Diversity in Valuing Social Contact and Risk Tolerance Lead to the Emergence of Homophily in Populations Facing Infectious Threats

Matthew J. Young[1,2 *], Matthew J. Silk[1,2], Alex J. Pritchard[1,2], Nina H. Fefferman[1,2]

March 31, 2022

1 *National Institute for Mathematical and Biological Synthesis (NIMBioS), University of Tennessee, Knoxville, TN, USA*

2 *Department of Ecology and Evolutionary Biology, University of Tennessee, Knoxville, TN, USA*

***** corresponding author

1 Abstract

How self-organization leads to the emergence of structure in social populations remains a fascinating and open question in the study of complex systems. One frequently observed structure that emerges again and again across systems is that of self-similar community, *i.e.*, homophily. We use a game theoretic perspective to explore a case in which individuals choose affiliation partnerships based on only two factors: the value they place on having social contacts, and their risk tolerance for exposure to threat derived from social contact (e.g., infectious disease, threatening ideas, etc.). We show how diversity along just these two influences are sufficient to cause the emergence of self-organizing homophily in the population. We further consider a case in which extrinsic social factors influence the desire to maintain particular social ties, and show the robustness of emergent homophilic patterns to these additional influences. These results demonstrate how observable population-level homophily may arise out of individual behaviors that balance the value of social contacts against the potential risks associated with those contacts. We present and discuss these results in the context of outbreaks of infectious disease in human populations. Complementing the standard narrative about how social division alters epidemiological risk, we here show how epidemiological risk may deepen social divisions in human populations.

2 Introduction

Many studies have considered the impact of individual homophilic behaviors on emergent social structures [1, 2, 3, 4, 5, 6]. Communities in which individuals are more likely to interact with others like them have been discussed widely as both positives (e.g. modularity potentially inhibiting disease spread [7, 8, 9]) and negatives (e.g. echo chambers into which important information may be less able to penetrate [10]). Few studies, however, have considered whether observable homophily in population organizational structure can itself have emerged from more fundamental self-organizing individual-level behaviors. For members of social species, there is a likely benefit from social contacts themselves, but also that benefit may be dampened by costs associated with risk that arises directly from that same contact. One clear example of such systems is infectious disease spreading through social populations.

During a disease outbreak, infection risk can be mitigated through preventative measures [11], such as social distancing, that have been presented as part of a behavioral immune system [12, 13]. Individual adherence to preventative behaviors has been associated with experimental measures of risk aversion [14]. Thus, all else equal, in response to an epidemic, more risk averse individuals can be expected to adhere more to social distancing. By definition, the act of social distancing impedes interactions; a phenomenon that is antithetical with human beings' need-to-belong [15]. Loneliness can have a diversity of health and mental costs [16]; pain associated with social circumstances shares neural pathways with physical pain [17]. This trade-off between the need-to-belong and the behavioral immune system has not gone unrecognized, and has been the subject of several laboratory-based studies [18, 19]. Outside of the laboratory, however, individuals live in a heterogeneous society with distinct predilections towards prioritizing social interactions against their own risk aversion.

Explicit consideration of this social heterogeneity can be explored by incorporating individually distinct strategies of decision-making. Both social isolation and infection have the capacity to cause very real physical and mental maladies; as such, there is unlikely to be a unique, globally optimal solution that accommodates needs across a diverse population. Contingent on risk categories (such as age, immunocompetence, nutrition, or underlying health factors; [20, 21]), individuals can face uncertainty as to the actual risks of contracting a disease and its severity, given those contingencies. This is significant, as risk-averse individuals have been suggested to be intolerant, or aversive, of uncertainty [22, 23, 24]. Importantly, risk assessment and aversion have a deep evolutionary history [25, 26], with evidence for them to be heritable traits [27, 28]. In humans, individual differences, such as the extraversion or sociability dimensions of personality, have been shown to predict scores on scales of risk-taking behavior, inclusive of health behavior [29, 30]. An infectious disease outbreak simultaneously introduces large-scale (often unequally perceived) risk factors and, thus, has the capacity to alter the social dynamics of a community [31, 32].

Risk aversion during a disease outbreak has been posited to be an 'impure public good' [14, 15, 33, 34]: individuals' actions contribute to public health,

but individuals may not have a vested interest in public health (either actual or self-perceived) and are, instead, acting out of self-preservation [33, 34]. Such a perspective emphasizes the importance of understanding how individuals' personal decision-making processes, in the midst of an outbreak, scale to have emergent social consequences. These emergent social consequences could have important implications for the progression of the outbreak while simultaneously altering people's experience of their social world and the benefits that they obtain from socializing. One such emergent consequence might be that individuals preferentially associate with alike individuals, which might be evident via social assortativity, i.e., homophily [6]. Homophilic associations have been demonstrated across numerous axes of individual variation, including shared values or characteristics [6], inclusive of personality [35]. Recently, risk aversion has been recognized as a key component of the structural emergence of homophily [36, 37]. For example Kovářík and van der Leij [37] demonstrate an association between risk aversion and social network structure in an economic context, with risk aversion correlating with local clustering/transitivity in empirical data and theoretical models. This is relevant as transitivity and homophily are often linked in social networks [38]. Subsequent experimental work has emphasized the correlation between homophily and risk aversion, but emphasized the need to explore what generates network homophily [36], and the role of risk aversion in this process.

Here we present a game theoretical approach to understanding how prioritization of individual strategies may shape the social associations of a population. We extend beyond prior research by exploring how balancing risk aversion with socialization can alter association dynamics without imposing a preexisting social structure (i.e., lacking a pre-imposed social network or hierarchy). We implement a socializing game, in which individuals have distinct strategies for prioritizing risk or socialization and demonstrate how these simple dynamics may lead to well-defined emergent structures that have implications for both disease risk and societal function. We then extend this model to incorporate potential pre-existing social relationships between agents to examine how social structure can reorganize based on risk aversion when exposed to extreme exogenous events [37], in this case an infectious disease epidemic.

