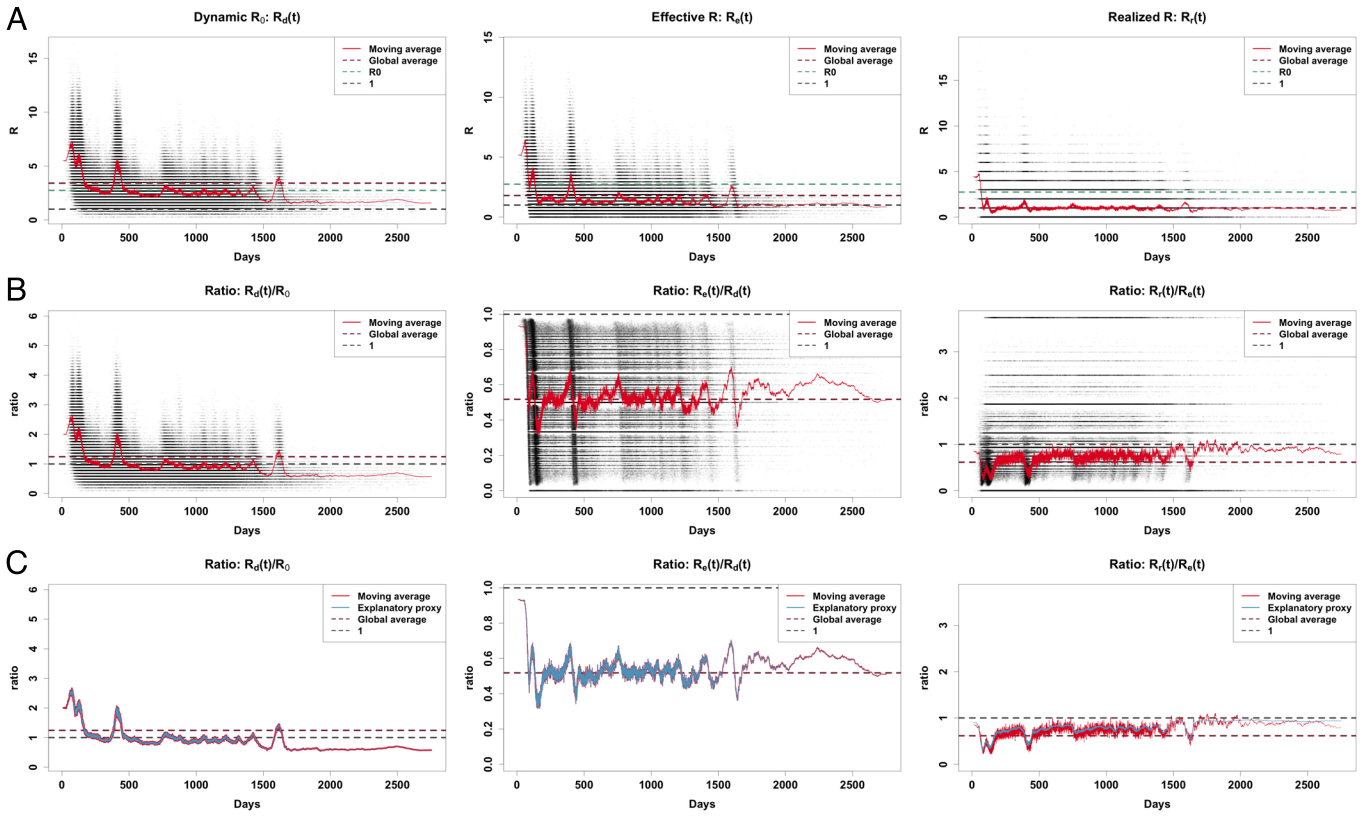

Fig. 2. Reproduction rates in the King County trajectory. Columns from *Left* to *Right* show dynamic basic reproduction numbers, effective reproduction numbers, and realized reproduction numbers; points show individual-level quantities, with population moving averages shown as solid red lines (grand means over all time points shown via dotted red lines). R_0 (green dashed line) and the unit reproductive rate (i.e., 1) are shown as references. Row (*A*) shows the reproduction numbers themselves over the course of the trajectory; note both their heterogeneity within and across time and the systematic deviation of moving averages from R_0. Row (*B*) shows the ratios between successive reproduction numbers (in order of "realism"). Particularly at the individual level, we see large differences between these quantities. Row (*C*) shows line plots of the moving averages of ratios from (*B*), with solid blue curves showing explanatory proxies (see text). In all three cases, the explanatory proxies exactly match or closely approximate the observed ratios.