3 Methods

We define a socializing game during an epidemic with susceptible-infected-susceptible (SIS) dynamics. We assume players are short-sighted and subrational; rather than being perfectly rational agents who can predict the behavior of other players and compute the optimal response, players repeatedly play games and incrementally update their behavior in the direction of the best response to their current environment.

Game Formalism

Let $M >> 0$ be a large constant. At each time t, for each player y, player x decides on a social action strategy $\vec{a}_x(t) = \{a_{x,y}(t) \in [0, M] : y \in G\}$, corresponding to an amount of social contact they want to have with each other player. [1] This social contact requires mutual agreement, so the amount of socializing that actually occurs between players x and y is

$$a'_{x,y}(t) := min(a_{x,y}(t), a_{y,x}(t)) \tag{1}$$

The utility of player x at time t is defined by

$$u_x(t) = \sigma_x f(a_x)\Delta t - \rho_x p(\vec{a'}_x, \Delta t) \tag{2}$$

Where f is a nondecreasing concave function (such as ln) which yields greater utility for higher inputs at diminishing rates, $a_x = \int a'_{x,y} dy$, Δt is the amount of time that passes between each time step, and $p(\vec{a'}_x m, \Delta t)$ is the expected probability of becoming infected as a result of socializing in each time step.[2] σ_x represents the value player x places on socializing. More social players will have higher values of σ_x and place more emphasis on socializing. ρ_x represents the value player x places on health. More health conscious (risk averse) players will have higher values of ρ_x, and place more emphasis on avoiding disease.

Epidemiological Calculations within the Game Formalism

Since the force of infection can be expected to scale proportionally to the amount of social contact among the players [39], then, assuming player x is uninfected,

$$p(\vec{a'}_x, \Delta t) = \Delta t \int \beta a'_{x,y} q_y dy \tag{3}$$

where β be the inherent transmissibility of the disease, and q_y is 1 if player y is infected, and 0 otherwise.

Let γ be the recovery rate of an infected individual from the disease, and assume that recovered individuals return to the susceptible population with no residual protective immunity. Then this defines a Susceptible-Infectious-Susceptible (SIS) model [40] with

[1] This assumes that all players in the population are socially connected to each other. If we relax this assumption by arranging players spatially and setting $a_{x,y} = 0$ for players more than a certain distance apart then, assuming players from each group are equally distributed spatially and so is the initial infection, this will yield identical results to our model. If we don't make those assumptions, or relax this restriction in other ways such as arranging players on a social network graph, then this is likely to create new dynamics, but is beyond the scope of this work.

[2] If f is an unbounded function, this raises the concern of a degenerate case where socialization rates increase infinitely and the cost of near-constant infection is made up for by unbounded gains from socializing. This is prevented via two features of the model. First, players do not change their behavior even when infected, so players behave as if they can be infected repeatedly, disincentivizing players from this case. Second, we bound socialization by a constant M so truly extreme scenario are ruled out directly.

$$\frac{dI}{dt} = -I\gamma + \int (1 - q_x) p(\vec{a'}_x) dx \tag{4}$$

where we explicitly model the contribution to the I class as the expected rate of recruitment determined by the behaviorally-driven force of infection, and $p(\vec{a'}_x) = p(\vec{a'}_x, \Delta t)/\Delta t$.

To ensure that p is a proper probability distribution, we require the maximum socialization $M \leq 1/\beta$, though in practice this limit is not relevant for small β.

Logic of the System

To simplify simulation and analysis of this system, we consider populations of players with σ_x and ρ_x determined randomly from discrete probability distributions. We define $\sigma_l < \sigma_h$, $\rho_l < \rho_h$, and assume each player x has an equal chance of having σ_x equal σ_l or σ_h, and ρ_x equal ρ_l or ρ_h. This effectively partitions the set of all players G into four homogeneous subsets of equal size G_j. Because all players within a group have the same utility function, they will make the same decision regarding $\vec{a}$. When referring to specific groups, we denote them identifiers based on their combined parameters: AS (asocial, $\sigma_j = \sigma_l$), SO (social, $\sigma_j = \sigma_h$), RT (Risk-Taking, $\rho_j = \rho_l$), RA (Risk-Averse, $\rho_j = \rho_h$). Combining these identifiers leads to the four subgroups: ASRT, ASRA, SORT, SORA.

Additionally, we assume that the utility functions of each player is common knowledge. We believe this is a realistic assumption because sociability and risk-averseness (in the context of health outcomes) are discernable personality traits which social contacts tend to know about each other. We also assume players know the frequency of infection within each subgroup, but do not know the infection status of any particular individual. So they can, for instance, choose to socialize more with groups that currently have fewer infected members, but cannot choose to socialize exclusively with uninfected individuals within a group. (Note: this is consistent with unreliable proximate cues of infection, especially for diseases with a pre-symptomatic phase; [41, 42, 43].)

Due to this information restriction, we can rewrite all variables using group-subscripts. The amount of socializing among these groups will be

$$a_{j,k}(t) = |G_k| a_{x,y}(t) \tag{5}$$

$$a'_{j,k}(t) = |G_k| a'_{x,y}(t) \tag{6}$$

where x is any player from G_j, and y is any player from G_k. This allows us to write the SIS dynamics in terms of these subgroups. Let $I_j(t)$ be the fraction of players in G_j who are infected at time t. Then

$$\frac{dI_j}{dt} = -I_j\gamma + \sum_k (1 - I_j) I_k \beta a'_{j,k} \tag{7}$$

and likewise the probability of becoming infected for an uninfected member of group j is

$$p(\vec{a'}_j, \Delta t) = \Delta t \sum_k I_k \beta a'_{j,k} \tag{8}$$

If we let the socialization function $f(a) = ln(a)$, and substitute, the utility function for a member of group j is

$$u_j(t) = \Delta t(\sigma_j ln(\sum_k a'_{j,k}) - \rho_j \sum_k I_k \beta a'_{j,k}) \tag{9}$$

Population Dynamics

Rather than rationally predicting other player's strategies and computing the optimal response to every circumstance, players gradually adapt their strategy over time in response to feedback from games they play (i.e., subrationality). In particular, each player starts with a set of strategies $\vec{a}$ corresponding to their pre-existing social preferences. At each time step, players play one round of the game with each other, and receive payoffs according to the results. Some time Δt passes, and the disease progresses according to the SIS equations and the amount of socialization played in the game. Each player is then informed of the new frequency of infection I_k for each subgroup of players. Each player then updates their social preferences in the direction of best response. In particular, we fix an update speed Δa, then for each group k, player j estimates u_j in the counterfactual situations where their socialization with group k had been $a'_{j,k} + n\Delta a\Delta t$, where $n \in \{-1, 0, 1\}$, and all other socialization levels remained unchanged. The player then increments $a_{j,k}$ by $n\Delta a\Delta t$ according to the n that yields the highest utility. This is computed simultaneously and independently for each k. [3]

This allows players to incrementally update their strategy in response to the changing environment of their peer's behavior and the state of the epidemic. By adjusting the ratio $\frac{\Delta a}{\Delta t}$ we can change the speed of social adaptation relative to the speed of the epidemic.

We assume infected individuals have a latent period in which they are contagious but pre-symptomatic, so they continue socializing at the same rate as uninfected individuals.

[3] Because $a_{j,k}$ does not directly control $a'_{j,k}$ except through mutual agreement via $a_{k,j}$, this can lead to a situation in which a player repeatedly increases $a_{j,k}$, but this yields no change in the actual dynamics. A value of $a_{j,k}$ much greater than $a_{k,j}$ behaves no differently from a value of $a_{j,k}$ that's only slightly greater than $a_{k,j}$ except that it takes longer to update if the incentives change causing player j to start decreasing it. To prevent cases such as this, we fix a small integer $n_c = 4$ and restrict $a_{j,k} \leq a_{k,j} + n_c\Delta a$

Extending the Model to Include Differentially Valued Contacts

Thus far, the model has remained agnostic among social contacts, allowing any interaction to serve equally to fulfill any individual's desire to socialize. To incorporate scenarios in which certain social contacts across groups are of greater personal value than others, we also fix a constant m corresponding to an additional utility gained from each individual's most high-valued relationships and social activities, and an exponential decay term $r \leq 1$. We then incorporate a distribution of utilities across interactions by multiplying the value of each marginal social interaction by $1 + mr^x$ and instead use

$$a_j := c \sum_k \int_0^{a'_{j,k}} 1 + mr^x dx \tag{10}$$

as the term characterizing the social contribution to the utility function, where c is a constant to scale the resulting function.

This simplifies to

$$a_j = c \sum_k \left(a'_{j,k} + \frac{m(1 - r^{a'_{j,k}})}{-ln(r)} \right) \tag{11}$$

when $r \neq 1$, and $a_j = c \sum a'_{j,k}$ when $r = 1$.

Changing the social contribution of the utility function in this way allows for the maintenance of social contacts across groups and the creation of new ones by creating an incentive for players to spread their social interactions among all of the groups. Of course, the benefits of socializing with members of different groups are then balanced against the differing infection exposure that results. (We explore boundary conditions for the system in Appendix 1 [44]. We find that for any set of parameters other than M and Δt, there exist M and Δt such that the observed behaviors of the model are consistent with the results presented in the main text.)

Metrics and Analysis

We record emergent patterns by quantifying social contacts within and across preference groups over time and the coupled outcomes in the prevalence of infection in those groups.

In order to characterize emergent homophily in the context of this composite utility behavior, we fix constants w and τ with the goal of fixing $\int_0^\tau 1 + mr^x dx = w$. For simulations presented here we use $\tau = 1, w = 1.25$. We let r range from 0 to 1 as a free variable, and set $m = \frac{ln(r)(w/r - \tau)}{r^\tau - 1}$. In this way, the area under the curve can be normalized by setting $c = r$. When $r = 1$ this yields a flat line, equivalent to the model with no social differentiation in value among individual contacts. As r decreases, m increases, and the ratio in benefit skews in favor of diversity among contacts across groups (i.e., allowing maintenance of particular

contacts, despite mismatch in axes of socialization and risk tolerance). As r goes to 0, the function approaches an infinitely steep exponential decay, which incentivizes socializing across descriptive groups. We then measure homophily using a modified version of Krackhardt's EI homophily index. [4] In particular, for each group j we define

$$Int_j = a'_{j,j} \tag{12}$$

$$Ext_j = \frac{1}{3}\sum_{k \neq j} a'_{j,k} \tag{13}$$

$$H_j = \frac{Int - Ext}{Int + Ext} \tag{14}$$

Where H_j measures the homophily in the socialization of each group. We then define the overall homophily of the population, H, as the average H_j. When players socializing equally among all groups, we get $H = 0$, as players socialize more in their own groups this increases up to a maximum of 1, and in the hypothetical scenario where players socialized more externally this decreases down to a minimum of -1.

In order to see how the parameter differences between groups impacts homophily, we define an independent variable $s \geq 0$, fix $\sigma_l = 4$, $\rho_l = 30$, and set $\sigma_h = 4 + 16s$, $\rho_h = 30 + 60s$. Thus, when s = 0 all four groups have identical parameters, and thus behave identically. As s increases the risk averse and high social groups become more distinct from the risk tolerant and low social groups.

4 Results and Discussion

Neutral Re-assortment of Social Contacts

In its most extreme expression, individuals were allowed to abandon all existing social contact with individuals who did not maximize their individual joint utility in both socializing and risk tolerance. Figure 1 shows the total socialization of players within each group over time for one instance of the simulation. Under this scenario, $r = 1$ and we see that players settle into an equilibrium level of socializing and infection, and that both levels are higher for groups with higher σ_j and lower ρ_j. The utility gained from socializing is subject to two levels of negative feedback. In the short term, higher levels of socialization receive less marginal utility due to the logarithm, meaning that increasing socialization is of less benefit to the more social groups. In the long term, higher levels of socialization result in higher levels of infection in the population, increasing the costs

[4] The standard EI homophily index uses total number of internal and external network connections, while we define them based on total socialization for our groups. Additionally, we define Ext using the average socialization with each external group rather than the sum, so that we get a score of 0 when players are indifferent among socializing among groups, rather than when precisely half of socialization is internal. Finally, we negate our measure relative to the typical EI homophily score so that perfect homophily yields a score of 1 rather than -1.