refers to the expected number of secondary cases produced by an individual infected at time t in a completely susceptible population. While R_0 is a scalar that describes the average transmissibility of disease of the whole system, $R_d(t)$ is a time-varying vector that describes the local transmissibility of disease based on an infected individual's network structure. Letting $d(t)$ be the number of alters of a given individual infected at time t, we have

$$R_d(t) = d(t) \cdot p_0. \quad [2]$$

Averaging this quantity over the set of all individuals infected at a given time provides a dynamic equivalent to R_0.

Loosening the assumption that all but the ego are susceptible, the third metric, the effective reproduction number ($R_e(t)$), has the same definition of $R_d(t)$ except that the population is now partially susceptible. In other words, $R_e(t)$ considers the fact that some individuals are no longer susceptible by the time the ego gets infected. Letting $d_s(t)$ be the number of susceptible alters of a given individual infected at time t, we have

$$R_e(t) = d_s(t) \cdot p_0. \quad [3]$$

As before, this is defined in terms of a particular infective; taking the average over the set of all infectives at time t leads to a population-level property.

The effective reproduction number is the finest-grained measurement of reproduction used by most literature. Yet, $R_e(t)$ still carries a strong assumption that all susceptible individuals will remain susceptible throughout the infectious period of the infected ego. This ignores the possibility that a susceptible alter can be infected (or "scooped") by other infectious individuals before the ego infects them. This is another way of understanding the endogenous competition mechanism: Multiple infected individuals can be in "competition" to infect the same susceptible individual when they are all connected with that susceptible. To examine whether and how much this mechanism reduces the transmissibility of the disease, we will compare $R_e(t)$ with the realized reproduction number $R_r(t)$. $R_r(t)$ is the number of cases that are actually produced by an individual infected at time t. Note that the first three metrics are all theoretical quantities regarding the expected transmissibility under certain assumptions, but the realized reproduction number is an observed quantity, directly derived from the disease diffusion trajectory.* As with the others, the average of the $R_r(t)$ over infectives provides a population-level index.

From Basic Reproduction Number to Its Realization

Fig. 2 shows the simulated reproduction trajectory of COVID-19 in King County using the metrics defined above. A key takeaway is the high level of heterogeneity in transmissibility across all metrics in Panel (*A*), and their deviation from R_0 (dotted green line).

*The spirit of the realized reproduction number is similar to the "actual reproduction number" in prior literature (17, 18), but the actual reproduction number is an estimated number defined in compartment models, while the realized reproduction number here is an observed number that can be directly summarized by examining the diffusion trajectory in network models.

This heterogeneity happens not only at the individual level (as demonstrated by the wide point clouds) but also over time, as reflected by variation in the population average transmission rates (red curves). These highly variable transmission patterns result from both heterogeneity in the underlying contact networks and from endogenous competition (10, 12), properties that are difficult to capture with compartment models and that are obviously omitted when using a scalar such as R_0 to summarize transmission across the whole system.

The Dynamic Reproduction Number Deviates from R_0 Due to Bias in Who Gets Infected. We now consider each transmission metric in turn. We begin with the dynamic R_0, $R_d(t)$ (Fig. 2, first column), representing the hypothetical infection potential among those actually infected. Beyond its heterogeneity, we observe that the global average is actually 24.4% higher than the scalar R_0. This is because not every individual is ultimately infected, and the uninfected do not contribute to the dynamic R_0 at any stage. As expected, individuals who get infected have more contacts, on average, than the whole population. In fact, by taking the ratio between Eqs. **1** and **2**, we can see that the ratio between dynamic R_0 and the scalar R_0 is simply $d(t)/\bar{d}$, the ratio between the degree of the individual infected at time t over the mean degree of the whole population. This explanation of the gap between $R_d(t)$ and R_0 is demonstrated by the *Left*-most panel in row C, which shows $\overline{d(t)/\bar{d}}$ as an explanatory proxy for $\overline{R_d(t)/R_0}$; the coincidence of the two curves visually confirms what is obtained from the ratio of equations above. Substantively, we can see that the dynamic reproduction number is higher when the disease penetrates into subpopulations that are well connected, while it drops when infecting those with fewer contacts.