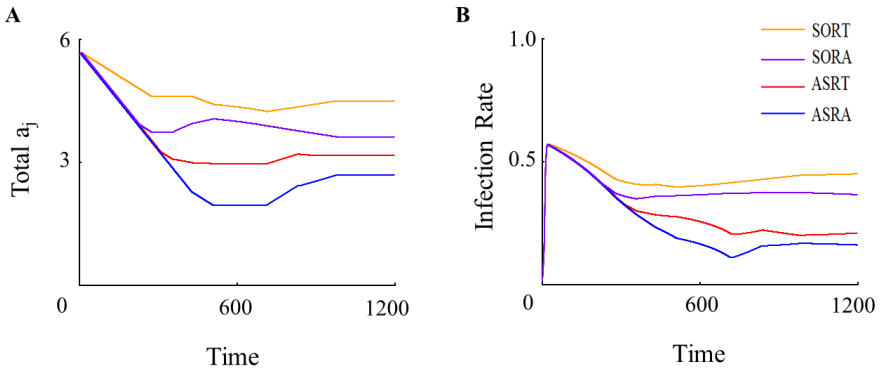

Figure 1: Illustration of the dynamics of the system over time. A. The socialization rates of all groups decline initially before stabilising at different levels, related to B. corresponding changes in the prevalence of infection in different groups. Model parameters are set as $\sigma_l = 4, \sigma_h = 12, \rho_l = 30, \rho_h = 60, r = 1$. Line colors designate the different groups as indicated in the figure. ASRT is asocial risk-tolerant, ASRA is asocial risk-averse, SORT is social risk-tolerant, SORA is social-risk averse.

of socializing. If we take the derivative of u_j with respect to $a'_{j,k}$ and rearrange, we find that

$$\frac{\partial u}{\partial a'_{j,k}} > 0 \text{ iff } \Sigma a'_{j,k} < \frac{\sigma_j}{\rho_j \beta I_k}. \tag{15}$$

Thus we see more socialization in groups with higher σ_j and lower ρ_j.

Simulations of this model repeatedly show the emergence of perfect homophily as time progresses. Each group segregates from the others and individuals only socialize within their own group.

Figures 1 and 2 shows how these social patterns emerge over time. The initial epidemic drives all socialization levels down, which in turn causes the infection to recede. However, when socialization levels recover, gains in socialization primarily occur within each group due to a natural stratification of epidemiological risk.

The differing levels of infection within each group mean that only more social players are willing to tolerate higher levels of infection to socialize more, so they begin increasing a sooner after the initial epidemic, and stably socialize more at equilibrium. As a result, other players are less willing to socialize with those players in particular. Because the positive term in each player's utility function does not vary based on player subgroup, socializing with less-infected subgroups yields the same payoff for a smaller risk.

Therefore, everyone wants to socialize with the safest, least social group - the one with high ρ_j and low σ_j (ASRA). However, everyone in this group also prefers to socialize with players from their own group. Consequently, players

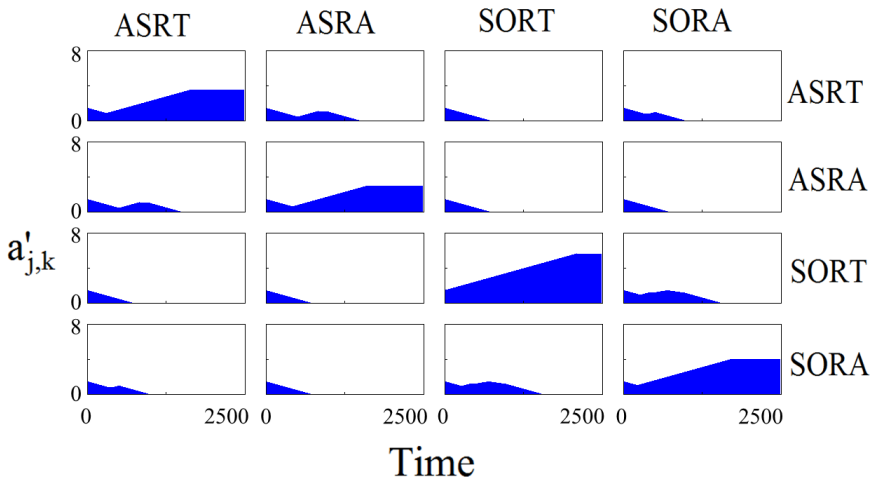

Figure 2: Within and between group socialization rates over time with $r = 1$. Within group socialization rates increase over time until the system stabilises with perfect homophily according to group membership. Model parameters are set as $\sigma_l = 4, \sigma_h = 12, \rho_l = 30, \rho_h = 60$. ASRT is asocial risk-tolerant, ASRA is asocial risk-averse, SORT is social risk-tolerant, SORA is social risk-averse.

from least social and most risk averse group can increase their utility by decreasing all $a_{2,k}$ to 0, where $k \neq 2$, and increasing $a_{2,2}$ to compensate. This lowers their risk while still allowing them to benefit from socializing. And because $a'_{2,2} = a_{2,2}$, they are able to socialize at exactly their preferred level, rendering other socialization with other groups unnecessary.

The continuation of the above mechanism propagated through successively more social and less risk averse groups recursively creates a linear hierarchy in the population. Every other group increases $a_{j,2}$, but this is not reciprocated, so $a'_{j,2} = 0$. The least infected group has effectively removed itself from the population Without the ability to socialize with their preferred group, they have to increase other $a_{j,k}$. The second least infected group (group 1 in the figure) then becomes the preferred social partner of each group, including themselves. As the same process repeats recursively, each group in turn segregates itself from the more infected groups and achieves its equilibrium socialization level internally. Eventually, all groups end up socializing internally, with more infected groups unilaterally being rebuffed in their attempts to socialize more by increasing their socialization with less infected groups ($a_{j,k} > 0$, but $a_{k,j} = 0$).