The Effective Reproduction Number Deviates from $R_d(t)$ Due to Local Exhaustion of Susceptibles. The *Middle* column of Fig. 2 shows the trajectory of the most frequently used time-varying metric, the effective reproduction number ($R_e(t)$). $R_e(t)$ is on average smaller than both the scalar R_0 and the dynamic R_0. As the *Middle* figure shows, on average, the effective reproduction number is roughly half (51.8%) of the dynamic R_0. Taking the ratio of Eqs. **2** and **3**, we can see that $R_e(t)/R_d(t) = d_s(t)/d(t)$: The ratio between effective reproduction number and dynamic R_0 at each time point is the proportion of the contacts to the focal infective that are susceptible at time t. In row (C), we plot the average of this quantity as an explanatory proxy against $\overline{R_e(t)/R_d(t)}$; the coincidence of the two curves confirms the source of the difference in the two metrics. We observe that the ratio between $R_e(t)$ and $R_d(t)$ is relatively high at the beginning of the diffusion, when most individuals are still susceptible, but it quickly drops to fluctuate around its average, going high when the disease penetrates into a mostly susceptible subpopulation, and going low when it spreads "deeper" into a community that has been penetrated. It should be noted that the nearly 50% gap between these metrics does not reflect the global loss of susceptibles in the population (which is much lower through much of the trajectory), but their local exhaustion in the communities through which the disease is actively spreading. Realistic contact networks are far less globally well-connected than random graphs (the discrete analog of mixing in simple compartment models), creating myriad "pockets" whose local infection rates can differ greatly from the population as a whole; when these pockets are consumed, local infection rates slow dramatically, despite the availability of susceptibles elsewhere in the population.

The Realized Reproduction Number Deviates from $R_e(t)$ Due to Endogenous Competition. Perhaps the most striking results lie in the *Right* column of Fig. 2, which shows the realized reproduction numbers ($R_r(t)$) for the King County trajectory. While all metrics examined here share a similar shape for their moving average, we see a further dip from the effective reproduction number to the realized reproduction number. Indeed, $R_r(t)$ fluctuates around 1 most of the time. While this seems counterintuitive, we note that the grand mean of $R_r(t)$ over all time must be equal to 1, because the total number of the cases infected (nominator) is equal to the total number of the cases ever being infectious (denominator) (ignoring a small number of patient zero(s) who were infected from outside the system). Put in substantive terms, when considered over an entire pandemic, the average number of secondary cases produced by each case is almost exactly equal to 1. Of course, this does not rule out periods during which the mean realized reproductive number is greater than 1–or below it. Here, we see a very brief period in which $\overline{R_r(t)} \gg 1$ at the very start of the pandemic, followed by a drop to $\overline{R_r(t)} \approx 1$ with short-term fluctuations over and under this value.

Comparing the effective and the realized reproduction numbers in the third column of row (B) offers further insight: On average, only 61.5% of the transmissibility predicted by the effective reproduction number is actually realized. The ratio varies drastically due to the stochasticity of the diffusion process and the network heterogeneity. Yet, on average, the ratio gets lower when the reproduction numbers peak, suggesting a role for increased competition among infectious individuals when penetrating into populations with a large number of connections: For a susceptible individual, the more connections, the higher the likelihood of having multiple contacts that overlap in their infectious periods. The moving-average ratio gets close to one at the later stages, when the number of actively infectious individuals drops, and the chance of competition lowers.

To unpack whether and how endogenous competition can explain the underrealization of reproduction from the prediction of effective reproduction numbers, we compare an explanatory proxy based on the endogenous competition model. Detailed in *SI Appendix*, this proxy considers the expected probability of infection as a function of the number of competitors an infected individual faces when attempting to infect a susceptible alter. The *Bottom-Right* panel of Fig. 2 shows that, despite the stochasticity contributing to great local variation in the disease diffusion process, the proxy stays close to the average ratio in its moving average, and effectively reproduces its temporal trend. This suggests that the endogenous competition mechanism sufficiently accounts for the discrepancy between the two reproduction numbers. In other words, the endogenous competition mechanism is a mechanism that lowers the transmissibility of infectious diseases but has not yet been captured by existing metrics such as the effective reproduction number; in our study case, failing to account for this mechanism can, on average overestimate disease transmission rates by 62.7%.

Alternative Sources are More Important than Alternative Two-Paths for Endogenous Competition in Diseases Like WT SARS-CoV-2

To further understand endogenous competition, we identify two mechanisms that can contribute to this phenomenon. Panel (*A*) of Fig. 3 illuminates the mechanisms by focusing

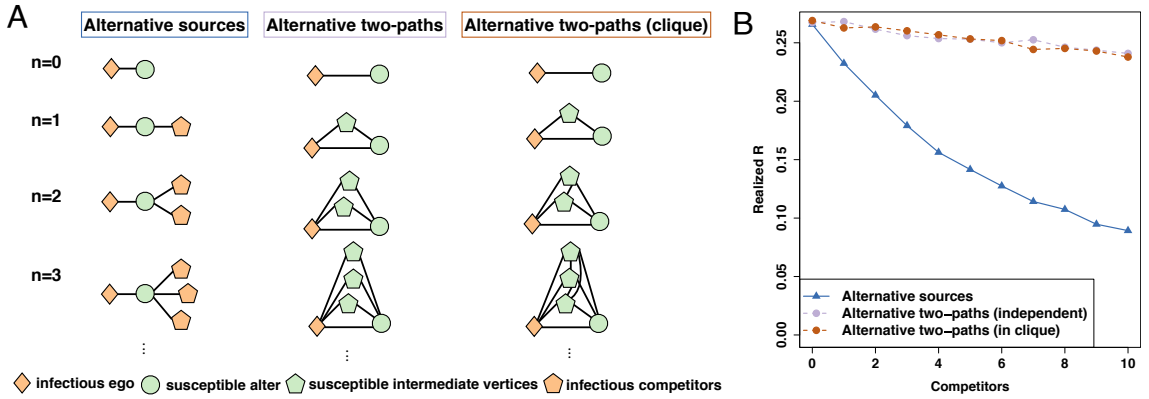

Fig. 3. Comparing two mechanisms behind endogenous competition: alternative sources vs. alternative two-paths. (*A*) Diagrams illuminating the alternative sources mechanism and the alternative two-paths mechanism (including independent two-paths and tow-paths that form a clique), where orange nodes are infectious, green nodes are susceptible, the diamond is the ego, the circle is the alter, and pentagons are competitors. (*B*) Simulation results of realized reproduction on a focal dyad as a function of the number of competitors for different mechanisms. Alternative source competitors play a larger role than alternative two-path competitors.