Retention of Preexisting Social Contacts

The perfect homophily generated above results from a scenario in which individuals drop all of their existing social relationships from outside of their group

and replace them with individuals from their own group. This assumption is unrealistic because individuals value aspects of their relationships other than the risk of infection, or may also be constrained in their choice of associations. This means that relationships are not perfectly replaceable, with each potential partner equivalently able to satisfy the desire for social contact. We therefore also explore an adjusted scenario for our model to account for some relationships being more valuable by increasing the benefits of heterogeneous socialization. In this scenario, we assume that possible social relationships and activities have varying values, and that players will drop the least valuable ones first when possible within a group, meaning the benefits obtained by socializing are limited if a player adheres too closely within any one group.

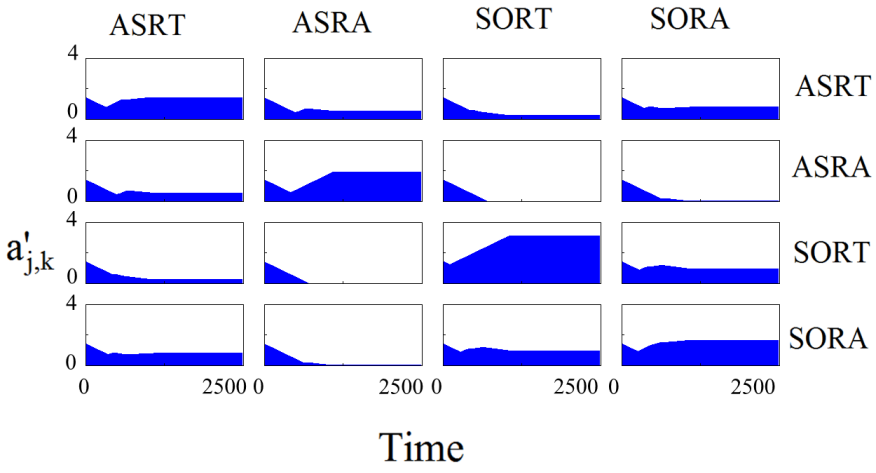

Figure 3: Within and between group socialization rates over time with $r = 1/2$ and $m = 3$. Within group socialization rates increase over time until the system stabilises with partial homophily according to group membership. Other model parameters are set as $\sigma_l = 4, \sigma_h = 12, \rho_l = 30, \rho_h = 60$. ASRT is asocial risk-tolerant, ASRA is asocial risk-averse, SORT is social risk-tolerant, SORA is social risk-averse. Final homophily scores in this figure for each group are 0.479, 0.782, 0.720, 0.482, respectively.]

Figure 3 shows the pairwise socialization between members of each group over time that result from adding these social constraints. We see now only partial homophily emerging. Players still socialize primarily within their own group at equilibrium, but still maintain certain associations with other groups as the benefits from the multiplier outweigh the increased infection risks. Pairs with more similar home-group infection risks still tend to socialize more than dissimilar pairs.

We can vary the value individuals place on particular, irreplaceable contacts by adjusting the steepness of the multiplier function (Fig. 4b). At low values of

r (less steep multiplier function) we see low levels of homophily. However, as we increase r we see that mixing between groups decreases until perfect homophily emerges at values of r close to one, as in as in Fig. 2.

We can also observe that the level of homophily increases as groups become more distinct in their strategies (Fig. 4a). We vary an independent variable $s \geq 0$, fix $\sigma_l = 4$, $\rho_l = 30$, and set $\sigma_h = 4 + 16s$, $\rho_h = 30 + 60s$. We see that homophily increases with s, meaning a population of groups with more disparate strategies will exhibit more homophily.

We can observe this trend more broadly by varying both r and s (Fig. 4c). We see that homophily increases with respect to each of r and s, though r has a larger impact, suggesting that pre-existing social connections are more influential than the intrinsic differences between groups in the extent of homophily that emerges.

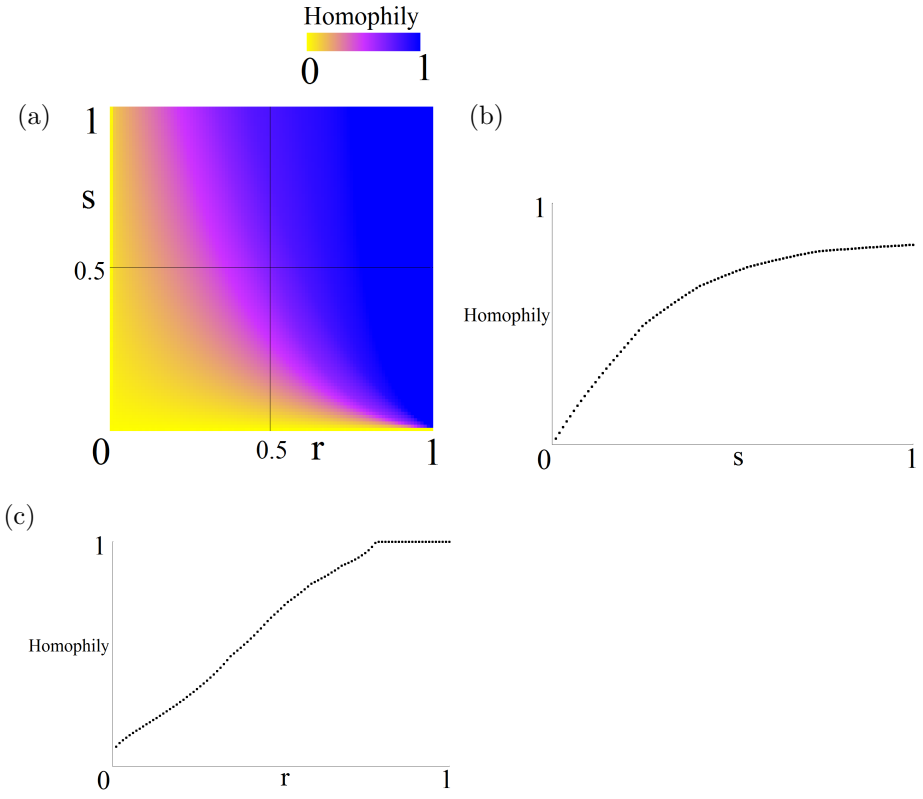

Figure 4: Homophily in socialization rates according to group membership as a function of both r and s (a), a function of s for $r = 0.5$ (b), and as a function of r for $s = 0.5$ (c). Other model parameters are set as $\sigma_l = 4, \sigma_h = 4 + 16s, \rho_l = 30, \rho_h = 30 + 60s$.