on the probability of realized reproduction from an infectious ego to a (focal) susceptible alter. The first mechanism of competition is the presence of alternative sources, where the susceptible alter is connected to other infectious individuals, generating direct competition among all infectious individuals connected to the susceptible alter. The second mechanism is alternative two-paths, in which the susceptible alter has no other infectious neighbors, but some of their alters are also connected to the infectious ego. This creates two-paths connecting the infectious ego and the susceptible alter, opening the possibility that one of these other alters will first become infected by ego and then "scoop" them by infecting the focal alter before ego doing so. We show the two most extreme scenarios for this mechanism, where either none of the shared contacts are connected (independent two-paths), or all of them are connected (creating a clique). Panel (*B*) of Fig. 3 shows simulation results (setup detailed in *Materials and Methods*) for the probability of infection of the focal alter by ego. We see that increasing the number of alternative sources substantially reduces the probability that ego will successfully infect the focal alter, while alternative two-paths have a minimal effect. The perhaps surprising weakness of the two-path mechanism can be understood as arising from the fact that having a two-path "scoop" a direct infection requires: 1) infection occurring over two edges, which is an event with base probability $p_0^2 \ll p_0$; and 2) kinetically, the combined time of the two infections must be smaller than the time that would have been taken by the direct infection (had the two-path not intervened). While this suggests that the relative strength of these endogenous competition mechanisms will depend on both the transmissibility of disease (p_0) and its time schedule, alternative sources can be expected to dominate whenever p_0 is not large (as is true of a great many diseases). Certainly, for wild-type SARS-CoV-2, we can see that we are in a regime in which alternative sources can be expected to be the dominant mechanism of endogenous competition.

It should be noted that reproduction numbers are only one of the many tools for characterizing a diffusion process. While the network structure that facilitates endogenous competition can undermine the transmissibility of infectious diseases from a reproduction number view, it may have positive influences in other aspects of the diffusion. We elaborate this point with further simulation analyses in *SI Appendix*.

Discussion

This paper demonstrates the discrepancy between the basic reproduction number and the realized scale of reproduction of infectious diseases and the mechanisms that contribute to this discrepancy. We highlight an overlooked mechanism in the literature, endogenous competition, that can reduce the scale of COVID-19 reproduction (relative to the effective reproductive number) by 39% in an empirically calibrated network epidemiological model. Simulation experiments show that the main contributor of endogenous competition is not alternative paths from an infected to a susceptible but to have multiple simultaneously infected individuals connected to and compete in infecting the same susceptible. *SI Appendix* shows that while different network structures can yield distinct reproduction trajectories, endogenous competition consistently undermines the transmissibility of infectious diseases.

This paper reminds us that reproduction numbers are theoretical values whose interpretations are associated with strong (but often tacit) modeling assumptions; even the effective reproduction number, which is time-varying and considers the shrinking susceptible population, carries assumptions of no endogenous competition, thereby overestimating the transmissibility of disease in realistic settings. Extant literature has posed caveats in comparing R_0 with the threshold 1 to determine the persistence of disease diffusion (19). This caution also applies to more sophisticated reproductive indices: Even if one can get the realized reproduction number R_r, as is shown above, it can stay close to or even below 1 for an extensive amount of time while the disease continues to spread, due to the heterogeneity in the contact structure across individuals and subpopulations.

It is important to bear in mind that simplified metrics are inevitable when trying to make sense of complex phenomena or to communicate their behaviors to nonspecialists, and we continue to find high value in the concept of a "reproduction number." However, we would argue that there is a strong need to caution users of such metrics to fully understand what they tell us at any given time in an epidemic. It is especially important to keep in mind the limits of both the theoretical R_0 and the seemingly more well-founded R_e that in contexts such as epidemic response (where such numbers may be taken as predictive of the evolution of epidemic, and/or used to evaluate interventions or otherwise guide policy). Future work should look to collect

and utilize contact tracing data to empirically evaluate the scale of endogenous competition. Similar to the progress from basic to effective reproduction numbers by considering the shrinking susceptible population, the quantification and incorporation of the endogenous competition mechanism should be the next step in achieving a realistic understanding of the disease reproduction process.

Materials and Methods

Spatially-Embedded Contact Network and Diffusion Model. We adopt the network data and the disease diffusion model from ref. 12. To summarize, the network data contain every individual in King County, WA, USA, based on the 2010 Census, with contacts based on a probabilistic model in which ties are a decaying function of their residential distance. Locations were based on households, placed within their Census blocks using a planar Halton sequence combined with an artificial elevation model (20). The distance between each pair of individual were then measured and employed to predict their probability of having a tie based on empirically calibrated spatial interaction functions employed in prior works (20, 21); a complete description of the modeling procedure can be found in ref. 12. The generated network contains 608,660 individuals, and each individual on average has 10.3 contacts (a.k.a. the mean degree), with SD 7.13.