Diversity as a Driver of Homophily

Here we demonstrate that the behavioural response of a population to an infectious outbreak can drive the emergence of homophily as a result of individual decision-making strategies. Explaining emergent organizational structures that impact the function and efficiency of populations, but arise solely from individual behaviors, is a fundamental challenge in understanding social systems [45, 46, 47, 48, 49]. The role of infectious disease is of particular interest due to both the successful function of a society (a benefit) and successful transmission of infection (a cost) relying on similar types of contacts and behaviors [50, 51, 52, 53]. Changing our assumptions to assign value to pre-existing relationships lessens, but does not remove, the emergence of homophily. Further, we see that increasing the magnitude of the differences between groups, in their utility functions, increases the level of homophily, while more similar populations are more willing to engage in intergroup interactions. (We provide an analytic explanation for general system behaviors in Appendix 2 [44].)

In our model, the unanticipated emergence of homophilous groups along axes prioritizing risk aversion or social interactions parallels experimental and theoretical work emphasizing risk aversion as a key component of the structural emergence of homophily [36, 37]. Uncertainty of the environment and individual differences in risk aversion have been posited to interact for the generation of network structure [37]. Interestingly, our results are similar, but achieve these similarities without making the previously included assumption of existing social structure. Even in our extended model that places value on existing relationships, homophilous ties emerge in response to a system-wide outbreak when both risk and risk perception are driven by behavior. (For exploration of the impact of relative values of the social and risk-aversion parameters, see Appendix 3 [44].) This shows that large epidemics have the capacity to disturb and reorganize social structure along axes of risk and social preference, as evidenced during the HIV and COVID-19 pandemics [54, 55, 56, 57]. Such population level emergence of assortative mixing has important consequences for anticipating the spread of disease in a heterogeneous society, especially as homophily will also increase modularity and alter the dynamics of the spread of both disease and information [58, 59, 60, 7].

Perhaps most intriguingly, however, our findings suggest that observable emergent community structure may arise from individual differences in categorical preference or assessment of the risks and benefits of social contact over time. While disease is an important example of socially contagious risk, it is certainly not the only one. Cultural norms and the perceived threat of homogenization eroding group identity [61, 62] may similarly act as a driving factor in constructing and maintaining social divisions among groups. Of course, individual game theoretic perspectives are not the only proposed mechanism for the emergence of such structures [63, 64, 65, 66]. However, our results contribute to understanding how simple, individual perceptions and behaviors may yield highly organized, and operationally beneficial, global outcomes.

Implications for Public Health

Emergent structure in society that arises out of individual needs to balance conflicting goals, can provide critical insight into how to influence the individual perceptions and behaviors from which they are formed. In the case of infectious disease outbreaks in human populations, the natural self-organization into homophilous groups offers immediate potential routes for public health intervention. Understanding the independent value propositions that drive community formation can allow the design of strategies that, while marginally less effective in the absolute reduction of infection risk, achieve meaningful reduction without incurring the same social costs. By considering self-organization rooted in multi-factorial utility we can begin to produce useful, quantitative tools to inform policy and improve real-world adoption of mitigation strategies.

Acknowledgements

We gratefully acknowledge helpful discussions with collaborator R.A. Bentley and S. Carrignon, and funding support for this work from NSF DEB 2028710.

Code Availability

Project code is available on the code hosting platform GitHub at https://github.com/kazarraha/SocDistModel

References

[1] Feng Fu, Martin A Nowak, Nicholas A Christakis, and James H Fowler. The evolution of homophily. *Scientific reports*, 2(1):1–6, 2012.

[2] Sergio Currarini, Jesse Matheson, and Fernando Vega-Redondo. A simple model of homophily in social networks. *European Economic Review*, 90: 18–39, 2016.

[3] Jan Treur. Mathematical analysis of the emergence of communities based on coevolution of social contagion and bonding by homophily. *Applied Network Science*, 4(1):1–30, 2019.

[4] Petter Holme and M. E. J. Newman. Nonequilibrium phase transition in the coevolution of networks and opinions. *Phys. Rev. E*, 74:056108, Nov 2006. doi: 10.1103/PhysRevE.74.056108.

[5] Michael Pearson, Christian Steglich, and Tom Snijders. Homophily and assimilation among sport-active adolescent substance users. *Connections*, 27(1):47–63, 2006.

[6] Miller McPherson, Lynn Smith-Lovin, and James M Cook. Birds of a feather: homophily in social networks. *Annual Review of Sociology*, 27: 415–444, 2001.

[7] Pratha Sah, Stephan T Leu, Paul C Cross, Peter J Hudson, and Shweta Bansal. Unraveling the disease consequences and mechanisms of modular structure in animal social networks. *Proceedings of the National Academy of Sciences*, 114(16):4165–4170, 2017.

[8] Marcel Salathé and James H Jones. Dynamics and control of diseases in networks with community structure. *PLoS computational biology*, 6(4): e1000736, 2010.

[9] Julian C Evans, David J Hodgson, Neeltje J Boogert, and Matthew J Silk. Group size and modularity interact to shape the spread of infection and information through animal societies. *bioRxiv*, 2021.