The spatially-embedded contact network offers a realistic approximation of many aspects of social contact patterns, including realistic mean degree, nonmonotone degree distributions with heavy upper tails, high transitivity, within-household clustering, and spatially biased interaction with power-law tail behavior. To further understand whether and how the reproduction process depends on certain network features, we replicate the diffusion experiment on two corresponding networks, a 2K graph (22) and a Bernoulli graph. The 2K graph has the same population size and the same degree and degree-mixing distributions as the spatial network but is otherwise uniformly random; it thus preserves properties related to first and second-order neighborhood sizes, while randomizing higher-order structure. The Bernoulli graph is a product of further randomization, where only population size and the mean degree is preserved. Results of 2K and Bernoulli graphs are reported in *SI Appendix*.

The disease diffusion model, adopted from ref. 12, uses a continuous-time individual-level SEIR framework, where each individual has four potential states: susceptible (never infected and at risk of being infected), exposed (already infected but not infectious yet), infectious (able to infect their contacts), and recovered (including death). The algorithm starts with five randomly selected patient zeros, who are at the state of exposed, and the rest of the population started at the susceptible state. Based on prior estimates of the wild-type COVID-19 features before implementation of public health measures including vaccination and shelter-in-place rulings, the time from exposed to infectious follows a Gamma distribution with (shape,scale) being ($k_S = 5.807, \theta_S = 0.948$); once becoming infectious, a Bernoulli trial determines the individual's fate of recovery or death, with death probability $p_D = 1.38\%$. The transition to recovery and death follow Gamma distributions of ($k_D = 4.566, \theta_D = 3.984$), and ($K_R = 5.834, \theta_R = 4.566$), respectively. Infection on each edge follows a Poisson process with rate $r = 1/82.87 = 0.012$, based on calibration of King County's COVID-19 death case in the early pandemic (12). The trajectory is completed when no further infections are possible.

Simulation Experiments Comparing Two Endogenous Competition Mechanisms. We first construct a set of networks illustrated in Panel (*A*) of Fig. 3 where the number of competitors ranges from 0 (no competitor but only the ego and the alter) to 10. For the alternative sources mechanism, we let both the ego and the infectious competitor be infected all at time 0, but for alternative two-paths, the only infected individual is the ego. We then run diffusion simulation 50,000 times for each network, and by definition, the realized reproduction number for the focal edge is p_r, the percentage of the times that ego successfully transmits the disease directly to the alter. As shown in panel (*B*), alternative sources provide a much more powerful source of endogenous competition than alternative two-paths.

The alternative path mechanism does not necessarily happen via a two-path; it can be achieved by a longer route, but on average, the longer the path, the less probability and the slower the infection will happen along it. Since the two-path mechanism is already ineffective in competing with direct infection, the longer-paths will be even less competitive (and, in practice, their contribution is negligible).

Data, Materials, and Software Availability. Replication data are deposited in Harvard Dataverse (https://doi.org/10.7910/DVN/PCBWO2) (23).

ACKNOWLEDGMENTS. We thank John R. Hipp for his helpful comments. C.T.B. was funded by the NIH grant 1R01GM144964-01. Z.W.A. was partially funded by a University of Washington Population Health Initiative Tier 3 grant.

Author affiliations: [a]Department of Sociology, University of Georgia, Athens, GA 30602; [b]Center for the Ecology of Infectious Diseases, Odom School of Ecology, University of Georgia, Athens, GA 30602; [c]Department of Sociology, University of Washington, Seattle, WA 98195; [d]Department of Statistics, University of Washington, Seattle, WA 98195; [e]Department of Sociology, University of California, Irvine, CA 92697; [f]Department of Statistics, University of California, Irvine, CA 92697; [g]Department of Computer Science, University of California, Irvine, CA 92697; and [h]Department of Electrical Engineering and Computer Science, University of California, Irvine, CA 92697

1. CDC, Behind the model: CDC.s tools to assess epidemic growth. https://www.cdc.gov/cfa-behind-the-model/php/data-research/rt-estimates/index.html. Accessed 20 May 2025.
2. World Health Organization, "Coronavirus disease 2019 (COVID-19): Situation report" (Tech. Rep. 73, 2020).
3. K. Dietz, The estimation of the basic reproduction number for infectious diseases. *Stat. Methods Med. Res.* **2**, 23–41 (1993).
4. A. Endo *et al*., Heavy-tailed sexual contact networks and monkeypox epidemiology in the global outbreak, 2022. *Science* **378**, 90–94 (2022).
5. R. Li *et al*., Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). *Science* **368**, 489–493 (2020).
6. Q. Li *et al*., Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *N. Engl. J. Med.* **382**, 1199–1207 (2020).
7. C. E. Mills, J. M. Robins, M. Lipsitch, Transmissibility of 1918 pandemic influenza. *Nature* **432**, 904–906 (2004).
8. K. Sun *et al*., Transmission heterogeneities, kinetics, and controllability of SARS-CoV-2. *Science* **371**, eabe2424 (2021).
9. P. L. Delamater, E. J. Street, T. F. Leslie, Y. T. Yang, K. H. Jacobsen, Complexity of the basic reproduction number (R_0). *Emerg. Infect. Dis.* **25**, 1–4 (2019).
10. M. Cevik, S. D. Baral, Networks of SARS-CoV-2 transmission. *Science* **373**, 162–163 (2021).
11. A. Goyal, D. B. Reeves, E. F. Cardozo-Ojeda, J. T. Schiffer, B. T. Mayer, Viral load and contact heterogeneity predict SARS-CoV-2 transmission and super-spreading events. *eLife* **10**, e63537 (2021).
12. L. J. Thomas *et al*., Spatial heterogeneity can lead to substantial local variations in COVID-19 timing and severity. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 24180–24187 (2020).
13. L. J. Thomas *et al*., Geographical patterns of social cohesion drive disparities in early COVID infection hazard. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2121675119 (2022).
14. D. M. Green, I. Z. Kiss, R. R. Kao, Parameterization of individual-based models: Comparisons with deterministic mean-field models. *J. Theor. Biol.* **239**, 289–297 (2006).
15. M. Morris, M. Kretzschmar, Concurrent partnerships and the spread of HIV. *AIDS* **11**, 641 (1997).
16. M. Morris, H. Epstein, M. Wawer, Timing is everything: International variations in historical sexual partnership concurrency and HIV prevalence. *PLoS One* **5**, e14092 (2010).
17. E. J. Amundsen, H. Stigum, J. A. Røttingen, O. O. Aalen, Definition and estimation of an actual reproduction number describing past infectious disease transmission: Application to HIV epidemics among homosexual men in Denmark, Norway and Sweden. *Epidemiol. Infect.* **132**, 1139–1149 (2004).
18. J. M. Gran, L. Wasmuth, E. J. Amundsen, B. H. Lindqvist, O. O. Aalen, Growth rates in epidemic models: Application to a model for HIV/AIDS progression. *Stat. Med.* **27**, 4817–4834 (2008).
19. J. Li, D. Blakeley, R. J. Smith, The failure of R_0. *Comput. Math. Methods Med.* **2011**, 527610 (2011).
20. C. T. Butts, R. M. Acton, J. R. Hipp, N. N. Nagle, Geographical variability and network structure. *Soc. Netw.* **34**, 82–100 (2012).
21. J. R. Hipp, C. T. Butts, R. Acton, N. N. Nagle, A. Boessen, Extrapolative simulation of neighborhood networks based on population spatial distribution: Do they predict crime? *Soc. Netw.* **35**, 614–625 (2013).
22. M. Gjoka, B. Tillman, A. Markopoulou, "Construction of simple graphs with a target joint degree matrix and beyond" in *2015 IEEE Conference on Computer Communications (INFOCOM)* (IEEE, 2015), pp. 1553–1561.
23. P. Huang, Z. W. Almquist, C. T. Butts, Replication data for: Endogenous competition and the under-realized reproduction of infectious diseases. Harvard Dataverse. https://doi.org/10.7910/DVN/PCBWO2. Deposited 9 May 2025.