[10] Matteo Cinelli, Gianmarco De Francisci Morales, Alessandro Galeazzi, Walter Quattrociocchi, and Michele Starnini. The echo chamber effect on social media. *Proceedings of the National Academy of Sciences*, 118(9), 2021.

[11] Anja Leppin and Arja R Aro. Risk perceptions related to sars and avian influenza: theoretical foundations of current empirical research. *International journal of behavioral medicine*, 16(1):7–29, 2009.

[12] Mark Schaller. The behavioural immune system and the psychology of human sociality. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366(1583):3418–3426, 2011.

[13] Mark Schaller. Parasites, behavioral defenses, and the social psychological mechanisms through which cultures are evoked. *Psychological Inquiry*, 17 (2):96–101, 2006.

[14] Trevor Collier, Stephen J Cotten, and Justin Roush. Risk aversion, guilt, and pandemic behavior. *Guilt, and Pandemic Behavior (December 1, 2020)*, 2020.

[15] Roy F Baumeister and Mark R Leary. The need to belong: desire for interpersonal attachments as a fundamental human motivation. *Psychological bulletin*, 117(3):497, 1995.

[16] Louise C Hawkley and John T Cacioppo. Loneliness matters: A theoretical and empirical review of consequences and mechanisms. *Annals of behavioral medicine*, 40(2):218–227, 2010.

[17] Naomi I Eisenberger and Matthew D Lieberman. Why rejection hurts: a common neural alarm system for physical and social pain. *Trends in cognitive sciences*, 8(7):294–300, 2004.

[18] Natsumi Sawada, Emilie Auger, and John E Lydon. Activation of the behavioral immune system: Putting the brakes on affiliation. *Personality and Social Psychology Bulletin*, 44(2):224–237, 2018.

[19] Donald F Sacco, Steven G Young, and Kurt Hugenberg. Balancing competing motives: Adaptive trade-offs are necessary to satisfy disease avoidance and interpersonal affiliation goals. *Personality and Social Psychology Bulletin*, 40(12):1611–1623, 2014.

[20] Mohitosh Biswas, Shawonur Rahaman, Tapash Kumar Biswas, Zahirul Haque, and Baharudin Ibrahim. Effects of sex, age and comorbidities on the risk of infection and death associated with covid-19: a meta-analysis of 47807 confirmed cases. *Age and Comorbidities on the Risk of Infection and Death Associated with COVID-19: A Meta-Analysis of*, 47807, 2020.

[21] Silvia Maggini, Adeline Pierre, and Philip C Calder. Immune function and micronutrient requirements change over the life course. *Nutrients*, 10(10): 1531, 2018.

[22] Paula Caligiuri and Ibraiz Tarique. Dynamic cross-cultural competencies and global leadership effectiveness. *Journal of world Business*, 47(4):612–622, 2012.

[23] Marco Lauriola, Irwin P Levin, and Stephanie S Hart. Common and distinct factors in decision making under ambiguity and risk: A psychometric study of individual differences. *Organizational Behavior and Human Decision Processes*, 104(2):130–149, 2007.

[24] S Petrocchi, P Iannello, G Ongaro, A Antonietti, and G Pravettoni. The interplay between risk and protective factors during the initial height of the covid-19 crisis in italy: The role of risk aversion and intolerance of ambiguity on distress. *Current Psychology*, pages 1–12, 2021.

[25] D. Caroline Blanchard, Guy Griebel, Roger Pobbe, and Robert J Blanchard. Risk assessment as an evolved threat detection and analysis process. *Neuroscience & Biobehavioral Reviews*, 35(4):991–998, 2011.

[26] Brian R Smith and Daniel T Blumstein. Fitness consequences of personality: a meta-analysis. *Behavioral Ecology*, 19(2):448–455, 2008.

[27] Kees van Oers, Piet J Drent, Piet de Goede, and Arie J van Noordwijk. Realized heritability and repeatability of risk-taking behaviour in relation to avian personalities. *Proceedings of the Royal Society of London B: Biological Sciences*, 271(1534):65–73, 2004.

[28] David Cesarini, Christopher T Dawes, Magnus Johannesson, Paul Lichtenstein, and Björn Wallace. Genetic variation in preferences for giving and risk taking. *The Quarterly Journal of Economics*, 124(2):809–842, 2009.

[29] M Zuckerman and D M Kuhlman. Personality and risk-taking: Common biosocial factors. *Journal of Personality*, 68(6):999–1029, 2000.

[30] Nigel Nicholson, Emma Soane, Mark Fenton-O'Creevy, and Paul Willman. Personality and domain-specific risk taking. *Journal of Risk Research*, 8 (2):157–176, 2005.

[31] Jay J Van Bavel, Katherine Baicker, Paulo S Boggio, Valerio Capraro, Aleksandra Cichocka, Mina Cikara, Molly J Crockett, Alia J Crum, Karen M Douglas, James N Druckman, et al. Using social and behavioural science to support covid-19 pandemic response. *Nature human behaviour*, 4(5): 460–471, 2020.

[32] John Paget. The influenza pandemic and europe: the social impact and public health. *Italian Journal of Public Health*, 6:257, 2009.

[33] Eli P Fenichel. Economic considerations for social distancing and behavioral based policies during an epidemic. *Journal of health economics*, 32(2):440–451, 2013.

[34] Clive Bell and Hans Gersbach. The macroeconomics of targeting: the case of an enduring epidemic. *Journal of Health Economics*, 28(1):54–72, 2009.

[35] Maarten H. W. van Zalk, Steffen Nestler, Katharina Geukes, Roos Hutteman, and Mitja D. Back. The codevelopment of extraversion and friendships: Bonding and behavioral interaction mechanisms in friendship networks. *Journal of personality and social psychology*, 118(6):1269–1290, 2020.

[36] Sergio Currarini and Friederike Mengel. Identity, homophily and in-group bias. *European Economic Review*, 90:40–55, 2016.

[37] Jaromír Kovářík and Marco J Van der Leij. Risk aversion and social networks. *Review of Network Economics*, 13(2):121–155, 2014.

[38] David V. Foster, Jacob G. Foster, Peter Grassberger, and Maya Paczuski. Clustering drives assortativity and community structure in ensembles of networks. *Phys. Rev. E*, 84:066117, 2011.

[39] Sergio Arregui, Alberto Aleta, Joaquín Sanz, and Yamir Moreno. Projecting social contact matrices to different demographic structures. *PLoS computational biology*, 14(12):e1006638, 2018.

[40] Roy M Anderson and Robert M May. *Infectious diseases of humans: dynamics and control*. Oxford university press, 1992.

[41] Adam Catching, Sara Capponi, Ming Te Yeh, Simone Bianco, and Raul Andino. Examining the interplay between face mask usage, asymptomatic transmission, and social distancing on the spread of covid-19. *Scientific reports*, 11(1):1–11, 2021.

[42] Armanda Cetrulo, Dario Guarascio, and Maria Enrica Virgillito. The privilege of working from home at the time of social distancing. *Intereconomics*, 55:142–147, 2020.

[43] Andrea K Townsend, Dana M Hawley, Jessica F Stephenson, and Keelah EG Williams. Emerging infectious disease and the challenges of social distancing in human and non-human animals. *Proceedings of the Royal Society B*, 287(1932):20201039, 2020.

[44] See supplemental material at [url will be inserted by publisher] for a discussion of the parameter constraints of the model (appendix 1), mathematical analysis of the model behaviors (appendix 2), and investigation of the effects of varying σ/ρ (appendix 3).

[45] Andrzej Kulig, Stanislaw Drozdz, Jaroslaw Kwapien, and Pawel Oswiecimka. Modelling subtle growth of linguistic networks. *Phys. Rev. E*, 91: 032810, 2015.

[46] Mark EJ Newman. Who is the best connected scientist? a study of scientific coauthorship networks. *Phys. Rev. E*, 64(016131), 2001.

[47] Kim R Hill, Robert S Walker, Miran Božičević, James Eder, Thomas Headland, Barry Hewlett, A Magdalena Hurtado, Frank Marlowe, Polly Wiessner, and Brian Wood. Co-residence patterns in hunter-gatherer societies show unique human social structure. *science*, 331(6022):1286–1289, 2011.

[48] Kenny Smith, Henry Brighton, and Simon Kirby. Complex systems in language evolution: the cultural emergence of compositional structure. *Advances in complex systems*, 6(04):537–558, 2003.

[49] Karlo Hock, Kah Loon Ng, and Nina H Fefferman. Systems approach to studying animal sociality: Individual position versus group organization in dynamic social network models. *PLoS One*, 5(12):e15789, 2010.

[50] Karlo Hock and Nina H Fefferman. Social organization patterns can lower disease risk without associated disease avoidance or immunity. *Ecological Complexity*, 12:34–42, 2012.

[51] Oyita Udiani and Nina H Fefferman. How disease constrains the evolution of social systems. *Proceedings of the Royal Society B*, 287(1932):20201284, 2020.

[52] Valéria Romano, Cédric Sueur, and Andrew JJ MacIntosh. The tradeoff between information and pathogen transmission in animal societies. *Oikos*, 2021.

[53] Mina Youssef and Caterina Scoglio. Mitigation of epidemics in contact networks through optimal contact adaptation. *Mathematical biosciences and engineering: MBE*, 10(4):1227, 2013.

[54] Julia C Buck and Sara B Weinstein. The ecological consequences of a pandemic. *Biology Letters*, 16(11):20200641, 2020.

[55] Barbara Critchlow Leigh. Reasons for having and avoiding sex: Gender, sexual orientation, and relationship to sexual behavior. *Journal of Sex Research*, 26(2):199–209, 1989.

[56] Jonathan M Ellen, Nancy Adler, Jill E Gurvey, Miranda BV Dunlop, Susan G Millstein, and Jeanne Tschann. Improving predictions of condom behavioral intentions with partner-specific measures of risk perception. *Journal of Applied Social Psychology*, 32(3):648–663, 2002.

[57] Pascaline Dupas. Do teenagers respond to hiv risk information? evidence from a field experiment in kenya. *American Economic Journal: Applied Economics*, 3(1):1–34, 2011.

[58] Nina H Fefferman, Matthew J Silk, Dana K Pasquale, and James Moody. Homophily in risk and behavior complicate understanding the covid-19 epidemic curve. *medRxiv*, 2021.

[59] Matthew O Jackson and Dunia López-Pintado. Diffusion and contagion in networks with heterogeneous agents and homophily. *Network Science*, 1 (1):49–67, 2013.

[60] Charles L Nunn, Ferenc Jordán, Collin M McCabe, Jennifer L Verdolin, and Jennifer H Fewell. Infectious disease and group size: more than just a numbers game. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1669):20140111, 2015.

[61] Juan Manuel Falomir-Pichastor, Daniel Muñoz-Rojas, Federica Invernizzi, and Gabriel Mugny. Perceived in-group threat as a factor moderating the influence of in-group norms on discrimination against foreigners. *European Journal of Social Psychology*, 34(2):135–153, 2004.

[62] Michael A Zarate, Berenice Garcia, Azenett A Garza, and Robert T Hitlan. Cultural threat and perceived realistic group conflict as dual predictors of prejudice. *Journal of experimental social psychology*, 40(1):99–105, 2004.

[63] Matthew O Jackson and Evan Storms. Behavioral communities and the atomic structure of networks. *Available at SSRN 3049748*, 2019.

[64] DAVE ELDER-VASS. For emergence: refining archer's account of social structure. *Journal for the Theory of Social Behaviour*, 37(1):25–44, 2007.

[65] Ranjay Gulati, Maxim Sytch, and Adam Tatarynowicz. The rise and fall of small worlds: Exploring the dynamics of social structure. *Organization Science*, 23(2):449–471, 2012.

[66] Mark EJ Newman, Duncan J Watts, and Steven H Strogatz. Random graph models of social networks. *Proceedings of the national academy of sciences*, 99(suppl 1):2566–2572, 2002.